Low–dose imaging for particle therapy with fluence–modulated proton computed tomography

Jannis Ivo Dickmann

München 2021

LOW–DOSE IMAGING FOR PARTICLE THERAPY
WITH FLUENCE–MODULATED
PROTON COMPUTED TOMOGRAPHY

NIEDRIGDOSISBILDGEBUNG FÜR DIE
PARTIKELTHERAPIE MIT FLUENZ–MODULIERTER
PROTONEN–COMPUTERTOMOGRAPHIE

Dissertation

an der Fakultät für Physik
der Ludwig–Maximilians–Universität München

vorgelegt von

Jannis Ivo Dickmann

aus Hanau

München, den 8. April 2021

ABSTRACT

Low–dose imaging for particle therapy with fluence–modulated proton computed tomography

PARTICLE THERAPY for the curative treatment of tumors with energetic charged particles allows for the precise deposition of a therapeutic dose in the cancerous tissue while sparing surrounding healthy tissue. This exploits the increased dose deposition of charged particles, typically protons or carbon ions, at the end of their range. The treatment planning with highly conformal doses, however, requires a precise volumetric knowledge of the patients stopping power relative to water (RSP). In current clinical practice, such images are acquired using x–ray computed tomography, which measures the interaction of photons with matter and is subsequently converted to RSP. This conversion leads to errors, which need to be considered as additional margins around the tumor and necessarily lead to a higher dose to healthy tissue. An imaging modality suggested to reduce such errors is proton computed tomography (pCT), which directly determines the RSP through measurements of the energy loss of protons in the patient.

Within this work, methods for dynamic modulation of the imaging fluence field have been established to reduce the imaging dose required for pCT acquisitions. With fluence–modulated pCT (FMpCT), the image is split into a region–of–interest (ROI), where good image quality is needed, and a non–ROI region, which is not used for treatment planning. In the context of particle therapy, the ROI covers the vicinity of the therapeutic beam. Outside of the ROI, imaging noise can be increased and imaging dose reduced. The calculation of modulated fluence patterns requires a forward model, that predicts the expected image noise for a given fluence setting. Such a forward model was realized using a Monte Carlo model of a specific pCT scanner and validated against experimental data. This allowed to carefully disentangle single contributions to pCT image noise, which was found to strongly depend on the heterogeneity of the object. Using the forward model, two FMpCT optimization algorithms were proposed: a simpler one, that only takes into account image variance, and a more sophisticated one, that considers both image noise and imaging dose objectives. The FMpCT scans were realized both in simulations and in experiments at the Chicago proton center using small fluence–modulated pencil beams, for which an interface to the control system needed to be established. Simulations were performed using anonymized x–ray CTs of patients undergoing photon therapy, and corresponding proton therapy treatment plans served for the definition of the imaging ROI. The imaging dose outside of the ROI could be reduced by 74 % compared to scans at uniform imaging fluence and at the same peak noise level. The imaging dose to critical structures such as the eyes could further be pushed down using the optimization and dose savings up to 87 % were achieved while maintaining the accuracy for treatment plan optimization on FMpCT images.

In addition, two methods for artifact reduction with un–modulated pCT scans were developed, with one directly addressing the physical reason for artifacts by merging two datasets at different incident energies. The second artifact correction method was purely empirical and made no assumption on the origin of image artifacts. It used a scan of a custom–built phantom with known RSP and allowed to almost halve the mean absolute RSP error of a prototype pCT scanner by 46 %.

In this work, the development and experimental realization of optimized FMpCT scans together with the improved accuracy of pCT opened an interesting perspective: towards adaptive particle therapy with daily image guidance without accumulation of excessive patient doses in healthy tissue.

ZUSAMMENFASSUNG

Niedrigdosisbildgebung für die Partikeltherapie mit fluenz–modulierter Protonen–Computertomographie

PARTIKELTHERAPIE zur kurativen Behandlung von Tumoren mit beschleunigten Ionen ermöglicht die präzise Verabreichung der therapeutischen Strahlendosis im Krebsgewebe während umliegendes, gesundes Gewebe geschont wird. Dabei nutzt man die erhöhte Dosisabgabe von geladenen Teilchen, typischerweise Protonen oder Kohlenstoffionen, am Ende ihrer Reichweite. Eine Bestrahlungsplanung mit hoch–konformen Dosisverteilungen erfordert jedoch eine präzise volumetrische Bildgebung des relativen Bremsvermögens (RSP) des Patienten. In der derzeitigen klinischen Praxis werden solche Schnittbilder mittels Röntgen–Computertomographie erstellt, und nachfolgend die dabei gemessene Interaktion von Photonen mit Materie in RSP umgerechnet. Diese Umrechnung führt zu Fehlern, welche bei der Bestrahlungsplanung als zusätzliches Toleranzvolumen berücksichtigt werden müssen und unumgänglich zu erhöhter Strahlendosis im gesunden Gewebe führen. Ein Bildgebungsverfahren, welches diese Fehler reduzieren kann, ist die Protonen–Computertomographie (pCT), bei der das RSP direkt durch Messung des Energieverlusts von Protonen im Patienten bestimmt wird.

In dieser Arbeit wurden Methoden zur dynamischen Modulation der Bildgebungsfluenz etabliert, mit denen die Dosis von pCT–Aufnahmen reduziert werden kann. Für fluenz–modulierte pCT (FMpCT) wird das Bild eingeteilt in eine Zielregion (ROI), in der hohe Bildqualität benötigt wird, sowie das restliche Volumen, das zur weiteren Bestrahlungsplanung nicht benötigt wird. Im Kontext der Partikeltherapie entspricht die ROI einer Umgebung um den therapeutischen Strahl. Außerhalb der ROI kann das Bildrauschen erhöht und die Bildgebungsdosis reduziert werden. Die Berechnung von modulierten Fluenzverteilungen wird ermöglicht durch ein Modell des für eine gegebene Fluenzmodulation zu erwartenden Bildrauschens. Ein solches Modell wurde mit Hilfe einer Monte Carlo Simulation eines spezifischen pCT Scanners erstellt und mit experimentellen Daten validiert. Dadurch war es außerdem möglich, einzelne physikalische Beiträge zum pCT Bildrauschen zu isolieren. Für den Rauschwert spielte dabei die Heterogenität des Bildgebungsobjektes eine

entscheidende Rolle. Basierend auf dem Rauschmodell wurden zwei FMpCT Optimierungsalgorithmen entwickelt: ein vereinfachter, bei dem lediglich das Bildrauschen berücksichtigt wird, und ein weiterentwickelter, mit dem sowohl gewünschte Bildrausch– als auch Bildgebungsdosis–Verteilungen erzielt werden können. FMpCT Aufnahmen wurden durch Modulation der Fluenz kleiner Strahlbündel sowohl in Simulationen als auch in Experimenten am Protonenzentrum in Chicago realisiert. Für die Experimente musste eine Schnittstelle zum Kontrollsystem des Beschleunigers etabliert werden. In Simulationen mit anonymisierten Röntgen-CTs von Photonentherapie–Patienten wurde die Bildgebungs–ROI basierend auf entsprechenden Protonen–Bestrahlungsplänen definiert. Die Bildgebungsdosis außerhalb der ROI wurde dabei um 74 % reduziert im Vergleich zu Aufnahmen bei gleichmäßiger Fluenz und bei gleichem maximalen Rauschwert. Die Bildgebungsdosis in kritischen Strukturen wie den Augen konnte dabei im Rahmen der Optimierung weiter verringert werden und Reduktionen von bis zu 87 % waren möglich. Dabei wurde die Genauigkeit der Bestrahlungsplanung nicht beeinträchtigt.
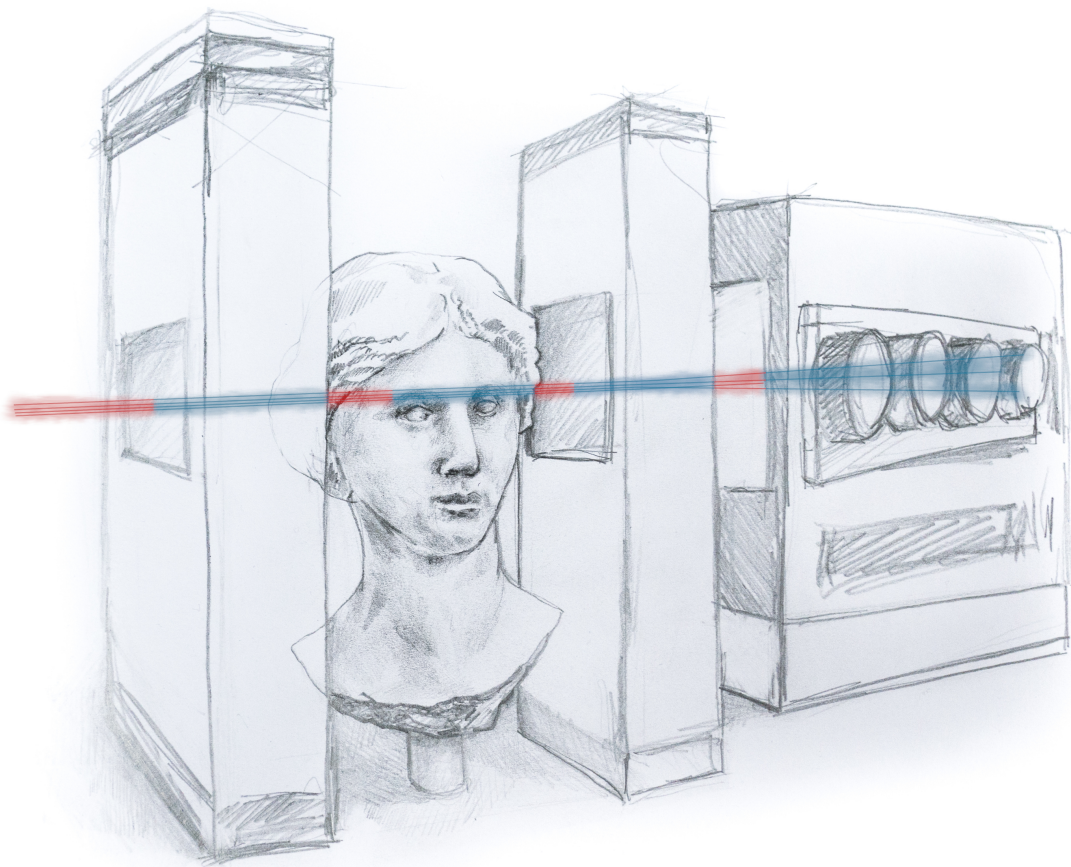
Zusätzlich wurden zwei Methoden zur Artefaktkorrektur in nicht–modulierten pCT Aufnahmen entwickelt. Dabei berücksichtigte eine Methode direkt die physikalische Ursache von Bildgebungsartefakten und führte selektiv zwei Datensätze bei unterschiedlichen Protonenenergien zusammen. Die zweite Artefaktkorrektur war vollkommen empirisch und machte keine Annahmen über die Ursache der Artefakte. Dabei wurde eine Aufnahme eines eigens hergestellten Phantoms mit bekanntem RSP genutzt. Mit der Methode wurde der mittlere absolute RSP–Fehler eines Prototypen–pCT–Scanners um 47 % reduziert.

Diese Arbeit hat mit der Entwicklung und experimentellen Realisierung von optimierten FMpCT Aufnahmen zusammen mit der verbesserten Genauigkeit von pCT eine interessante Perspektive eröffnet: hin zur adaptiven Partikeltherapie mit täglicher Bildgebung unter Vermeidung übermäßiger Dosisbelastung in gesundem Gewebe.

Nothing in life is to be feared, it is only to be understood.
Now is the time to understand more so we may fear less.

— Marie Skłodowska Curie

CONTENTS

Abbreviations: CT – computed tomography, FMpCT – fluence–modulated proton CT

LIST OF ABBREVIATIONS

Notation	Description
ADC	analog–to–digital converter
C	center of a grayscale display
CBCT	cone–beam (x–ray) computed tomography
CT	(x–ray) computed tomography
CTV	clinical target volume
DECT	dual energy (x–ray) computed tomography
EMpCT	energy–modulated proton computed tomography
FDK	Feldkamp–Davis–Kress algorithm for image reconstruction
FDK–DDB	Feldkamp–Davis–Kress image reconstruction with distance–driven binning
FMpCT	fluence–modulated proton computed tomography
FWHM	full width at half maximum
GTV	gross tumor volume
IQR	inter–quartile range
LDPE	low–density polyethylene
LINAC	linear accelerator
MAPE	mean absolute percentage error
MCS	multiple Coulomb scattering
MLP	most likely path
MRI	magnetic resonance imaging
OAR	organ at risk
PBS	pencil beam scanning
pCT	proton computed tomography
PMMA	polymethylmethacrylate/acrylic

Notation	Description
PMP	polymethylpentene
PTV	planning target volume
RMS	root mean square (error)
ROI	region–of–interest
RSP	stopping power relative to water
TPS	treatment planning system
W	width of a grayscale display
WEPL	water–equivalent path length

Throughout this work the abbreviations *C* and *W* are used with plots to define the center of a grayscale display (C) and the width of a grayscale display (W). Thereby, the center indicates the value of the central gray level and the width defines the range from black to white.

1 | INTRODUCTION

THIS FIRST CHAPTER shall introduce a general framework within which this thesis operates. The challenges and advantages of modern image–guided particle therapy for the treatment of tumors are discussed. It is laid out how proton computed tomography, the main imaging modality used in this thesis, may help improve the accuracy of cancer therapy with charged particles while decreasing the radiation dose administered to the patient. Eventually, the use of fluence–modulation for proton computed tomography is motivated and the aim and scope of this thesis are defined.

1.1 PARTICLE THERAPY

Healthy cells in the human body continuously reproduce at a controlled rate and with an efficient system of detection and correction of errors. A malfunction of these processes, initiated by external factors such as radiation or smoking, or due to systemic factors or viruses, can cause an uncontrolled and malign proliferation of cells, which in general terms is called cancer [1, 2] and can occur in any organ of the body. The rapid growth of cells is initially localized in a primary tumor. However, systems to detect cell reproduction errors become insufficient inside the tumor causing cell mutations which can eventually invade healthy tissues or spread to other organs and form metastases. Cancer can have both local effects due to the expansion of the tumor, as well as various systemic effects such as a loss of weight or a disruption of the hormonal system. If untreated, the disease is fatal, in most cases not due to the primary tumor, but due to disruption of organs from metastases. Cancer is the second leading cause of death globally according to the World Health Organization and over–proportionally affects low– and middle–income countries [3].

To prevent a fatal outcome, early diagnosis and efficient treatment are mandatory. Primary tumors and metastases can be mainly treated in three ways [4]: (1) removal of the primary tumor by surgery, (2) systemic treatment with chemotherapy or other therapies affecting the patient's immune or hormone system, and (3) eradication of tumor cells using radiation therapy. This work will only focus on the treatment with radiation therapy assuming that the other two options are used supplementally, or are exhausted, or not possible. While radiation can be a cause of cancer itself, it can also be used to treat tumors. Due to their insufficient repair mechanisms, tumor cells are often more susceptible to cell death induced by radiation [5]. This allows to target cancerous cells by administering radiation dose either

internally by bringing a radioactive source close to the tumor (brachytherapy) or by exposing the patient to an external beam of ionizing radiation (external beam radiation therapy). The most common sources of ionizing radiation for external beam radiation therapy are x–ray photons, electrons, protons and carbon ions. Due to their relatively low cost and good clinical outcomes, megavoltage x–ray beams generated by compact linear accelerators (LINACs) [6] are most frequently used in oncological clinics today. The photon beams are administered from various angles while focusing the fluence on the intersection with the tumor using a multi–leaf collimator [7]. Thereby, the therapeutic dose can be concentrated in the tumor volume and dose to healthy tissue is reduced [8]. This opens a therapeutic window within which the tumor cells are eradicated and do not regrow while healthy tissue is spared and can regenerate [9], ideally without severe side effects. The relevant physical interaction processes of photons that lead to that dose deposition are described in section 2.2.

Treatment of cancer using protons and carbon ions, or in short particle therapy [11, 12], is less frequently used due to the higher cost of machines accelerating ions compared to LINACs. However, particle therapy shows more favorable clinical outcomes for many types of cancer [13–15], due to the different mode of dose deposition, which is described in detail in section 2.1. While the dose deposition of a single photon beam exponentially decreases with depth, charged particles and in particular protons exhibit a so–called Bragg peak within which the depth–dose curve rapidly increases when the particle's energy decreases, eventually bringing the particle to a halt. This effect can be exploited to carefully tailor the proton's incident energy such that the Bragg peak — and with it the maximum dose deposition — occurs inside the tumor. Using many particle beams with different locations, energies and intensities allows to uniformly cover the tumor volume with an approach called pencil beam scanning (PBS). Since no dose is deposited after the stopping point of the particle, cancer treatment using charged particles allows to better spare healthy tissue and the treatment is generally more tolerable with less severe comorbidities [13], thereby
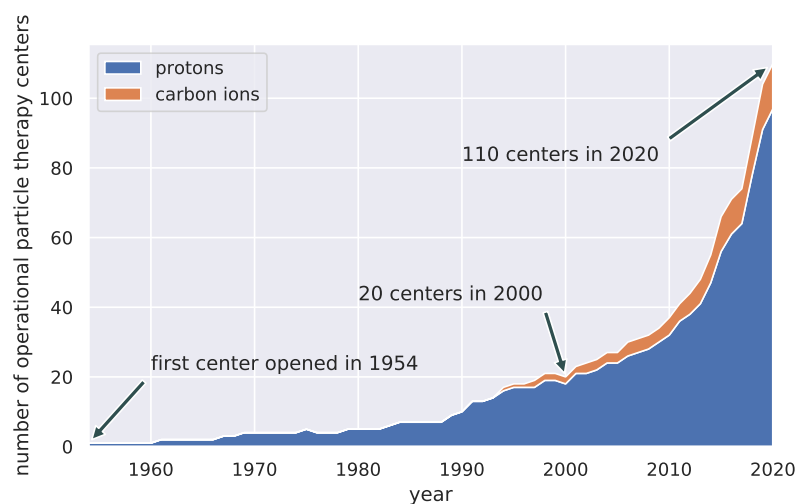
Figure 1.1: Number of operational particle therapy centers worldwide from 1954 to today with a rapid growth within the last decade. *Data source:* PTCOG and Levin et al. [10].

widening the above–mentioned therapeutic window. Additionally, therapy with charged particles also has biological advantages [16, 17]. While the vast majority of patients today are still treated with photons, the number of facilities offering proton or carbon ion treatment has grown dramatically within the last decade [10, 18], as shown in fig. 1.1. It can also be seen that carbon ion therapy plays a minor role compared to proton therapy, which again is mostly due to economic rather than clinical considerations [19]. While this ratio is different in Germany — with two centers in Heidelberg and Marburg offering carbon ion and proton therapy and three centers in Berlin, Essen and Dresden offering only proton therapy — particle therapy with protons is still more common. Therefore, in this work, only proton therapy is considered specifically.

1.2 TREATMENT WORKFLOW OF IMAGE–GUIDED PARTICLE THERAPY

The benefit of particle therapy, which is the precise targeting of the tumor using Bragg peaks, is also its biggest challenge. Even small changes in the positioning of the patient or the internal anatomy can cause Bragg peaks to fall outside of the tumor volume causing damage to healthy tissue while leaving parts of the tumor untreated and allowing it to regrow. In this worst case scenario the patient would suffer from severe side effects while the tumor growth could not be stopped. To avoid this, and to ensure a safe and precise delivery of the therapeutic dose, particle therapy relies heavily on frequent imaging [20, 21] using various imaging modalities both before [22, 23] and during the treatment [24, 25]. Specifically, during the treatment workflow of image–guided particle therapy, which is illustrated in fig. 1.2, imaging is performed during the following steps and using various imaging modalities.
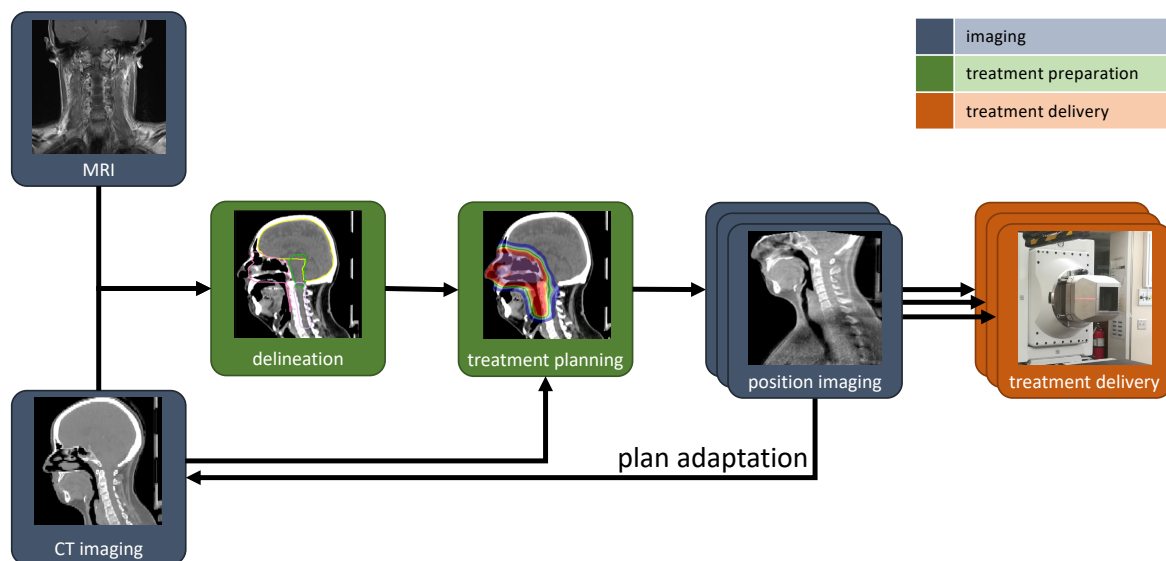

Figure 1.2: Illustration of the treatment workflow of modern image–guided particle therapy.

DELINEATION   Before any treatment planning can begin, the tumor volume and critical organ structures need to be delineated [26, 27]. This task is performed based on computed tomography (CT) and ideally also magnetic resonance imaging (MRI). While the good soft tissue contrast of MRI allows to identify the tumor volume more easily, CT images ensure the correct geometry which can be distorted with MRI. For this purpose, MRI is (deformably) registered to the CT. In some cases, delineation can be complemented by nuclear imaging, in particular positron emission tomography (PET) or single photon emission CT (SPECT). Apart from the gross tumor volume (GTV), which is the visible extent of the tumor, the clinical target volume (CTV) and planning target volume (PTV) are delineated. The CTV extends the GTV to ensure that all suspected microscopic malignant diseases are covered that are not visible with imaging. The PTV contains the CTV and an additional geometric margin to ensure that the CTV is irradiated even under delivery uncertainties and considering errors in the calculated particle range due to imaging inaccuracies. For definition of the PTV, four–dimensional (time resolved) MRI or CT imaging can be used [28]. Moreover, organs at risk (OARs) such as the brainstem are delineated to monitor and reduce treatment dose to critical structures.

TREATMENT PLANNING   When tumor and organ structures are delineated, a medical doctor prescribes a therapeutic dose to the PTV which can be up to 70 Gy for definitive treatment of the tumor. Additionally, between one and typically up to four beam directions are defined from which the tumor can be reached without crossing critical structures or very heterogeneous tissue [17, 29]. Constraints are set to minimize dose to healthy tissue and in particular to OARs while ensuring that the prescription in the PTV is met. This optimization task is performed using a treatment planning system (TPS) operated by a medical physicist. To date, the TPS solely relies on CT imaging for calculation of the particles' ranges. To achieve this, the CT image is passed through a calibration curve [30] to convert the linear attenuation coefficient of x–ray photons to the stopping power of charged particles. This process involves additional uncertainties which need to be considered in the PTV margin.

PATIENT POSITIONING   To deliver the therapeutic dose, the patient needs to be positioned on the treatment couch exactly as during the acquisition of the treatment planning CT. For tumors of the head and neck region, this is often aided by using a personalized thermoplastic mask that can be attached to the treatment couch. After an initial positioning of the patient, a set of two–dimensional (planar) radiographs is taken and compared to digitally reconstructed radiographs of the planning CT. Modern centers additionally use volumetric imaging using cone–beam computed tomography (CBCT) to ensure a correct patient positioning [24].

PLAN ADAPTATION   If major anatomic changes, which are likely not to be covered by the PTV margins, are detected during CBCT imaging prior to treatment or during a routine MRI or CT scan in between treatment sessions, the treatment needs to be interrupted and treatment planning needs to be performed again. Since CBCT imaging does not (yet) offer sufficient image quality for treatment planning, a CT scan would need to be additionally performed in this case. Due to good positioning using thermoplastic masks, generally large PTV margins and infrequent MRI or CT scans in between scans, a plan adaptation is rare

in current clinical practice. However, patients may benefit from more frequent volumetric imaging and daily plan adaptation [25, 31], as indicated by the recent clinical availability of MR–LINACs [32, 33], which allow MRI in treatment position right before or after the treatment delivery, and online plan adaptation.

––––––––––

In conclusion, modern particle therapy would not be possible without the extensive use of imaging before and during the treatment. At the same time, improvements in imaging are key to increase treatment accuracy and safety [34]. Specifically, PTV margins [26] could be reduced if the particle range calculation becomes more accurate due to improved imaging, or if patient positioning [35] is more accurate and anatomical changes are detected reliably due to frequent imaging in treatment position. In both cases, margin reductions would directly translate into a reduced dose to healthy tissue which improves treatment tolerability and further widens the therapeutic window of particle therapy.

1.3 PROTON COMPUTED TOMOGRAPHY

An imaging modality suggested to improve accuracy of treatment planning is proton computed tomography (pCT), which was initially proposed in 1963 by Cormack [36] and later realized by Hanson et al. [37]. Instead of measuring the linear attenuation coefficient of photons with x–ray CT and then converting it to proton stopping power, a process which involves conversion errors, pCT directly measures the patient's stopping power relative to water (RSP) by determining the energy loss of protons traversing the patient. The RSP can then be used in the TPS to calculate the range of protons in the patient. The detection principle and the required technical instrumentation of pCT are described in detail in sections 3.1 and 3.2. The general idea of pCT is to make use of the fact that protons continuously lose energy in small steps when traversing a medium [38, 39]. For pCT acquisitions, the initial proton energy is fixed and increased from what is typically used for treatment (where protons stop inside the tumor), to typically $200\,MeV$ so that protons can traverse at least an average adult head and exit with a reduced energy. A detector then measures the protons' residual energy, which can be converted to a water–equivalent path length (WEPL). The WEPL is a line integral through RSP along the path of the proton and as such can be used to reconstruct a volumetric image of the patient using dedicated reconstruction algorithms such as the one described in section 3.5.3.

Early prototype pCT scanners were shown to be as accurate as state–of–the–art clinical x–ray CT scanners [40, 41], and may outperform the standard of care with future technology improvements [42, 43]. Apart from the improved accuracy, pCT scanners also administer less radiological dose to the patient [44] compared to getting the equivalent information using x–ray CT. With this, they could allow for frequent imaging prior to every treatment session [34], which, as discussed before, could allow for daily adaptive particle therapy [25, 31]. Moreover, pCT scans employ protons through the treatment system and are therefore, by design, acquired in treatment position and even using the same coordinate system. This is beneficial to avoid a positioning mismatch between scans acquired in two different rooms.

Consequently, pCT could supplement or replace many of the imaging modalities in section 1.2 and fig. 1.2. It cannot replace MRI for structure delineation due to its low soft tissue contrast, but it could be used instead of x–ray CT as a basis for treatment planning. CBCT imaging could also be replaced by pCT allowing for a precise patient positioning as well as an immediate plan adaptation.

1.4 AIMS & SCOPE

The main focus of this work was to further reduce imaging dose of pCT by employing modulated fluence fields during the tomographic acquisition. Fluence, and with it imaging dose, is inversely related to noise in the image. The goal of fluence–modulated proton computed tomography (FMpCT) [45, 46] is to achieve a low noise level only where this is needed, namely inside a region–of–interest (ROI). In the context of particle therapy, this ROI will be the vicinity of the treatment beam path, which typically only covers a fraction of the imaged volume. In this work the ROI will be defined as the volume receiving at least 10 % of the treatment dose. Inside the ROI, the full radiological information will be maintained, but outside of the ROI, fluence can be reduced and with it imaging dose. The imaging dose saving outside of the ROI is particularly meaningful, since particle therapy, by using Bragg peaks, has very little treatment dose outside of the treatment field. Or in other words: FMpCT may allow to image the patient more frequently without compromising the low–dose benefit of particle therapy. A major challenge of this work was to solve the inverse problem of calculating imaging fluences that would result in a desired image noise and imaging dose distribution. This was achieved using two different optimization algorithms that needed to account for the noise characteristics of a prototype pCT scanner to allow for subsequent experimental acquisitions.

Within this work the following objectives were pursued aiming to enable and experimentally realize optimized FMpCT acquisitions and, at the same time, improve accuracy of pCT in general.

OBJECTIVE I *To understand and model pCT image noise to enable FMpCT.* Controlling image noise as part of a fluence optimization requires a complete understanding and the ability to predict image noise for a given fluence setting. Monte Carlo simulations were used to predict image noise contributions for a specific prototype pCT scanner and compared to experimental scans of the same objects. In particular the impact of object heterogeneities was studied.

OBJECTIVE II *To develop optimization algorithms for FMpCT and demonstrate their feasibility using Monte Carlo simulations.* With a detailed forward model of image noise in pCT, the task of finding optimal fluence settings achieving a desired image noise distribution becomes an inverse problem, which can be solved using optimization. The performance of the optimizer was tested in Monte Carlo simulations which also allowed to estimate potential dose savings.

OBJECTIVE III   *To show the experimental feasibility of FMpCT using a prototype pCT scanner.* Applying optimized fluence patterns in an experimental scan required to establish an interface with the beam delivery system of a proton therapy facility. Resulting experimental image noise maps were compared to the predictions from Monte Carlo simulations as well as to the prescribed image noise.

OBJECTIVE IV   *To investigate the clinical applicability of FMpCT by using patient data and develop a fast and flexible optimizer with clinically relevant optimization targets.* Images of pediatric patients undergoing cancer therapy were imported to a Monte Carlo pCT simulation to apply FMpCT to plausible scenarios. For this, an improved optimization algorithm was developed that does not only take into account image noise, but also imaging dose and in particular imaging dose to critical structures such as OARs.

OBJECTIVE V   *To investigate and improve the accuracy of a prototype pCT scanner.* Besides a reduction of imaging dose, which is the main topic of this work, pCT could also benefit from an improvement of image accuracy. In conjunction with the work on FMpCT, the accuracy of pCT images in terms of the RSP was investigated and methods applied to improve it.

––––––––––

In chapter 2, this thesis will introduce the fundamental physical processes, the understanding of which allows us to treat patients with particle therapy and to perform x–ray CT and pCT scans. In chapter 3, the concept of pCT imaging is detailed together with the necessary technical instrumentation and reconstruction algorithms. Since this thesis makes frequent use of Monte Carlo simulations, the computational models as well as the use of patient data in simulations are described in chapter 4. After those fundamental chapters, the first objective, the understanding of image noise, is discussed in chapter 5. In chapter 6, an optimization algorithm for FMpCT is proposed and its performance assessed in a simulation study, thereby covering the second objective. The third objective, the experimental realization of FMpCT, is pursued in chapter 7. Based on the conclusions of chapters 6 and 7, the optimization algorithm is re–designed and a novel optimizer with joint dose and variance objectives is presented in chapter 8. It is applied to patient data and to clinically relevant optimization targets, which discusses the fourth objective. Finally, chapter 9 puts forward the fifth objective trying to improve image accuracy of pCT.

2 | INTERACTION OF PROTONS AND PHOTONS WITH MATTER

T HE UNDERSTANDING of the interaction of charged particles with matter is a key component to allow for their usage to treat cancer patients. This chapter will summarize the fundamental physical interactions that are relevant for proton therapy, but that also enabled the development of pCT scanners. It is mainly based on the reviews of Newhauser and Zhang [38], Lomax [39], and Durante and Paganetti [47]. To understand the benefits of particle therapy over conventional x–ray radiotherapy, it is important to understand the interaction of x–ray photons with matter, which is also discussed.

2.1 INTERACTION OF PROTONS

2.1.1 Interaction mechanisms

While traversing a medium, protons interact with the atoms of the medium either via the Coulomb field of their electrons or nucleus, or via nuclear interactions. With regards to particle therapy and pCT, one is most interested to predict the energy loss and correspondingly the dose deposited in each interaction. Additionally, directional changes of the primary protons are relevant to predict the spatial distribution of the dose deposition for particle therapy and to understand the proton's curved trajectory in the case of pCT. Protons used for particle therapy are accelerated to energies between $70\,\text{MeV}$ and up to $250\,\text{MeV}$ for deep–seated tumors [39]. This puts them in the moderately relativistic range [47] with up to $\beta \approx 0.6$, where β is the particle speed relative to the speed of light. For pCT, energies of $200\,\text{MeV}$ are the current standard [48] and are foreseen for cranial applications. In this energy range, physical interaction mechanisms are well understood. They are Coulomb interactions as well as inelastic nuclear interactions and emission of Bremsstrahlung, which are briefly introduced in the following.

INELASTIC COULOMB INTERACTION   The most frequent interaction of protons with matter occurs when they come close to an electron of the atom's shell and interact with its electric field. Since the rest mass of a proton is much higher than the one of an electron, the electron is removed from the shell and the proton loses a small fraction of its energy while its direction of travel only changes negligibly. Repeated inelastic Coulomb interactions result

in a quasi–continuous deceleration of the proton while the cumulative deflection can be neglected due to the much larger mass of protons compared to electrons.

ELASTIC COULOMB SCATTERING   If the proton comes closer to the atom's nucleus it can interact with the Coulomb field of the nucleus resulting in a repulsive, but elastic scattering of the proton. Due to the higher mass of the nucleus, the direction of travel of the proton changes. Repeated elastic Coulomb scattering events are often referred to as multiple Coulomb scattering (MCS) and cause a steady broadening of a proton beam as it penetrates a medium. While this process is elastic (i. e. no energy is lost in the center–of–mass frame), some energy is transferred from the proton to the nucleus. This becomes the dominant energy loss mechanism for energies below $10\,keV$ shortly before the proton comes to a halt.

INELASTIC NUCLEAR INTERACTIONS   In case of a direct impact of the proton on the atom's nucleus, the primary proton is removed while creating secondary particles, including secondary protons, neutrons and recoil nuclei. Photons can be emitted during the de–excitation of the nucleus. Since the emission angle and energy of a secondary proton are only loosely correlated to the angle and energy of the primary proton, they can cause spurious signals in a pCT scan and need to be filtered out appropriately. Elastic nuclear interactions are negligible for proton energies relevant for therapy and imaging.

BREMSSTRAHLUNG   While the generation of Bremsstrahlung through interactions with nuclei is generally possible, it can be neglected for particle therapy and imaging applications [38].

———

While some conclusions for the development of particle therapy and particle imaging applications can be drawn by the individual analytical description of these physical effects, the most powerful application lies in the use of Monte Carlo simulations which take all interactions into account. In sections 2.1.2 to 2.1.5 the most important analytical descriptions of particle interaction with applications to therapy and imaging are summarized. Monte Carlo simulations will be discussed in section 4.1.

2.1.2 Stopping power

The loss of a proton's energy E is described by the material's stopping power S, which is defined as the ratio

$$S = -\frac{dE}{dx},$$

(2.1)

and thus as the negative energy variation (i. e. the energy loss) per unit step length in x. Often, the stopping power is normalized to a material's density ρ such that the mass stopping power is calculated as $S_{mass} = S/\rho$. For proton energies above $1\,MeV$, the stopping power is dominated by the quasi–continuous energy loss due to inelastic Coulomb interaction with shell electrons (electronic stopping power or collision stopping power). Energy loss due to elastic Coulomb scattering is referred to as nuclear stopping power and the radiative stopping power describes the energy loss due to generation of Bremsstrahlung. The latter two contributions are negligible for most relevant proton energies.
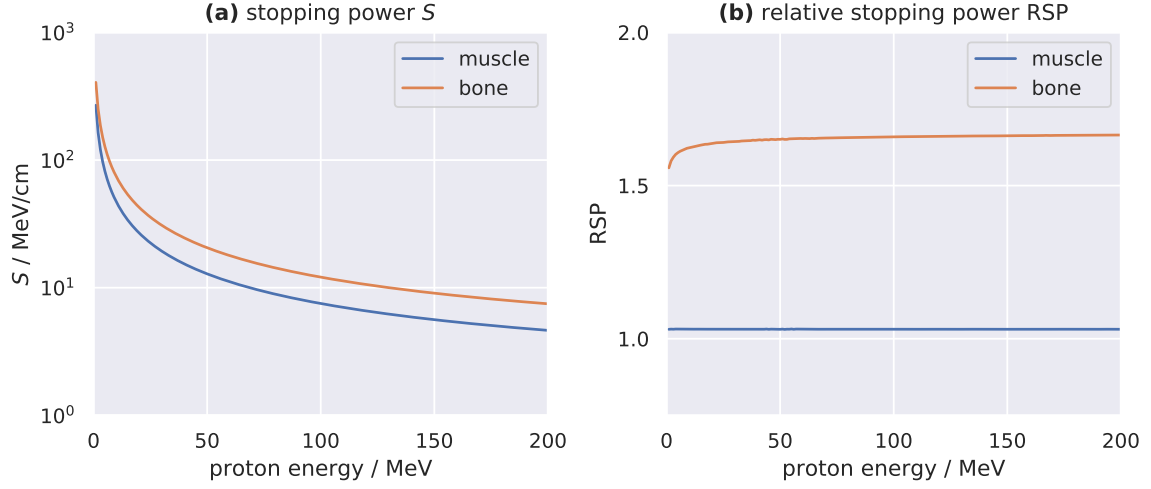
Figure 2.1: (a) Stopping power S and (b) relative stopping power as a function of the proton energy. *Data source:* PSTAR [53].

The (dominant) electronic stopping power of any particle can be estimated for energies above $1\,\mathrm{MeV/u}$ by a formula that is attributed to Bethe [49] and Bloch [50]. As a function of the particle's relativistic speed β, the stopping power is calculated as

$$S(\beta) = 4\pi r_e^2 m_e c^2 \rho_e \frac{Z^2}{\beta^2} \left[\ln\left(\frac{2m_e c^2 \beta^2}{\langle I\rangle(1-\beta^2)} \right) - \beta^2 - \frac{C}{Z_m} - \frac{\delta}{2} \right], \tag{2.2}$$

where r_e and m_e are the classical electron radius and mass, Z is the atomic number of the particle, and ρ_e and $\langle I\rangle$ are the electron density and mean excitation energy of the traversed material. C and δ were not considered in the original formulation and are correction terms. C/Z_m with the atomic number of the material Z_m is relevant for low energies when the particle velocity approaches the shell electron velocity and it cannot be assumed that shell electrons are at rest [51]. The density correction term δ considers the polarization of the medium by the particle and is not relevant for clinical proton energies [38].

At energies below $1\,\mathrm{MeV/u}$ the particle's speed becomes comparable to the orbital velocity of shell electrons ($\beta \approx 0.008$). The positively charged ion then pulls the negatively charged electrons with it causing a shielding of its charge. Barkas [52] described this shielding by replacing the atomic number of the particle by an empirical effective charge

$$Z_{\mathrm{eff}} = Z \cdot \left(1 - e^{-125\beta Z^{-\frac{2}{3}}} \right), \tag{2.3}$$

which goes to effectively zero for very low energies and reduces the stopping power.

The stopping power values for various elements and material compositions, including approximations for human tissues, are tabulated for protons in the PSTAR database [53]. This includes all relevant corrections and is validated against experimental data. The stopping power of protons is plotted in fig. 2.1 (a) for muscle and bone tissue. Within the range relevant for pCT, the stopping power changes by more than one order of magnitude and peaks for low energies at the end of the proton beam's range. This can be expected from

eq. (2.2) due to its dependency on $1/\beta^2$ while considering that the Barkas effect is not relevant for the plotted energies. However, this very pronounced peak of energy deposition is not observed for realistic beams due to the stochastic nature of the energy loss described in section 2.1.3.

Due to its strong dependency on the particle energy, which changes while the particle traverses an object, the stopping power is not a good choice as a quantity for imaging or treatment planning. Instead, pCT scanners measure the RSP, which is the stopping power of the medium S^m divided by the stopping power of water S^w, and thus

$$\text{RSP} = \frac{S^m}{S^w}. \tag{2.4}$$

The RSP is mostly independent of the value of particle's energy, which can be appreciated by the approximation of eq. (2.2)

$$\text{RSP}(\beta) = \frac{S^m(\beta)}{S^w(\beta)} \approx \frac{\rho_e^m}{\rho_e^w} \left(\frac{f(\beta) - \ln\langle I^m \rangle}{f(\beta) - \ln\langle I^w \rangle} \right) \approx \frac{\rho_e^m}{\rho_e^w} \approx \frac{\rho^m}{\rho^w} \tag{2.5}$$

with the functional dependency on β

$$f(\beta) = \ln \left(\frac{2m_e c^2 \beta^2}{1 - \beta^2} \right). \tag{2.6}$$

This rough estimation ignores the two corrections, but shows that, if the logarithm of the material's excitation energy is similar to the one of water, the RSP is independent of β and approximately equal to the ratio of the material's electron or mass density divided by the one of water. Indeed fig. 2.1 (b) shows the RSP of muscle and bone across the relevant energies for imaging, which changes only little with the proton's energy and mostly for low energies. For muscle in particular, which has a RSP around one and thus is similar to water, there is no relevant change with energy. Given the RSP of a material, one can calculate its stopping power at any energy by multiplying the RSP with the tabulated stopping power of water at that energy.

In some literature, the term stopping power ratio is used instead of relative stopping power, but both refer to the same quantity.

2.1.3 Range & range straggling

For particle therapy it is important to be able to calculate the proton's range in a material. This quantity is linked to the material's stopping power. However, due to the stochastic nature of the energy loss and deflections, two protons of the same initial energy and traversing the same material will never have the exact same range. A good assumption, in particular for clinically relevant energies, is the so–called continuous slowing down approximation (CSDA), which simply ignores random range fluctuations and deflections and integrates the inverse stopping power. The CSDA range of a proton with incident energy E_0 in a material with stopping power S is calculated as

$$R_{\text{CSDA}}(E_0) = \int_0^{R_{\text{CSDA}}} dx = \int_{E_0}^0 dE \left(\frac{dE}{dx} \right)^{-1} = -\int_{E_0}^0 \frac{dE}{S}. \tag{2.7}$$

To account for (generally small) deflections an additional detour factor $f_{detour} < 1$ can be considered to calculate the projected range $R_{proj.} = f_{detour} \cdot R_{CSDA}$, which is the average depth to which the particles of a beam arrive projected on the initial direction of the beam.

Range straggling is defined as the spread (in terms of the standard deviation) of ranges of protons with the same initial energy going through a medium. The precise analytical description of range straggling is extensive and many, mostly empirical, theories exist. The most comprehensive description which was applied to image noise of pCT by Schulte et al. [54] is the straggling theory of Tschalär [55]. For many applications, however, it is sufficient to assume that the range straggling σ_R is proportional to the range in the medium R. During the design of a prototype pCT scanner, Bashkirov et al. [56] assumed

$$\sigma_R(R) = 1.1\,\% \cdot R \tag{2.8}$$

for protons in materials relevant for pCT, which goes back to measurements of Janni [57]. Using this formula, a proton beam of 200 MeV with a CSDA range of 259.6 mm in water according to eq. (2.7) would have a range spread of 2.9 mm.

2.1.4 Deflection due to scattering

As a proton traverses a material it will frequently scatter on the Coulomb field of the material's nuclei, which is referred to as MCS. As a consequence, an initially localized and parallel proton beam will widen up, the deeper it penetrates the medium. To describe MCS, the notation of Schulte et al. [54] is adopted here. They describe the trajectory of a proton along the coordinate d (see section 3.4 and fig. 3.5 for a sketch of the coordinates). Within the u–d–plane, where u is a coordinate perpendicular to d, scattering can be described using the parameter vector

$$y_1 = \begin{pmatrix} u_1 \\ \theta_1 \end{pmatrix}, \tag{2.9}$$

where θ_1 is the angle of the proton's direction of travel with the d–axis at a given depth d_1 and u_1 is the proton's u–coordinate at d_1. According to Eyges [58] who used Fermi's scattering theory, the likelihood L of finding the proton with the parameters y_1, given that it entered the medium at $u_0 = \theta_0 = 0$, can be modeled as a bivariate Gaussian

$$L\left(y_1 \,\middle|\, y_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}\right) = \exp\left(-\frac{1}{2}y_1^\mathsf{T}\Sigma_1^{-1}y_1\right), \tag{2.10}$$

where Σ_1 is the symmetric positive–definite scattering matrix. The scattering matrix describes the variances and covariances of y_1 and θ_1, i. e. the spread that accumulated when the proton traveled from d_0 to d_1. It can be written as

$$\Sigma_1 = \begin{pmatrix} \sigma_{u_1}^2 & \sigma_{u_1\theta_1}^2 \\ \sigma_{u_1\theta_1}^2 & \sigma_{\theta_1}^2 \end{pmatrix}. \tag{2.11}$$

There are several works calculating the elements of the scattering matrix, including the one of Eyges [58] and Highland [59]. The most commonly used scattering theory for pCT is

the one proposed by Lynch and Dahl [60], which was used by Schulte et al. [54] to estimate the matrix elements as

$$\sigma^2_{u_1} = E_0^2 \cdot \left(1 + 0.038 \ln \frac{d_1 - d_0}{X_0}\right)^2 \cdot \int_{d_0}^{d_1} dd \, \frac{(d_1 - d)^2}{\beta^2(d)p^2(d)X_0}, \tag{2.12}$$

$$\sigma^2_{\theta_1} = E_0^2 \cdot \left(1 + 0.038 \ln \frac{d_1 - d_0}{X_0}\right)^2 \cdot \int_{d_0}^{d_1} dd \, \frac{1}{\beta^2(d)p^2(d)X_0}, \tag{2.13}$$

$$\sigma^2_{u_1\theta_1} = E_0^2 \cdot \left(1 + 0.038 \ln \frac{d_1 - d_0}{X_0}\right)^2 \cdot \int_{d_0}^{d_1} dd \, \frac{d_1 - d}{\beta^2(d)p^2(d)X_0}. \tag{2.14}$$

Here, $E_0 = 13.6 \, \text{MeV}/c$ and 0.038 are empirical constants that were found to match experimental data. X_0 is the material's radiation length, which for water is $X_0 = 361 \, \text{mm}$. The speed β and momentum p of a proton at depth d depend on the initial energy and were found by Schulte et al. [54] by fitting a polynomial to $1/\beta^2(d)p^2(d)$ from a Monte Carlo simulation.

This description of the Gaussian spread of a particle in terms of a likelihood for the coordinates u and θ will later be used in section 3.4 to estimate the proton's most likely path inside a medium given location and direction measurements of a pCT scanner.
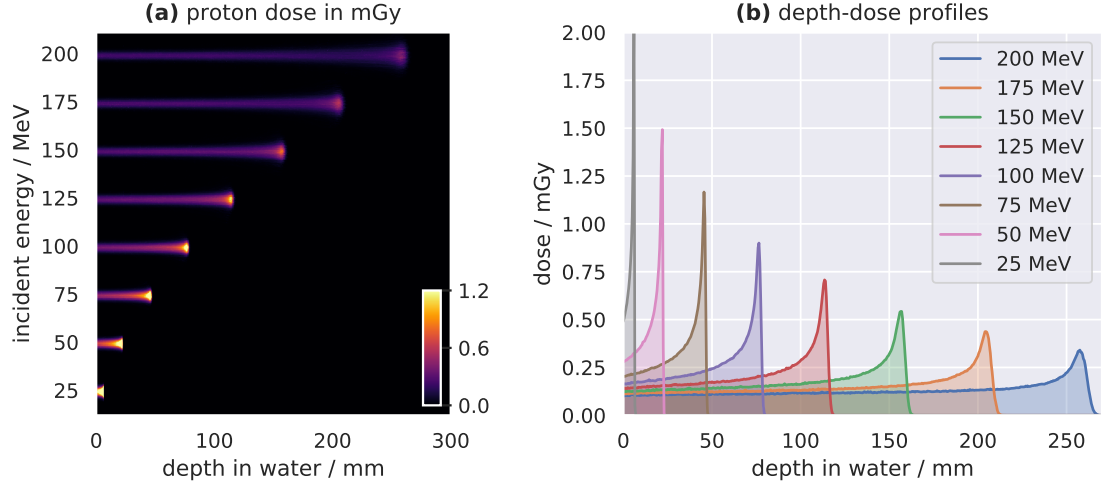
2.1.5 Depth–dose curve of protons

The dose deposit $D = dE/dm$ is defined as the energy dE absorbed in a medium per unit mass dm. Using the mass stopping power $S_{\text{mass}}(E, \vec{r})$ at a given point $\vec{r}$, the dose deposit $D(\vec{r})$ at that point from a proton beam with a fluence spectrum $\Phi(E, \vec{r})$ at the same location can be calculated as an energy integral

$$D(\vec{r}) = \int_0^\infty dE \, \Phi(E, \vec{r}) \cdot S_{\text{mass}}(E, \vec{r}). \tag{2.15}$$

As the proton beam penetrates a medium its mean energy decreases, which in turn increases the medium's mass stopping power experienced by the proton. At the same time, nuclear interactions remove primary protons from the beam, which reduces the fluence. However, a precise three–dimensional calculation of the dose deposition as needed for particle therapy is complex and affected by range straggling, the removal of primary protons due to nuclear interactions, the impact of secondary particles and deflections due to MCS. While the equations presented here describe the individual quantities, the most complete picture can be obtained using Monte Carlo simulations with toolkits such as GEANT4 [61].

Dose deposition maps of perfectly monoenergetic proton beams in water have been calculated in fig. 2.2 (a) for eight different incident proton energies. Each beam consisted of 10 000 particles with an initial Gaussian spatial distribution with a full width at half maximum (FWHM) of 5.8 mm, which is a typical beam size for PBS proton therapy although

Figure 2.2: (a) Dose maps of proton beams in water at various incident energies and (b) corresponding depth–dose curves laterally integrated across the beam. *Data source:* GEANT4 [61].

this may depend on the beam energy in reality. In fig. 2.2 (b) the corresponding depth–dose curves obtained as laterally integrated profiles for each beam are displayed. It can be observed that for the smallest incident energies the peak in the dose deposition, the so–called Bragg peak, is very pronounced. As the energy, and with it the range in water, increases, the Bragg peak is spread out due to range straggling. While a single proton within the beam still experiences its highest energy loss at the end of the range, the physical location of this dose deposition is less confined within a beam. Prior to the Bragg peak there is a dose plateau which results from the slowly increasing stopping power at lower energies while more and more primary particles are removed due to nuclear interactions. The beams of the highest energies are considerably more spread out laterally compared to the low energy beams at the end of their range due to more deflections over the longer penetration depth. For protons with a range of 200 mm about 20 % of the primary particles are removed from the beam by nuclear interactions prior to the Bragg peak [39].

2.2   INTERACTION OF PHOTONS

The interaction of x–ray photons with matter is conceptually different from the interaction of charged particles. For sufficiently thin absorbers, a certain fraction of photons will pass through without interacting with the object. The remaining photons are either absorbed or scattered. This removal of photons from the primary beam is described by the attenuation coefficient μ and the differential equation

$$dN = -\mu N dx, \tag{2.16}$$

where N is the number of photons and x is the depth in the absorber. This leads to an exponential decay described by the Beer–Lambert law

$$N(x) = N_0 \cdot e^{-\mu x}. \tag{2.17}$$

The attenuation coefficient μ can be separated into contributions from different physical effects, namely Rayleigh scattering (μ_{RS}), the photoelectric effect (μ_{PE}), Compton scattering (μ_{CS}), and pair production (μ_{PP}), such that

$$\mu = \mu_{RS} + \mu_{PE} + \mu_{CS} + \mu_{PP}. \tag{2.18}$$

The four interaction processes are briefly summarized below and their individual contributions for two different materials, water and bone, are shown in fig. 2.3.

RAYLEIGH SCATTERING If the photon interacts with a nucleus of a dimension that is smaller than its wavelength, its energy can be temporarily absorbed by the atom's shell electrons followed by an emission of a photon of the same wavelength, but in an arbitrary direction. Rayleigh scattering, therefore, is an elastic scattering process. While it is not a dominant interaction it can still impact the quality of imaging setups if not accounted for.

PHOTOELECTRIC EFFECT If the photon's energy is sufficiently high, it can ionize atoms of the material and remove one of the shell electrons [63]. The remaining photon energy is converted to kinetic energy of the electron and the photon is absorbed completely. Eventually, a relaxation can cause further emission of radiation. The photoelectric effect dominates for low energies of up to $30\,keV$ for water and $60\,keV$ for bone.

COMPTON SCATTERING In addition to the photoelectric effect, photons can inelastically scatter on weakly bound electrons of the absorbing material [64]. The photon transfers part
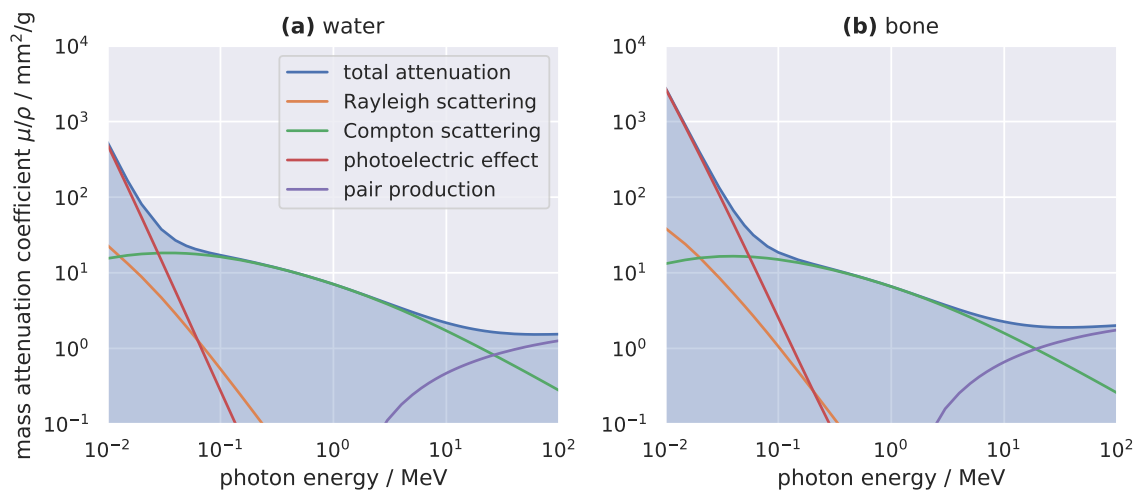
Figure 2.3: Mass attenuation coefficient and its contributions for x–ray photons in (a) water and (b) bone. The legend applies to both figures. *Data source:* NIST [62].

of its energy to the electron and is emitted in a different direction. The energy loss of the photon is directly linked to its scattering angle. Compton scattering is the dominant effect for intermediate energies around 1 MeV.

PAIR PRODUCTION In the vicinity of a nucleus or electron, and if the energy of a photon exceeds 1022 keV and thus twice the rest mass of an electron, it can transfer its energy to an electron–positron pair. The photon is absorbed and the remaining energy is transferred as kinetic energy to the electron and the positron. Pair production becomes the dominant effect at around 20 MeV. Electron–positron pairs can also be produced in the vicinity of electrons, in which case the process is referred to as triplet production.

––––––––

The relevant energy ranges for x–ray CT are from 50 keV to 140 keV making the photo-electric effect and Compton scattering the most dominant interaction processes. Rayleigh scattering can also impact the image quality. Radiotherapy typically employs x–ray beams with photons of few and up to 15 MeV making Compton scattering and pair production the most important processes of dose deposition. In both cases, the attenuation of the primary photon beam is governed by an exponential decay described by the Beer–Lambert law in eq. (2.17).

2.3 IMPLICATIONS ON THE TREATMENT WORKFLOW

2.3.1 Photon & particle therapy

The physical processes described in this chapter have important implications on the use of ionizing radiation for radiation therapy in general, and in particular on the choice between photons and charged particles for treatment delivery. The first and foremost goal of cancer therapy is the eradication of the tumor and, therefore, the dependable deposition of the prescribed dose. Secondly, the sparing of healthy tissue is required to ensure a safe and tolerable treatment delivery. With their exponential decay of deposited dose described by the Beer–Lambert law in eq. (2.17), photons do not seem an ideal candidate for this task, in particular for deep–seated tumors. For a single photon beam, the highest dose deposition will be close to the patient's skin (after a small build–up region) and a tumor, for example 100 mm below the skin, will receive considerably less dose. However, this is only a hypothetical problem and can be solved by exposing the patient to photon beams from many angles using a treatment gantry and superimposing the dose inside the tumor volume [6]. While the tumor receives dose from all angles, healthy tissue and in particular tissue far away from the tumor, will only receive treatment dose when they are aligned with the tumor volume. When OARs are aligned with the tumor, the intensity can be reduced and OARs spared additionally. A modulated intensity can also be used to make the dose distribution in the tumor flat and as close as possible to the prescribed dose.

Protons, on the other hand, allow us to place the point of the highest dose deposition, the Bragg peak, inside the tumor. A combination of several incident energies allows to have a flat dose deposition inside the tumor, even from a single gantry angle [65]. This

(a) photon plan   **(b)** proton plan
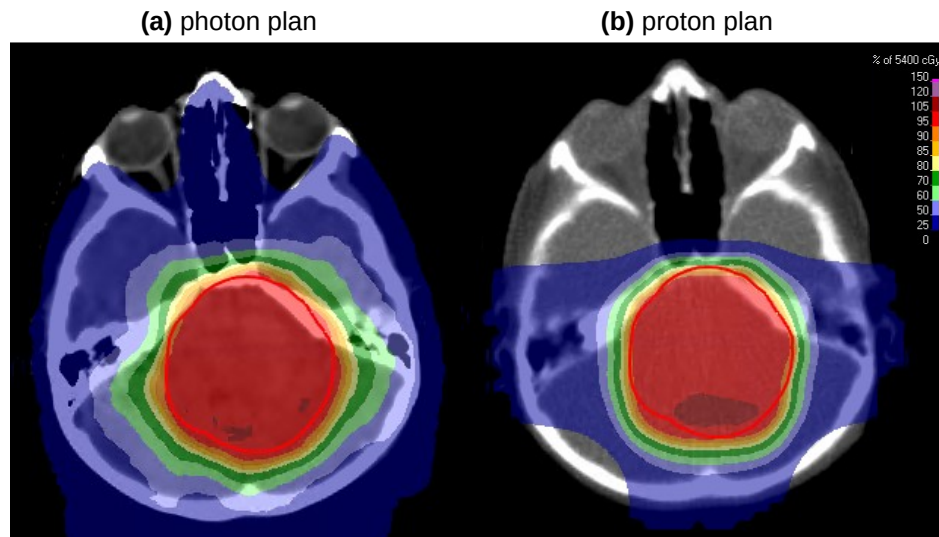
Figure 2.4: Exemplary dose maps of (a) a clinically delivered photon treatment and (b) a correspond-
ing proton treatment planned for the same patient. The PTV is indicated by a red line.
The photon plan is shown on top of an x–ray CT image while the proton plan is shown
on top of a pCT image explaining the different contrasts and resolution. *Treatment plans:*
photon plan and delineations from University Hospital, LMU Munich, proton plans
created with RayStation (*RaySearch Laboratories AB*, Stockholm, Sweden).

combination of Bragg peaks is referred to as spread–out Bragg peak. While the dose
deposition inside and close to the tumor volume can be comparable to the one with photons,
the big advantage of protons is that they deposit close to no dose behind the tumor and
considerable less dose before. Several spread–out Bragg peaks from various angles allow to
further push down dose in healthy tissue while maintaining the flat dose deposition in the
tumor. Even though attempts are made to employ proton therapy with a continuous rotation
[66] such as for photons, currently typically only up to three or four fields are superimposed.
Nevertheless, proton therapy has been shown to produce superior dose maps compared to
x–ray radiotherapy [14], in particular in the low–dose region. It has also been shown to have
superior clinical outcomes with less side effects for several clinical indications [13, 15].

Figure 2.4 (a) shows a clinically delivered photon plan for a pediatric patient with a central
brain tumor. The PTV is nicely covered with a sharp gradient, but a low–dose bath spans
the whole head, except for the eyes, which were additionally spared as they are OARs. In
contrast, fig. 2.4 (b) shows a corresponding proton treatment plan for the same patient and
with the same PTV. The dose gradient around the PTV is considerably sharper compared
to the photon plan. Three beam directions were used that contribute to a low–dose region
around the ears and at the back of the skull. Apart from that, the nose and specifically the
eyes receive close to zero dose during the treatment. The area covered by relevant treatment
doses is considerably smaller for the proton plan compared to the photon plan.

2.3.2 Imaging for particle therapy

To optimize treatment plans for particle therapy, such as the one shown in fig. 2.4 (b), the TPS depends on a volumetric map of the patient's the RSP. However, this information is not directly available in a clinical setting. Instead, x–ray CT images are used to obtain the relevant information. From a physical perspective, such images show the photon attenuation coefficient, which depends strongly on the photon energy and consequently on the CT scanner's spectrum, which varies between models and manufacturers and even between scan protocols. For this reason, CT images are shown as CT values in the Hounsfield (HU) scale, which involves a normalization forcing the attenuation coefficient of water to be at $0\,\mathrm{HU}$ and that of air to be at $-1000\,\mathrm{HU}$. As discussed in this chapter, both the energy loss of protons and the attenuation processes of photons are related to shell electrons of the absorbing material. In particular, the Bethe–Bloch equation in eq. (2.2) as well as the various contributions to the photon attenuation coefficient in eq. (2.18) depend on the electron density ρ_e. This makes it possible to establish a relationship between the RSP and the CT values. In the state–of–the–art stoichiometric calibration [67], a phantom with inserts of known RSP is scanned and a piecewise–linear calibration curve fitted to the data, relating CT values to RSP. This method was shown to introduce proton range errors of up to $3\,\%$ [68, 69].

More recently, spectral CT has become available for proton treatment planning [70]. Dual energy computed tomography (DECT) scanners employ two different x–ray spectra or two detectors with varying spectral sensitivity, allowing to probe the attenuation coefficient at two different effective energies. Since the contributions of the attenuation coefficient follow simple mathematical laws, a linear operation on the two resulting images enables the calculation of estimated maps of the electron density ρ_e and an effective atomic number $Z_{m,eff}$ of the material [71]. $Z_{m,eff}$ can then be converted to the material's ionization potential $\langle I \rangle$ using a piece–wise linear function [72, 73]. From there, the Bethe–Bloch equation can be used to calculate the RSP. Thereby, DECT can reduce the uncertainty of RSP maps to around $1\,\%$ [22, 30, 74–78]. In the future, DECT scanners are likely to be superseded by photon–counting CT detectors [79], which directly measure the energy of incident photons and may further improve accuracy of treatment planning [80].

Considering the uncertainties involved with imaging in particle therapy, in particular due to the conversion of photon interaction to proton stopping power, it would be beneficial to avoid the conversion step and directly employ protons for imaging of the RSP. The following chapter 3 will describe in detail the concept of pCT and the required technical instrumentation while referring back to this section for the underlying physical processes.

3

PROTON COMPUTED TOMOGRAPHY

PROTON COMPUTED TOMOGRAPHY is the main imaging modality under investigation in this work. In this chapter, the detection principle of pCT, the necessary technical instrumentation and important algorithms for data processing are discussed. Since this work was strongly coupled to a specific prototype pCT scanner [56, 81, 82], the arguments will focus on this scanner with generalizations to other systems, where required. The notation introduced in this chapter will be carried on to the following chapters.

3.1 DETECTION PRINCIPLE

3.1.1 The water–equivalent pathlength

PCT scanners make use of energy measurements to produce volumetric maps of the RSP. As introduced in eq. (2.1), a material's stopping power is defined as the negative energy variation (i.e. the energy loss) per unit step length $S^m = -dE/dx$. For a proton with an initial energy E_0 traveling along a curved path $\mathcal{C}$, this results in an energy loss of

$$\Delta E = \int_{\mathcal{C}} dx\, S^m(x) = -\int_{E_0}^{E_1} dE\ = E_0 - E_1, \tag{3.1}$$

where E_1 is the proton's energy at the end of the path. The stopping power is a function of x and can vary along the path. If instead of the stopping power the line integral is performed on the $RSP = S^m/S^w$, the integral becomes

$$\int_{\mathcal{C}} dx\, RSP(x) = \int_{\mathcal{C}} dx\, \frac{S^m(x)}{S^w(x)} = -\int_{E_0}^{E_1} \frac{dE}{S^w(E)} \equiv WEPL, \tag{3.2}$$

where the right hand side of the equation is defined as the water–equivalent pathlength or WEPL. The WEPL is measured in units of length can be understood as the length in water that is required to cause the same energy loss.

While the RSP is not directly accessible to measurements, the WEPL can be determined by measuring the energy of protons exiting an object. Assuming that the incident proton

energy is constant at $E_0 = 200\,MeV$ and that the exit energy E_1 was determined by a pCT detector, the WEPL can be calculated as

$$WEPL(E_1) = -\int_{200\,MeV}^{E_1} \frac{dE}{S^w(E)} = \int_{E_1}^{200\,MeV} \frac{dE}{S^w(E)}, \tag{3.3}$$

where the stopping power of water $S^w(E)$ is tabulated and could be approximated by a polynomial. Since $S^w(E)$ is strictly positive and $E_1 \leqslant 200\,MeV$, a higher residual energy E_1 will always result in a smaller WEPL and vice versa, or in other words eq. (3.3) is a monotonically decreasing function.

Experimentally, it is difficult to determine absolute energy measurements, which is why detectors are typically directly calibrated to produce WEPL values as described in section 3.2.3. Nevertheless, it is useful to think of the detection process as an energy measurement. In fact, detectors are referred to as energy detectors.

To produce RSP maps of the patient, the integral in eq. (3.2) needs to be inverted. By integrating over the spatial dependency of the RSP, information is lost, that is not encoded in a single energy or WEPL measurement. However, it can be restored by measuring two–dimensional maps of the WEPL from many directions. This restoration, which is called *reconstruction*, is described in section 3.5.

Apart from protons, also helium ions [83–85] and carbon ions [86] can be used for tomographic acquisitions, but are not covered in this thesis.

3.1.2 Single particle tracking proton CT

Keeping in mind the physical interaction processes of protons described in sections 2.1.3 and 2.1.4, the WEPL detection principle summarized in eq. (3.3) imposes several challenges to pCT which are outlined in the following.

- While in theory a single proton would be sufficient to solve eq. (3.3), the stochastic nature of the energy loss (range straggling) requires that at each point a set of protons is measured to calculate an average WEPL in order to reduce image noise.

- Nuclear interactions cause spurious detector signals since they do not follow the assumption of a continuous energy loss as described by the Bethe–Bloch formula in eq. (2.2). To avoid distortions of the image, particles undergoing nuclear interactions either in the object or in the detector are filtered out. Details of data filtration are described in section 3.3.

- Due to MCS, the path $\mathcal{C}$ is not a straight line between the exit and entrance coordinate and the exact path is not known. A straight path is an assumption for many reconstruction algorithms for x–ray CT, but also curved paths can be taken into account with dedicated algorithms for pCT, if they can be estimated. But even then will an incorrect path assumption result in a blurred image. Therefore, ideally, a pCT scanner measures the entrance and exit coordinates as well as the travel directions of each proton individually, to best estimate the proton path and avoid image blurring. Details of the path estimation are described in section 3.4.

In particular the requirement for path estimation and data filtration demand that pCT scanners record the WEPL and coordinates of individual protons (instead of the average WEPL of a set of protons). Such scanners are referred to as *single particle tracking* pCT scanners. This requires that the electronics is capable of particle detection at around 1 MHz or above to allow for reasonably low scan times of a few minutes. Such hardware has become available in the last two decades, initiating the development of several prototype scanners in academic [81, 87, 88] as well as early commercial environments [89]. Data from single tracking detectors is often referred to as list mode data, since they consist of a list of single proton measurements.

3.1.3 Integration mode proton CT

Single particle tracking pCT setups are very costly due to the requirements on electronics and because the components are not available from mass production and need to be built individually. Systems are typically also slow to operate due to the limited particle flux that is dictated by the maximum possible particle detection rate. As an alternative, simple and comparably cheap detectors can be utilized as WEPL detectors, such as x–ray flatpanel detectors [90] or proton range detectors used for quality assurance [91]. Since such detectors do not detect single particles, but instead integrate the WEPL over all protons incident to a given pixel of the detector, they are referred to as *integration mode* pCT detectors. Thin pixelated flatpanel detectors can be used by scanning the incident proton energy to locate the Bragg peak which requires imaging doses that are considerably higher than those of single particle tracking systems and are between 50 mGy and up to several Gray per tomography [92–94]. This limits the potential clinical applicability of integration mode detectors. However, there may be applications in the context of small animal imaging [95, 96]. This thesis will focus exclusively on single particle tracking systems.

3.2 TECHNICAL INSTRUMENTATION OF A PROTOTYPE PROTON CT SCANNER

In this work, the phase II prototype pCT scanner [56, 81, 82, 97, 98] developed at the Loma Linda University (LLU) and the University of California Santa Cruz (UCSC) was used extensively both for experiments and in simulations. The single particle tracking scanner consists of two tracking detectors, one front tracker prior and one rear tracker after the object, as well as a downstream energy detector. The incident proton energy is not measured and assumed to be equal to the accelerator setting. Protons are emitted from a proton beamline and the object to be scanned is mounted on a rotation stage. The experimental setup is shown in fig. 3.1, where it is mounted in a treatment room at the Northwestern Medicine Chicago proton center. A corresponding schematic drawing can be found in fig. 3.2, where also the coordinate system used to describe the measurements is indicated.
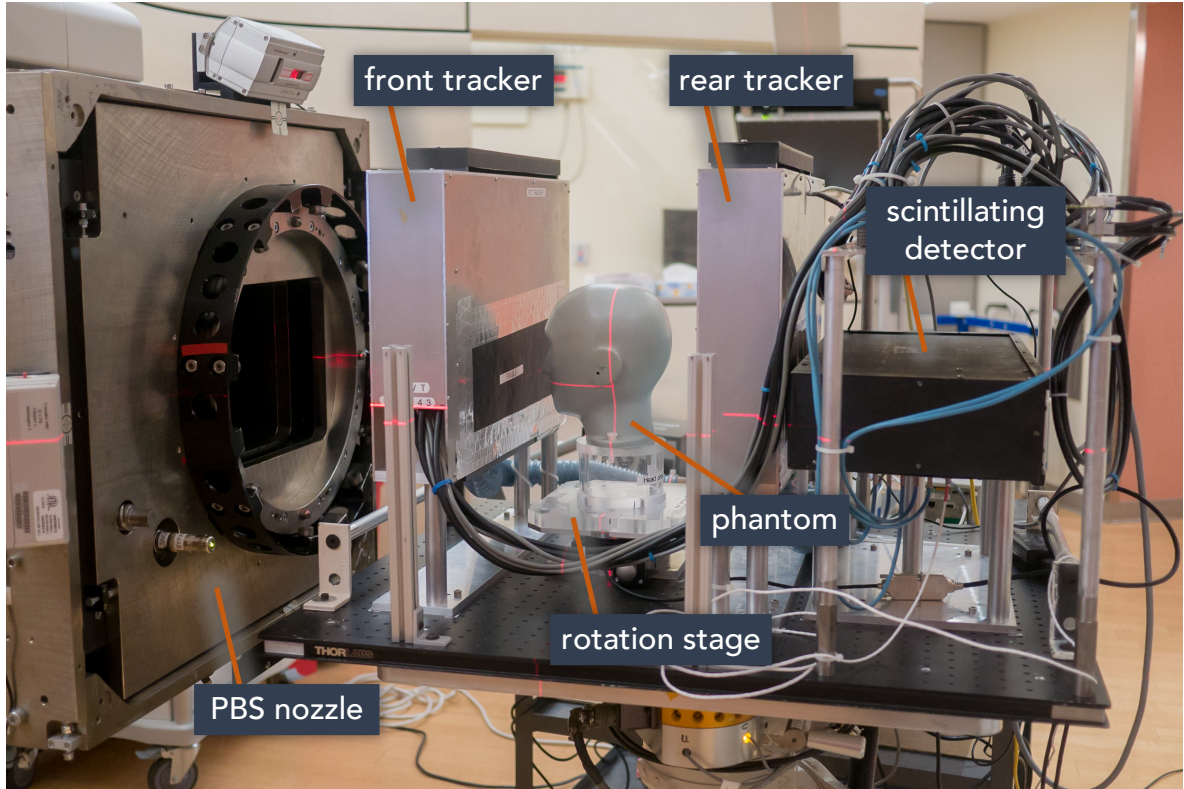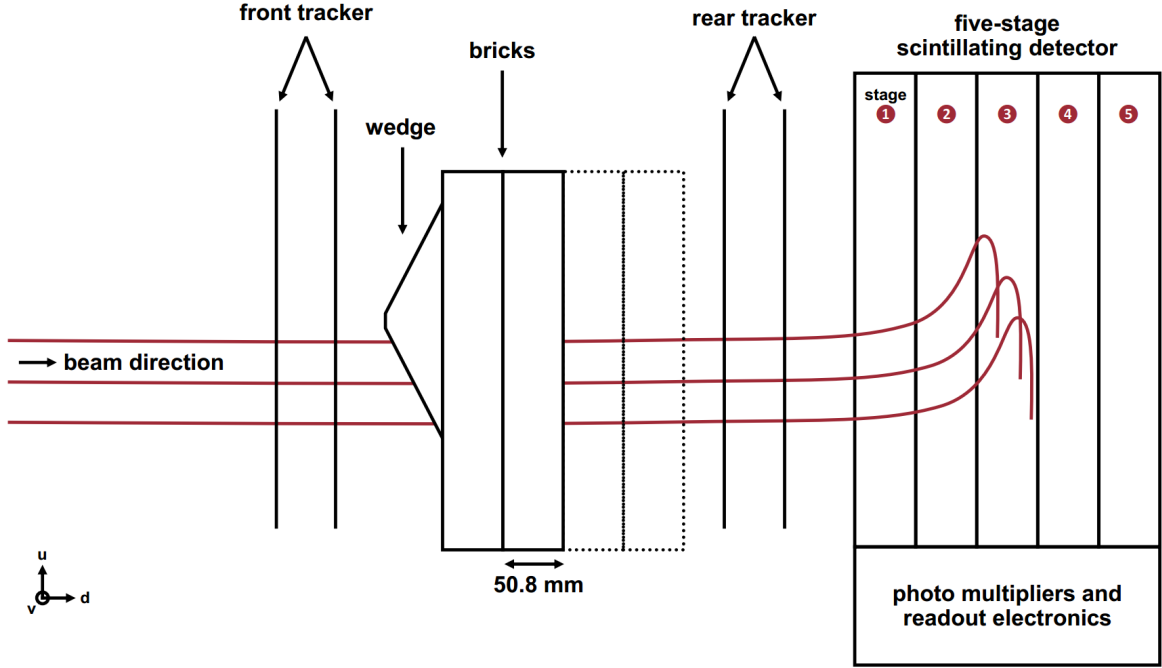
Figure 3.1: Photograph of the prototype pCT scanner used in this work mounted in the Chicago proton center. Abbreviations: PBS – pencil beam scanning.

3.2.1 Tracking detectors

Each of the two identical tracking modules [82], which can be seen in fig. 3.1 as aluminum housings with a black opening window, measures the position of incident protons at two distinct locations along the longitudinal coordinate d as indicated in fig. 3.2. The two position detectors within each tracking module are interspaced by 50 mm [97] and measure the horizontal u– and the vertical v–coordinate. Each position detector consists of two stacked silicon strip planes, which can determine either the u– or the v–coordinate. Each plane is 0.4 mm thick. The strip pitch in both planes is 0.228 mm with 1536 strips in the u–plane and 384 strips in the v–plane [82]. The total area of each tracking plane is 352 mm × 88 mm. Each tracking plane consists of four modules which are horizontally glued together with an insensitive gap of 0.6 mm. The tracking detectors are staggered to avoid that protons hit more than one gap allowing to estimate the coordinates even if data are missing from one of the eight planes. The total material budget of both tracking modules is $8 \cdot 0.4\,\mathrm{mm} = 3.2\,\mathrm{mm}$ of silicon which corresponds to an energy loss of approximately 2.8 MeV at an incident energy of 200 MeV and 7.5 MeV at 50 MeV [53].

Figure 3.2: Schematic drawing of the prototype pCT scanner including the phantom used for calibration which consists of bricks and a double wedge. The illustration of Bragg peaks is not to scale. The coordinates (u, v, d) are indicated below on the left, where the d–coordinate goes into the plane and the origin of the coordinate system is at the center between the two trackers, where the phantom is located. (cc) Adapted from Dickmann et al. [100] under the CC BY 3.0 license.

3.2.2 Energy detector

The setup's energy detector [56] is located behind the rear tracker and consists of five scintillating detectors, which are wrapped with reflective material and individually coupled to a photomultiplier tube. Each scintillating detector is referred to as a *stage* and the whole system is referred to as a five–stage energy detector. The segmented detector design has been found to minimize WEPL noise and compromises between noise caused by range straggling and by intrinsic noise of the scintillator.

The noise in a monolithic energy detector as a function of the WEPL W would be a sum of the range straggling up to the detector and the intrinsic measurement uncertainty of the detector. The water–equivalent range up to the detector is $W + S_0$, where S_0 accounts for all materials except for the scanned object, like the tracking detector. The magnitude of range straggling can be estimated using eq. (2.8). The detection noise of the detector is proportional to the residual water–equivalent range in the detector $R_0 - S_0 - W$, where R_0 is the total water–equivalent range of protons at the initial energy. Bashkirov et al. [56] empirically estimated this uncertainty as $p \cdot (R_0 - S_0 - W)\delta E$, where $p = 1.8$ results from a fit of the tabulated R versus E relationship [53] and $\delta E = \sigma_E/E$ is the relative energy resolution of the scintillator. This results in a total WEPL uncertainty of approximately

$$\sigma_{W,\text{monolith}} \approx \sqrt{(1.1\,\% \cdot (W + S_0))^2 + (p \cdot (R_0 - S_0 - W) \cdot \delta E)^2} \qquad (3.4)$$

for a monolithic detector. With an expected energy resolution of about $\delta E = 2\,\%$ [56], image noise will be dominated by the detection noise and will be far from the range straggling limit of $\sigma_S = 1.1\,\% \cdot (W + S_0)$, in particular for low WEPLs. To avoid this, Bashkirov et al. [56] proposed a segmented design, for which the WEPL noise can be approximated by

$$\sigma_{W,\text{segmented}} \approx \sqrt{(1.1\,\% \cdot (R_0 - R_i))^2 + (p \cdot R_i \cdot \delta E)^2}, \qquad (3.5)$$

where R_i is the residual water–equivalent range of the proton at the entrance of the final stage S_i, in which the proton stops. This stage is referred to as *stopping stage*. For an energy resolution of $\delta E = 2\,\%$, in total five stages were sufficient to reduce the WEPL uncertainty below 3 mm per proton for all WEPLs up to 260 mm. This is just slightly above the energy straggling limit, which means, that the second summand under the square root in eq. (3.5) is small compared to the first.

The final detector [56] consists of five stages made from UPS–923A [101], which is a polystyrene based scintillator material with optimized light output. The RSP of the material is about 1.038 [56]. To cover the dynamic WEPL range of 260 mm, each stage is 51 mm thick, which corresponds to a WEPL of 52.9 mm. The lateral dimension of the scintillating blocks is 100 mm × 400 mm with an additional trapezoidal light guide at the side of the photomultipliers. Each stage is wrapped in a 65 μm thick reflective film. The schematic drawing in fig. 3.2 shows the segmentation and is approximately to scale.

3.2.3 Calibration

Each incident proton will trigger five energy measurements (one for each stage) in the analog–to–digital converters (ADCs) attached to the photomultiplier tubes. These ADC counts are converted to WEPL values in a two–step calibration procedure [56, 82, 102], based on calibration measurements with a phantom of known geometry and RSP.

In a fist step an energy scale is established for each stage from ADC data acquired with 200 MeV protons incident to the detector and with no absorber in place. Data in such an acquisition will exhibit two peaks in a histogram of ADC counts of a single stage. The peak at higher ADC counts corresponds to the expected mean energy deposition in that stage for 200 MeV protons. The expected mean energy was estimated using a Monte Carlo simulation modeling the full geometry of the setup [103] as

$$E_n^{G4} = \{25.25, 28.01, 32.76, 42.62, 67.71\}\,\text{MeV}, \qquad (3.6)$$

where n is the stage number. The second, much smaller peak corresponds to a zero energy deposit and is called the pedestal. It can be caused by protons triggering the acquisition in one stage, but then scattering out of the detector and not reaching the other stages. A noisy reading or cosmic events can also cause a zero–energy reading. Using these two known energies, a linear relationship can be established between the ADC count a and the energy E deposited in a given stage. This gives

$$E(a) = g \cdot (a - p), \qquad (3.7)$$

where g is the gain and p is the pedestal, which need to be determined for each stage. Due to the geometric shape of the scintillating detector, the gain will vary as a function of the
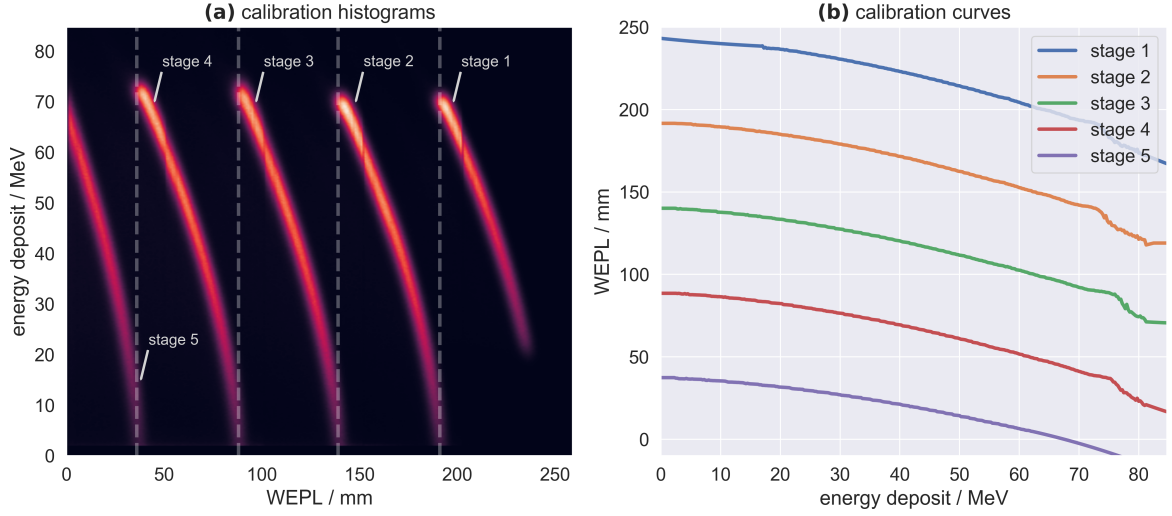
Figure 3.3: (a) Calibration histograms and (b) resulting calibration curves for the WEPL calibration of the scanner.

lateral coordinates u and v. The coordinates for each proton are known from the tracking detector and can be extrapolated along the d–coordinate to the center of each stage. Then, the gain calibration is performed on subsets of the data with a grid size of $10\,\text{mm} \times 10\,\text{mm}$ resulting in a location dependent gain $g = g(u, v)$.

A second step requires five acquisitions using a double wedge–shaped phantom with known geometry and RSP. The phantom is illustrated in fig. 3.2 and a photograph is shown in fig. 3.6 (e). The double wedge consists of polystyrene with an RSP of 1.030 and has a physical thickness varying between $0\,\text{mm}$ and $50.8\,\text{mm}$. By adding between zero and four bricks of the same material and with a thickness of $50.8\,\text{mm}$, the whole dynamic WEPL range of the detector can be scanned. The bricks are shifted by $20\,\text{mm}$ with respect to the wedge's edge to prevent protons going through the wedge but exiting at the side of a brick. The WEPL of an incident proton only depends on the energy deposit to the stopping stage. This stage is defined as the furthest stage with an energy deposit of at least $1.0\,\text{MeV}$ or $1.5\,\text{MeV}$, depending on the noise level in the detector. From the data of the calibration phantom, a two–dimensional histogram is created based on the energy deposit to the stopping stage calculated with eq. (3.7) and the WEPL of the corresponding proton. The WEPL can be calculated by intersecting the known geometry of the calibration phantom with the trajectory of the proton, which can be estimated from the tracking data. Such a histogram is shown in fig. 3.3 (a), where the contributions of all stopping stages are summed. As expected, for each stage, the energy deposit decreases as the WEPL increases. When the WEPL reaches $36\,\text{mm}$, protons stop in stage 4 instead of stage 5 and the energy deposit increases rapidly. This point is indicated by a dashed line and more stage transitions occur at $88\,\text{mm}$, $139\,\text{mm}$, and $191\,\text{mm}$. Stages 1 to 4 have a maximum energy deposit of approximately $70\,\text{MeV}$. The maximum energy deposit in stage 5 is lower, since $200\,\text{MeV}$ protons do not have sufficient energy to fully penetrate all five stages. Stopping stage 1 has a minimum energy deposit of $20\,\text{MeV}$. This is because an energy deposit of $20\,\text{MeV}$ to the first stage is used to trigger the data acquisition. From the histograms, calibration curves as depicted in fig. 3.3 (b) can be

found by fitting the most frequent WEPL for each energy deposit bin. Histograms for the five stopping stages are treated individually resulting in five independent calibration curves.

For processing of subsequent proton data, first the energy of the proton is calculated according to the gain attributed to the hit location and the pedestal. Then, the proton's WEPL is calculated by determining its stopping stage and applying the corresponding calibration curve. After the calculation of the energy, data cuts can be applied to filter spurious detector signals. Furthermore, the location–dependent gain can be proportionally re–adjusted on the fly based on protons at the edge of the scanners field–of–view which are assumed to go through air. This avoids that a change of the scanners temperature or other environmental factors distort the results over time.

The list mode output of the scanner is a proton–by–proton list of five measurements: the hit location and direction vector at the front tracker and at the rear tracker as well as the WEPL.

3.2.4 Other prototype scanners

Apart from the prototype pCT scanner used in this work, several other groups have built single particle tracking pCT systems. The specifics of the systems are briefly summarized in the following. All scanners are named by the city in which the proton center where they are operated is located while acknowledging that most of them have been built as part of an international effort. Some of the listed scanners have not been completed yet, but first experimental tests were reported. Historic efforts that are not operational anymore as well as systems intended for radiography are not considered.

CHICAGO, UNITED STATES   ProtonVDA Inc. built a pCT scanner [89, 104] employing a single monolithic energy detector and using scintillating fibers for tracking. The system is intended for commercialization and employs proton beams with varying incident energy to allow for a compact and noise–efficient design of the energy detector. To allow for a smaller dimension, the system only has a single tracking plane prior and after the object and does not allow for direction measurements and instead estimates them from the incident beam directions.

TRENTO, ITALY   The relatively small prototype system [88] with an aperture of 50 mm × 200 mm consists of two silicon strip tracking modules allowing for position and direction measurement as well as a yttrium aluminum garnet scintillating detector. RSP accuracies better than 1 % were reported.

MANCHESTER, UNITED KINGDOM   The PRAVDA consortium developed a high–speed tracking detector [105] capable of particle detection rates of 200 MHz when particles are uniformly distributed on the detector [87]. This is possible because each tracking module consists of three silicon strip detectors rotated at 60 degrees to one another allowing one detector to suffer from pile–up while the remaining two detectors still produce a usable position reading. The scanner was first operated in South Africa, but has since been moved to Manchester to build a follow–up scanner with the name OPTIMA.

BERGEN, NORWAY This scanner, which is currently in the final design phase [106, 107] does not employ separate tracking detectors, but uses the known location and direction of a proton pencil beam as entrance tracking and features a pixelated range telescope that can both measure the proton's position, direction and WEPL. This compact system will most likely have an inferior spatial resolution [108], but could be more easily implementable in a clinical environment due to its compact design. Experiments with subsystems are currently performed elsewhere, but Bergen is expected to have a proton center by 2023.

WIENER NEUSTADT, AUSTRIA This very early prototype pCT system [109] consists of three silicon strip tracking planes prior and three after the object with an aperture of $25\,\text{mm} \times 50\,\text{mm}$ and a plastic scintillator range telescope. The particle tracking rate is currently at $500\,\text{Hz}$, resulting in scan times of several hours per projection. The system is intended for prototyping a larger functional system in the future.

———————

It is worth noting that these systems have been built within the last five years, initiated by promising results of the first prototype scanner of the United States pCT collaboration in 2012 [110, 111]. This prototype is no longer in operation and was the precursor of the phase II scanner [81] used in this work.

3.3 DATA FILTRATION

Nuclear interactions in the object or in the energy detector can cause spurious signals and need to be filtered out to prevent a distortion of the pCT image. Such interactions violate the assumption of a continuous energy loss and of small angle scattering. For this purpose, the two scattering angles ϑ and φ are calculated as

$$\vartheta = \arccos\left(\frac{\vec{\rho}_{\text{in},u} \cdot \vec{\rho}_{\text{out},u}}{|\vec{\rho}_{\text{in},u}| \cdot |\vec{\rho}_{\text{out},u}|}\right), \tag{3.8}$$

$$\varphi = \arccos\left(\frac{\vec{\rho}_{\text{in},v} \cdot \vec{\rho}_{\text{out},v}}{|\vec{\rho}_{\text{in},v}| \cdot |\vec{\rho}_{\text{out},v}|}\right), \tag{3.9}$$

where $\vec{\rho}_{\text{in/out},u}$ is a two–dimensional vector with the u– and the d–component of the proton's direction vector at the front (in) and rear (out) tracker, respectively, and $\vec{\rho}_{\text{in/out},v}$ is the corresponding vector with the v– and the d–component. These scattering angles are used here to filter data, but they can also be used to reconstruct images of the relative scattering power [112].

Figure 3.4 (a) shows, for data of a homogeneous slab of $152\,\text{mm}$ water, distributions of the WEPL and the two scattering angles. Note, that the histograms are displayed with a logarithmic scale. Apart from the prominent central peak around a WEPL of $152\,\text{mm}$, a plateau is measured at WEPLs larger than the central peak. About $15\,\%$ of the data are outside of the central peak. For the two scattering angles, the distributions are more Gaussian and even the 95th percentile is within the central peak.

To filter out nuclear interactions, the list mode data are grouped together based on the front tracker hit location in a grid of typically $2\,\text{mm} \times 2\,\text{mm}$. Within each of these
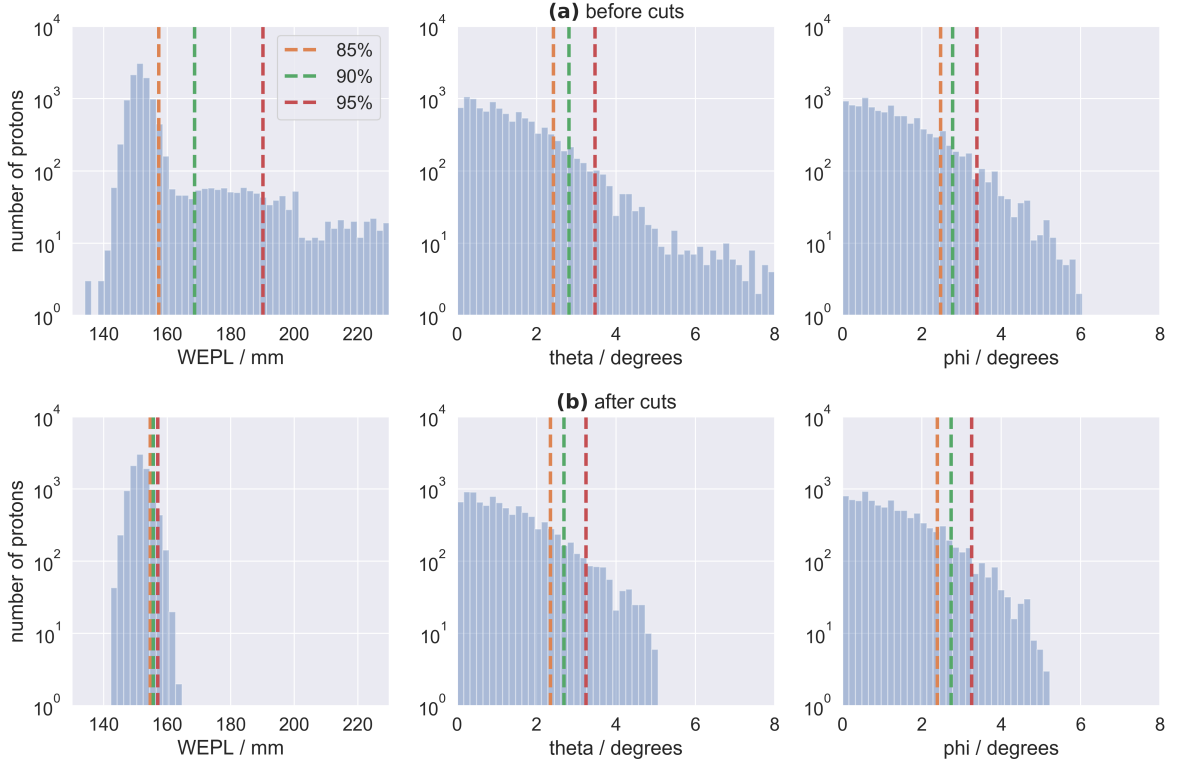
Figure 3.4: Histograms of WEPL and scattering angles ϑ and φ for a homogeneous slab of water (a) before and (b) after data cuts. The 85th, 90th and 95th percentile of the distributions is indicated by dashed lines. The histograms are displayed with a logarithmic scale for the number of protons.

bins, the 30.85th percentile and the median of the WEPL and scattering angle data are calculated. The difference between 30.85th percentile and the median corresponds to a 0.5–standard–deviation interval, which better represents the central peak and is more robust to outliers compared to a direct calculation of the standard deviation of the non–Gaussian data. Subsequently, in each bin, protons are rejected that fall out of a three standard deviation interval around the median value. For experimental data this removes about 20 % of the data in air and 15 % in a homogeneous phantom [99].

Figure 3.4 (b) shows the WEPL and angle distributions after the data cuts. A dominant filtration is observed for the WEPL data where data cuts efficiently removed all protons outside of the main peak. The percentiles as indicated by dashed lines are now all inside the main peak. For the angular data only minor changes are observed, in particular for the location of the percentiles.

3.4 PATH ESTIMATION

An important subsystem of a pCT scanner is the tracking detector, which allows to measure the protons' position and direction both before and after the scanned object. With
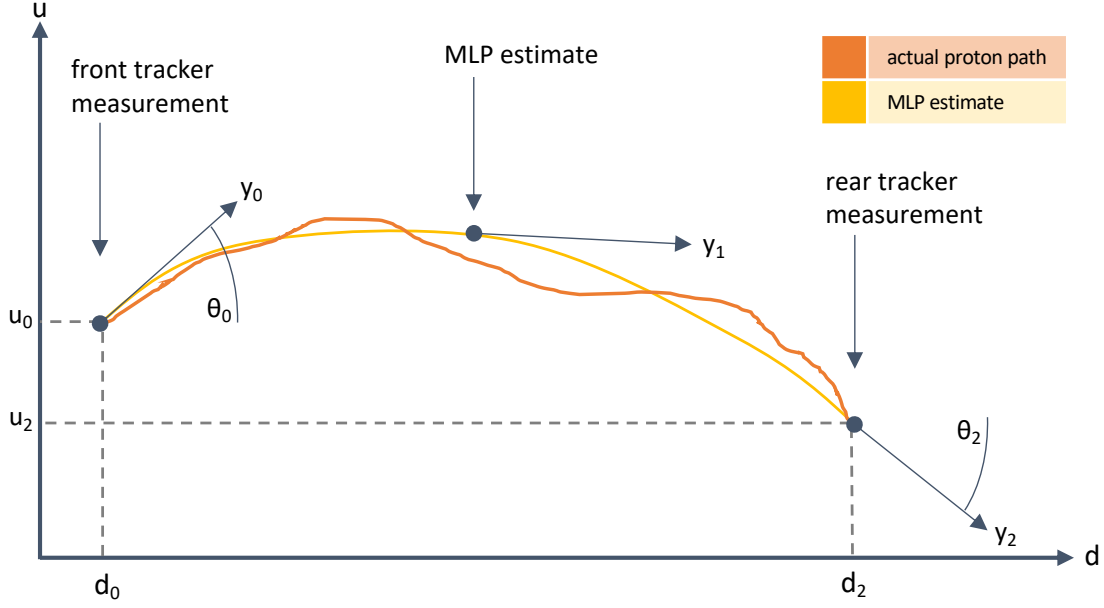
Figure 3.5: Schematic of the coordinates used in the most likely path estimation.

this information, the path of each proton can be estimated, which considerably improves image resolution [113, 114]. In this work the original path estimation algorithm [54] suggested for pCT was used. Since then, several other algorithms for path estimation have been suggested [115–119], which all have only little impact on the spatial resolution and mostly improve computation speed.

The most likely path (MLP) calculation suggested by Schulte et al. [54] makes use of the scattering matrix defined in eq. (2.11) which describes the broadening in u–direction due to scattering of a proton beam while traveling in water along a coordinate d, as depicted in fig. 3.5. Using the parameter vector $y_1 = (u_1, \theta_1)^T$ with the lateral coordinate u_1 and the scattering angle θ_1, this allows to model the likelihood of scattering from $y_0 = (0,0)^T$ to any point y_1 using eq. (2.10). This can be generalized to any entry point y_0 using the coordinate transform $y_1' = y_1 - R_0 y_0$ with the rotation matrix

$$R_0 = \begin{pmatrix} 1 & d_1 - d_0 \\ 0 & 1 \end{pmatrix}, \tag{3.10}$$

which uses the small angle approximation $\sin \theta = \theta$. With this, the likelihood to scatter from $y_0 = (u_0, \theta_0)^T$ to y_1 becomes

$$L(y_1|y_0) = \exp\left(-\frac{1}{2} \left(y_1^T - y_0^T R_0^T \right) \Sigma_1^{-1} \left(y_1 - R_0 y_0 \right) \right). \tag{3.11}$$

Analogously the likelihood of scattering from y_1 to any y_2 is defined by

$$L(y_2|y_1) = \exp\left(-\frac{1}{2} \left(y_2^T - y_1^T R_1^T \right) \Sigma_2^{-1} \left(y_2 - R_1 y_1 \right) \right), \tag{3.12}$$

$$R_1 = \begin{pmatrix} 1 & d_2 - d_1 \\ 0 & 1 \end{pmatrix}, \tag{3.13}$$

with Σ_2 defined analogously to Σ_1.

For pCT, it is known from the tracking measurements that the proton passed through y_0 and y_2 and one would like to calculate the probability that it also passed through any point y_1 with $d_0 < d_1 < d_2$. This can be expressed using the chain rule for joint probabilities as

$$L(y_1 \cap y_2 | y_0) = L(y_2 | y_1; y_0) \cdot L(y_1 | y_0) = L(y_2 | y_1) \cdot L(y_1 | y_0). \tag{3.14}$$

Using eqs. (3.11) and (3.12) this yields

$$L(y_1 \cap y_2 | y_0) = \exp\left(-\chi^2\right) \tag{3.15}$$

with

$$\chi^2 = \frac{1}{2} \left(\left(y_1^\mathsf{T} - y_0^\mathsf{T} R_0^\mathsf{T} \right) \Sigma_1^{-1} (y_1 - R_0 y_0) + \left(y_2^\mathsf{T} - y_1^\mathsf{T} R_1^\mathsf{T} \right) \Sigma_2^{-1} (y_2 - R_1 y_1) \right). \tag{3.16}$$

Differentiation of χ^2 with respect to u_1 and θ_1 gives

$$\nabla \chi^2 = \left(\Sigma_1^{-1} + R_1^\mathsf{T} \Sigma_2^{-1} R_1 \right) y_1 - \Sigma_1^{-1} R_0 y_0 - R_1^\mathsf{T} \Sigma_2^{-1} y_2. \tag{3.17}$$

This equation can be set to zero and solved for y_1 yielding an analytical expression for the MLP

$$y_{\mathrm{MLP}} = \left(\Sigma_1^{-1} + R_1^\mathsf{T} \Sigma_2^{-1} R_1 \right)^{-1} \left(\Sigma_1^{-1} R_0 y_0 + R_1^\mathsf{T} \Sigma_2^{-1} y_2 \right). \tag{3.18}$$

Equation (3.18) can be used to estimate the MLP (i.e. the coordinate u_1) at any depth d_1 in between the trackers. The dependency on d_1 is hidden inside the rotation and scattering matrices. While for simplicity the derivation was restricted to one lateral coordinate within the u–d–plane, the v coordinate is independent and can be treated analogously. Naturally, this will not be the exact proton path, but merely a best guess based on the available data. Investigations using Monte Carlo simulations have shown, that the true proton path for trajectories without nuclear interactions is within the uncertainty envelope, which can be calculated based on estimated uncertainties of the measurements y_0 and y_2 [54]. The calculation of the scattering matrices can only account for the radiation length of one scattering material, which typically is water. The formalism has been extended in other works [119–121] to account for material heterogeneities based on a first reconstruction, but the effect on the reconstructed image was reported to be small [121]. Additionally, the accuracy of the MLP can be improved if the outer hull of the object is known or can be estimated from a rough first reconstruction without path estimation. In that case the incident direction vector and the exit direction vector are projected along a straight line to the object's hull, assuming that the proton will not scatter in air.

3.5 IMAGE RECONSTRUCTION

Image reconstruction describes the mathematical process of calculating volumetric maps of a quantity given line integrals through the quantity from different angles. This inversion process can be performed directly using analytical equations or iteratively as part of an optimization. In sections 3.5.1 and 3.5.2, the analytical equations for direct image reconstruction

are derived assuming that line integrals are performed along straight lines. Section 3.5.3 generalizes the direct image reconstruction to curved proton paths and section 3.5.4 discusses how image variance maps can be calculated from list mode pCT data.

All reconstruction operations used in this work were implemented using the *Reconstruction Toolkit (RTK)* of Rit et al. [122], which itself is based on the *Insight Toolkit (ITK)* of McCormick et al. [123].

3.5.1 Image reconstruction in parallel beam geometry

The simplest case of CT image reconstruction is the so called parallel beam geometry. The image $f(x, y)$ in that case is two–dimensional and the detection system measures line integrals through $f(x, y)$ along parallel straight lines, which are parameterized by

$$x \cos \theta + y \sin \theta = \xi, \tag{3.19}$$

where θ is the rotation angle of the CT system and ξ is the coordinate along the one–dimensional detection system. The measurements $p(\theta, \xi)$ can then be described as

$$p(\theta, \xi) = \int dx dy\, f(x, y) \delta(x \cos \theta + y \sin \theta - \xi), \tag{3.20}$$

where $\delta(\cdot)$ is the Dirac delta function, which describes a ray–line at rotation angle θ and intersecting with the detector element ξ. The full set of measurements $p(\theta, \xi)$ is called *sinogram*.

To invert the integral equation one can calculate the one–dimensional Fourier transform of eq. (3.20) with respect to the coordinate ξ, which yields

$$P(\theta, u_\xi) = (\mathcal{F}p)(\theta, u_\xi) \tag{3.21}$$

$$= \int d\xi\, p(\theta, \xi) e^{-2\pi i u_\xi \xi} \tag{3.22}$$

$$= \iint dx dy\, f(x, y) e^{-2\pi i u_\xi (x \cos \theta + y \sin \theta)}, \tag{3.23}$$

$$= \iint dx dy\, f(x, y) e^{-2\pi i (u_\xi \cos \theta x + u_\xi \sin \theta y)}, \tag{3.24}$$

where u_ξ is the spatial frequency associated to ξ.

At the same time, the two–dimensional Fourier transform of the image $f(x, y)$ is

$$F(u_x, u_y) = \iint dx dy\, f(x, y) e^{-2\pi i (u_x x + u_y y)}. \tag{3.25}$$

By comparing eqs. (3.24) and (3.25), one can find the relationship

$$P(\theta, u_\xi) = F(u_\xi \cos \theta, u_\xi \sin \theta), \tag{3.26}$$

or in other words: the one–dimensional Fourier transform of the sinogram $p(\theta, \xi)$ along ξ is equal to the two–dimensional Fourier transform of the image $f(x, y)$ in polar coordinates. This relationship is referred as the *Fourier slice theorem*. It would allow to reconstruct the

image $f(x, y)$ by filling its Fourier space in polar coordinates by calculating the Fourier transform of lines of the sinogram $p(\theta, \xi)$. A resampling to Cartesian coordinates and an inverse Fourier transform then unveils the image. However, this can cause artifacts due to the resampling and is computationally expensive. Therefore, it is worthwhile to perform the coordinate change analytically, as described in the following.

To achieve this, eq. (3.26) is transformed back to position space while performing the coordinate transform $u_x = u \cos \theta$ and $u_y = u \sin \theta$ with $du_x du_y = |u| du d\theta$, and thus

$$f(x, y) = \int_0^\pi d\theta \int_{-\infty}^\infty du \; |u| P(\theta, u) e^{-2\pi i u (x \cos \theta + y \sin \theta)} \tag{3.27}$$

$$= \int_0^\pi d\theta \int_{-\infty}^\infty du \; K(u) P(\theta, u) e^{-2\pi i u \xi}, \tag{3.28}$$

where $K(u) = |u|$ is the so called *ramp kernel*. In this form the integration along u can be expressed as a convolution with the reconstruction kernel

$$k(\xi) = \mathcal{F}^{-1} K(u) = \int_{-\infty}^\infty du \; |u| e^{-2\pi i u \xi} = \frac{-1}{2\pi^2 \xi^2}, \tag{3.29}$$

and yielding the reconstruction formula

$$f(x, y) = \int_0^\pi d\theta \; p(\theta, \xi) \circledast k(\xi) \bigg|_{\xi = x \cos \theta + y \sin \theta}. \tag{3.30}$$

Equation (3.30) is called *filtered backprojection*. To reconstruct the image $f(x, y)$, the measurements $p(\theta, \xi)$ are first convolved with the reconstruction kernel $k(\xi)$. This kernel amplifies high frequencies. Subsequently, for each rotation angle θ, the filtered data are backprojected ("smeared back") into image space after rotating the data by the corresponding rotation angle. This is repeated for all rotation angles in an interval from 0 to π.

For x–ray CT, the image shows the attenuation coefficient μ and, according to eq. (2.2), the line integrals are defined as

$$p(\theta, \xi) = -\ln \frac{N}{N_0} \tag{3.31}$$

for a photon counting detector, or as intensities for an energy integrating detector. For pCT, the image shows the RSP and, according to eq. (3.2), the line integrals are defined as

$$p(\theta, \xi) = \text{WEPL}. \tag{3.32}$$

The reconstruction algorithm in eq. (3.30) is typically discretized, which, following the notation of Rädler et al. [124], results in

$$f(x, y) = \frac{\pi}{N_P} \sum_{n=1}^{N_P} h(\theta_n, x \cos \theta_n + y \sin \theta_n), \tag{3.33}$$

where N_P is the total number of acquired projections, θ_n is the projection angle corresponding to the n–th projection. The $h(\theta, \xi)$ are the filtered projection values with a discretized detector location $\xi = j\Delta\xi$ and

$$h(\theta_n, j\Delta\xi) = \Delta\xi \sum_{m=-D/2}^{D/2-1} p(\theta_n, m\Delta\xi) \cdot g((j-m)\Delta\xi), \tag{3.34}$$

where $g(\xi)$ is the discretized ramp filter and $D \in 2\mathbb{N}$ its finite extent. The most common discretization is the one of Ramachandran and Lakshminarayanan [125], which is calculated as

$$g(j\Delta\xi) = \begin{cases} (2\Delta\xi)^{-2} & \text{if } j = 0, \\ 0 & \text{if } j \text{ even}, \\ -(j\pi\Delta\xi)^{-2} & \text{if } j \text{ odd.} \end{cases} \tag{3.35}$$

Equation (3.33) queries the filtered projections in eq. (3.34) at $\xi_n = x\cos\theta_n + y\sin\theta_n$, but the reconstruction grid in x and y does not always coincide with the center of a detector pixel in ξ (in fact, it only does for few angles). For this, an interpolation between adjacent detector pixels at $j\Delta\xi$ and $(j+1)\Delta\xi$ is necessary. This interpolation is particularly important in this work, since it affects image noise [126]. The final discretized filtered backprojection formula becomes

$$f(x, y) = \frac{\pi\Delta\xi}{N_P} \sum_{n=1}^{N_P} \sum_{m=-D/2}^{D/2-1} p(\theta_n, m\Delta\xi) \cdot \Big[g((j-m)\Delta\xi) \cdot (1-a) $$
$$+ g((j+1-m)\Delta\xi) \cdot a \Big] \tag{3.36}$$

with the interpolation weights

$$a(\xi_n) = \frac{\xi_n - j\Delta\xi}{\Delta\xi}. \tag{3.37}$$

3.5.2 Image reconstruction in cone beam geometry

If the rays between the source and the detector are not parallel as in the parallel beam geometry, the reconstruction algorithm needs to be modified. In this work a cone beam geometry will be considered, where the flat detector is two–dimensional and extends both in the u and v direction. The rays are assumed to diverge from a common source point and the opening angle in u direction is called the *fan angle* α, whereas the opening angle in v direction is the *cone angle* γ. Images are then directly reconstructed into the volume with the coordinates x, y and z instead of reconstructing the volume for each slice (for each value of z) independently.

The Feldkamp–Davis–Kress (FDK) algorithm [127] generally uses eq. (3.36) for reconstruction with modifications to account for the cone–beam geometry. The fan angle can simply be taken into account by resorting the raw data, since a ray from a projection with rotation angle θ and a fan angle α is equivalent to a ray from rotation angle $\theta + \alpha$ and no fan angle. For the cone angle, an additional pre–factor $\cos\gamma$ needs to be considered, accounting for the

fact that rays close to the source are more densely sampling the volume. In general, however, it is often useful to think of the reconstruction process as if it were performed in parallel beam geometry.

3.5.3 Image reconstruction using distance–driven binning

In this work FDK image reconstruction with distance–driven binning (FDK–DDB) was used to produce volumetric pCT images from the WEPL data. The FDK–DDB algorithm proposed by Rit et al. [113] extends the FDK algorithm to account for the curved path of protons, which can be estimated from the tracking information. Depth in this context refers to the depth coordinate d in between the trackers and along the beam direction and binning describes the process of creating pixelated projection values $p(\theta, \xi, d)$ from the list mode proton data. Again, this will be introduced in parallel beam geometry to simplify the notation, with a subsequent extension to cone–beam geometries.

For each proton and each depth d, the protons lateral coordinate ξ can be calculated by evaluating eq. (3.18). To calculate the projection values, for each projection angle θ a two–dimensional virtual detection grid is established with bins of the size $\Delta\xi \times \Delta d$. This pixel size can be chosen by the user. Each pixel in that grid is assigned a subset of the proton list mode data to which all protons belong that intersect the virtual pixel at $\xi = j\Delta\xi$ and $d = k\Delta d$. For each bin, the projection value calculates as

$$p(\theta, \xi, d) = \frac{1}{N_{\theta,\xi,d}} \sum_{i=1}^{N_{\theta,\xi,d}} \text{WEPL}_i \ , \tag{3.38}$$

where $N_{\theta,\xi,d}$ is the number of protons intersecting the bin and WEPL_i their WEPL value. Assuming that the MLP best describes the true proton path, the projection values at distance d produce a projection image which is least affected by blurring due to MCS for contrasts located at depth d.

This can be considered in the filtered backprojection by assigning an optimal depth d_{DDB} to each combination of rotation angle θ and image pixel (x, y). The optimal depth is the depth that a proton needs to travel to reach the pixel (x, y) given that it travels at a rotation angle θ, and can be calculated as

$$d_{\text{DDB}} = x \sin\theta + y \cos\theta. \tag{3.39}$$

With this, eqs. (3.33) and (3.34) become

$$f(x, y) = \frac{\pi}{N_P} \sum_{n=1}^{N_P} h(\theta_n, x\cos\theta_n + y\sin\theta_n, x\sin\theta_n + y\cos\theta_n), \tag{3.40}$$

$$h(\theta_n, j\Delta\xi, k\Delta d) = \Delta\xi \sum_{m=-D/2}^{D/2-1} p(\theta_n, m\Delta\xi, k\Delta d) \cdot g((j-m)\Delta\xi). \tag{3.41}$$

It is important to note, that the basic filtered backprojection algorithm is not modified, but merely repeated at every binning depth. Additionally, a second interpolation is needed to

query d_{DDB} in steps $k\Delta d$. However, this interpolation is also omissible and can be replaced by the nearest neighbor of the query point, since projection values only slowly change with the depth d.

The extension of the FDK–DDB algorithm to a cone–beam geometry is comparable to the extension of filtered backprojection to FDK. The projections for each rotation angle will have three dimensions, namely u and v as lateral coordinates and d as the binning depth. The proton's MLP estimate will consequently also model a two–dimensional track with coordinates u and v at every binning depth d. To account for the divergence of protons, the virtual detector grid will need to be scaled in u and v direction with the binning depth d to allow for a constant fluence at each depth. This can be done by setting $\Delta u(d) = \rho(d) \cdot \Delta u_0$ and $\Delta v(d) = \rho(d) \cdot \Delta v_0$ with reference voxel sizes Δu_0 and Δv_0 and

$$\rho(d) = \frac{d + SID}{SDD}, \tag{3.42}$$

where SID is the distance of the cone beam source to the image's isocenter and SDD is the distance between the source and the depth at which $\Delta u(d) = \Delta u_0$, which typically is the physical location of the rear tracker measurement.

In the initial study, Rit et al. [113] reported an improvement of the spatial resolution which reduced the blurring around inserts from 3 mm when assuming straight rays down to 1.5 mm when using FDK–DDB. The blurring in this case was measured for ideal data and at the center of a 200 mm diameter water phantom with aluminum inserts. Even though the FDK–DDB is fairly simple and elegant, it showed competitive results in comparative studies [114, 128]. Several other reconstruction algorithms have been proposed for pCT including direct [129–131] and iterative [128, 132–135] algorithms.

3.5.4 Reconstruction of image variance

Within this work the reconstruction of image variance $\text{Var}[f]$ is essential for the fluence control algorithms to ensure that the variance level is at the desired level. The image variance in a voxel (x, y, z) can be calculated by repeating an experiment or simulation and the corresponding reconstructions N times, where $N \gg 10$, with a subsequent calculation of the variance of RSP values voxel–by–voxel as

$$\text{Var}[f(x, y, z)] = \frac{1}{N} \sum_{n=1}^{N} (f_n(x, y, z) - \mu(x, y, z))^2 \tag{3.43}$$

where f_n are the statistically independent realizations of the image f and

$$\mu(x, y, z) = \frac{1}{N} \sum_{n=1}^{N} f_n(x, y, z) \tag{3.44}$$

is the expectation value. This method, however, is often not feasible due to the time requirements for repeating a simulation or experiment N times and also running the reconstruction algorithm for each of the datasets. To enable a fast optimization of imaging fluence, variance is ideally calculated on a single dataset with a computation effort similar to

a standard image reconstruction. This is possible using variance reconstruction [136], which was originally proposed for x–ray CT early on [137].

Rädler et al. [124] applied variance reconstruction to pCT imaging, which again is introduced here in parallel beam geometry with a subsequent generalization to the FDK–DDB algorithm. The variance of the projection values, $\sigma_p^2(\theta, \xi)$ can directly be estimated from the list mode data as the standard error of the mean

$$\sigma_p^2(\theta, \xi) = \frac{\sigma_{WEPL}^2(\theta, \xi)}{N_{\theta, \xi}}, \tag{3.45}$$

where $\sigma_{WEPL}^2(\theta, \xi)$ is the variance of the WEPL values of all protons intersecting the detector element ξ at rotation angle θ and $N_{\theta, \xi}$ is their number. This calculation makes no assumption about the origin of the projection noise, but directly calculates it based on the available list mode data.

To calculate the image variance, the projection values can be treated as independent random variables. Equation (3.36) consequently is a weighted sum of M random variables X_i, whose variance can be calculated, in general terms, as

$$\text{Var}\left[\sum_{i=1}^{M} c_i X_i\right] = \sum_{i, i'=1}^{M} c_i c_{i'} \text{Cov}[X_i, X_{i'}], \tag{3.46}$$

where $\text{Cov}[X_i, X_{i'}]$ is the covariance matrix, which for $i = i'$ becomes $\text{Cov}[X_i, X_i] = \text{Var}[X_i]$ and c_i and $c_{i'}$ are arbitrary prefactors.

For eq. (3.36) the random variable X are the projection values p. The sum over M elements is expressed as three interleaved summations: the sum with the variable n over all rotation angles, the sum with the variable m over the filter extent and then the interpolation sum over just two summands denoted in square brackets in eq. (3.36).

In the following, it is assumed that there is no covariance between projection values and between rotation angles. Then, the covariance matrix is

$$\text{Cov}\left[p(\theta, \xi), p(\theta', \xi')\right] = \begin{cases} \sigma_p^2(\theta, \xi) & \text{if } \theta = \theta' \text{ and } \xi = \xi' \\ 0 & \text{otherwise} \end{cases} \tag{3.47}$$

Note, that this does *not* mean that $\text{Cov}[X_i, X_{i'}] = 0$ for $i \neq i'$, since in eq. (3.46) two different values of i and i' can refer to the same projection value. In fact, if eq. (3.36) is expressed with eq. (3.46), each projection value p occurs twice in the sum over M summands: once with the prefactor $c_\downarrow$ and once with the prefactor $c_\uparrow$, which are defined as

$$c_\downarrow = \frac{\pi \Delta \xi}{N_P} \cdot g((j - m) \cdot \Delta \xi) \cdot (1 - a), \tag{3.48}$$

$$c_\uparrow = \frac{\pi \Delta \xi}{N_P} \cdot g((j + 1 - m)\Delta \xi) \cdot a. \tag{3.49}$$

Consequently, each σ_p^2 in eq. (3.47) then needs to be considered four times in eq. (3.46): once with the prefactor $c_\downarrow^2$, once with the prefactor $c_\uparrow^2$ and two times with the prefactor $c_\downarrow \cdot c_\uparrow$. With this, the image variance can be calculated as

$$\text{Var}\,[f(x,y)] = \sum_{n=1}^{N_p} \sum_{m=-D/2}^{D/2-1} \Big\{ c_\downarrow^2 \cdot \sigma_p^2(\theta_n, m\Delta\xi)$$
$$+ 2 \cdot c_\downarrow c_\uparrow \cdot \sigma_p^2(\theta_n, m\Delta\xi) \tag{3.50}$$
$$+ c_\uparrow^2 \cdot \sigma_p^2(\theta_n, m\Delta\xi) \Big\}.$$

In full, this leads to the variance reconstruction formula

$$\text{Var}\,[f(x,y)] = \left(\frac{\pi\Delta\xi}{N_P}\right)^2 \sum_{n=1}^{N_P} \Big\{ (1-a)^2 \cdot V(\theta_n, j\Delta\xi)$$
$$+ 2(1-a)a \cdot C(\theta_n, j\Delta\xi, (j+1)\Delta\xi) \tag{3.51}$$
$$+ a^2 \cdot V(\theta_n, (j+1)\Delta\xi) \Big\},$$

where V and C are variance and covariance terms which are defined as

$$V(\theta_n, j\Delta\xi) = \sum_{m=-D/2}^{D/2-1} g^2((j-m)\Delta\xi) \cdot \sigma_p^2(\theta_n, m\Delta\xi), \tag{3.52}$$

$$C(\theta_n, j\Delta\xi, (j+1)\Delta\xi) = \sum_{m=-D/2}^{D/2-1} g((j-m)\Delta\xi)g((j+1-m)\Delta\xi) \cdot \sigma_p^2(\theta_n, m\Delta\xi). \tag{3.53}$$

The expression in eq. (3.51) correctly considers the effect of interpolation on image variance for each pixel separately. For example, a pixel located at the center of the image could always be aligned with the central detector pixel and thus always be assigned $a = 0$. Its variance would, therefore, be higher than another pixel where $0 < a < 1$, since the covariance terms C are always negative for the filter g defined in eq. (3.35).

Equation (3.52) is a convolution with the variance filter g^2 (which is the square of the ramp filter g) and as Wunderlich and Noo [136] suggest, eq. (3.53) can also be written as a convolution using the covariance filter

$$g_C(j\Delta\xi) = g(j\Delta\xi) \cdot g((j+1)\Delta\xi), \tag{3.54}$$

which then results in

$$C(\theta_n, j\Delta\xi, (j+1)\Delta\xi) = \sum_{m=-D/2}^{D/2-1} g_C((j-m)\Delta\xi) \cdot \sigma_p^2(\theta_n, m\Delta\xi). \tag{3.55}$$

As Rädler et al. [124] suggest, it is instructive to ignore the local pattern caused by interpolation and only consider the average effect of interpolation based on the average

value of a over all pixels, which simplifies the reconstruction formula. For this, it is assumed that the projection variance values σ_p^2 are locally constant, which leads to the approximation

$$V(\theta_n, j\Delta\xi) \approx \sigma_p^2(\theta_n, j\Delta\xi) \cdot \sum_{m=-D/2}^{D/2-1} g^2(m\Delta\xi) \approx \frac{\sigma_p^2(\theta_n, j\Delta\xi)}{12(\Delta\xi)^4}, \tag{3.56}$$

$$C(\theta_n, j\Delta\xi, (j+1)\Delta\xi) \approx \sigma_p^2(\theta_n, j\Delta\xi) \cdot \sum_{m=-D/2}^{D/2-1} g_C(m\Delta\xi) \approx -\frac{\sigma_p^2(\theta_n, j\Delta\xi)}{2\pi^2(\Delta\xi)^4}. \tag{3.57}$$

The change of j and m to query the variances is allowed here, since it is equivalent to convolve the projection with the filter or the filter with the projection. The second part of each equation was found by assuming that m goes from $-\infty$ to ∞, which is a valid assumption since the filter g falls quickly to zero.

With that, eq. (3.51) can be factorized as

$$\text{Var}\left[f(x,y)\right] = \left(\frac{\pi\Delta\xi}{N_P}\right)^2 \sum_{n=1}^{N_P} \underbrace{\frac{\sigma_p^2(\theta_n, j\Delta\xi)}{12(\Delta\xi)^4}}_{V(\theta_n, j\Delta\xi)} \underbrace{\left\{(1-a)^2 + 2\eta(1-a)a + a^2\right\}}_{f_{\text{interp}}(a)}, \tag{3.58}$$

where

$$\eta = \frac{C(\theta_n, j\Delta\xi, (j+1)\Delta\xi)}{V(\theta_n, j\Delta\xi)} \approx -\frac{12}{2\pi^2}. \tag{3.59}$$

The part in curly brackets in eq. (3.58) is defined as $f_{\text{interp}}(a)$ and describes the effect of interpolation on image noise. Assuming that a is distributed uniformly in the interval $[0, 1]$, the term can be reduced to a prefactor as

$$f_{\text{interp}} = \int_0^1 da\, f_{\text{interp}}(a) = 2/3 - 2/\pi^2 \approx 0.464. \tag{3.60}$$

The variance reconstruction formula then simplifies to

$$\text{Var}\left[f(x,y)\right] = f_{\text{interp}} \left(\frac{\pi\Delta\xi}{N_P}\right)^2 \sum_{n=1}^{N_P} V(\theta_n, j\Delta\xi) \tag{3.61}$$

$$= f_{\text{interp}} \left(\frac{\pi\Delta\xi}{N_P}\right)^2 \sum_{n=1}^{N_P} \sum_{m=-D/2}^{D/2-1} g^2((j-m)\Delta\xi) \cdot \sigma_p^2(\theta_n, m\Delta\xi), \tag{3.62}$$

where in eq. (3.62) the simplification of eq. (3.56) was undone by replacing $V(\theta_n, j\Delta\xi)$ by its original definition in eq. (3.52). The covariance terms $C(\theta_n, j\Delta\xi, (j+1)\Delta\xi)$ did not vanish, but were merely expressed as multiples of $V(\theta_n, j\Delta\xi)$ under the assumption of locally constant variance values.

Interestingly, eq. (3.62) is very similar to eq. (3.36), apart from the different pre–factor and the fact that the variance projection values are convolved with the squared elements of the reconstruction filter in eq. (3.35). While a convolution with the reconstruction filter

$g(\xi)$ steepens edges in the image and amplifies high spatial frequencies, the variance reconstruction filter $g^2(\xi)$ only has a mild blurring effect.

As a good approximation, the convolution can completely be disregarded and replaced by a constant, resulting in the formula

$$\text{Var}\left[f(x,y)\right] = f_{\text{interp}} \cdot f_{\text{filter}} \left(\frac{\pi \Delta \xi}{N_P}\right)^2 \sum_{n=1}^{N_P} \sigma_p^2(\theta_n, m\Delta\xi), \tag{3.63}$$

where the additional constant calculates as

$$f_{\text{filter}} = \sum_{j=-\infty}^{\infty} g^2(j\Delta\xi) = \frac{1}{(2\Delta\xi)^4} + \frac{2}{(\pi\Delta\xi)^4} + \frac{2}{(3\pi\Delta\xi)^4} + \frac{2}{(5\pi\Delta\xi)^4} + \ldots \approx \frac{1.33}{(2\Delta\xi)^4}, \tag{3.64}$$

and which is equivalent to the assumption in eq. (3.56). With this, the variance reconstruction is a simple backprojection operation of the projection variance values. This particular simplification was suggested for the work presented in chapter 8 and in Dickmann et al. [138], but similar approaches were used in previous works [139]. If not otherwise stated, variance in this work is reconstructed using eq. (3.62) and eq. (3.64) is only used in chapter 8.

The extension of eqs. (3.51), (3.62) and (3.63) to the FDK–DDB algorithm is straightforward. The additional term for the cone angle must be squared, since it occurs both in $c_\downarrow$ and $c_\uparrow$. The extension of the projection to three dimensions, including the binning depth, involves another interpolation. However, Rädler et al. [124] suggest to consider the projection to be constant with the binning depth (which is a good approximation since protons travel along the binning depth) and perform a nearest neighbor interpolation, which does not need to be considered with additional co–variance terms.

3.5.5 Image reconstruction parameters

Unless otherwise stated, all reconstructions presented in this work employed a projection grid of $280 \times 80 \times 280$ voxels for distance–driven binning with a uniform grid size of $1\,\text{mm}$ in u, v and d. The reconstruction volume also had $280 \times 280 \times 80$ voxels with a uniform voxel size of $1\,\text{mm}$ in x, y and z, where z is the coordinate along slices which is aligned with v. For data filtration, the list mode data was grouped in bins of $2\,\text{mm} \times 2\,\text{mm}$ at the front tracker. The number of projections was 90 in chapters 6 and 7 and 360 in chapters 5, 8 and 9. The reduced number of projections in chapters 6 and 7 was required to allow for a feasible acquisition time in proof–of–concept experimental step–and–shoot acquisitions. It may be possible to reduce the waiting time in between two projections in future acquisitions as discussed in chapter 7, allowing for more projections to be acquired, which is beneficial for image quality.
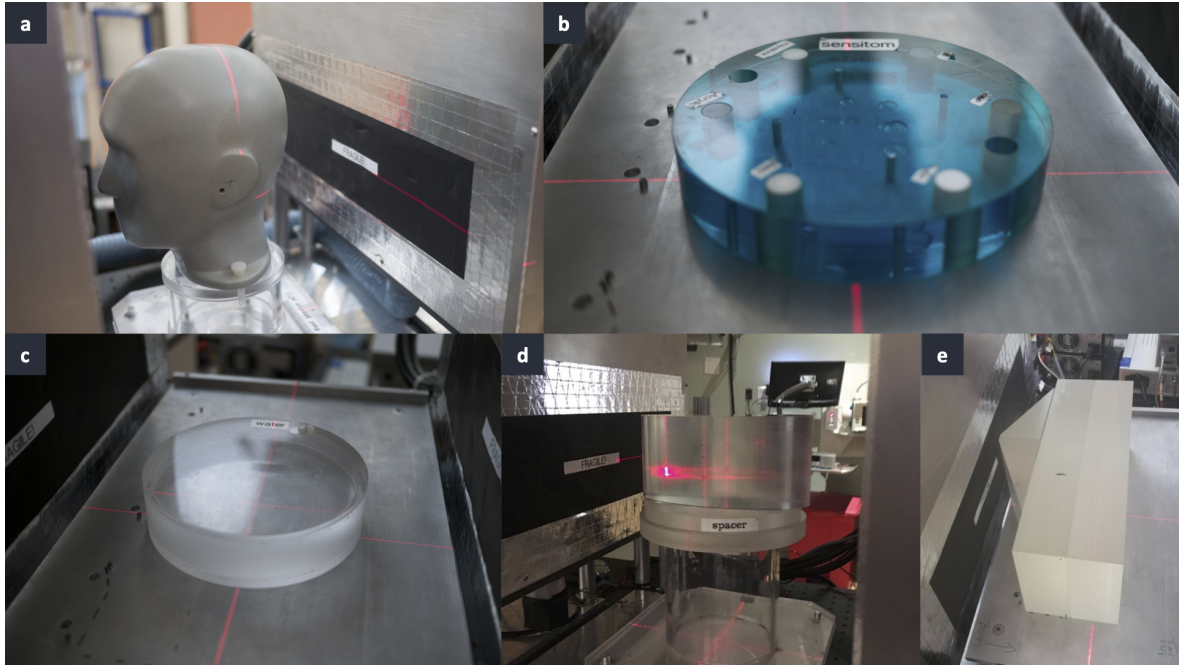
3.6 PHANTOMS

Within this work several phantoms were used both for calibration of the scanner and evaluation of its imaging performance. Details on the phantoms are given below as well as in table 3.1. Images of all phantoms while mounted in the pCT scanner are shown in fig. 3.6.

Table 3.1: Dimensions, materials and RSP values of all phantoms used in this work. RSP values were determined by water–column measurements. *Abbreviations:* PMMA – polymethyl-methacrylate/acrylic, PMP – polymethylpentene, LDPE – low–density polyethylene.

phantom	shape & dimensions		materials	RSP
calibration	double–wedge and bricks:		RSP value from [56]:	
	– brick thickness	50.8 mm	– polystyrene	1.030
	– brick width	320.0 mm		
	– brick height	120.0 mm		
water	PMMA cylinder with water filling:		RSP values from [100]:	
	– outer diameter	150.5 mm	– purified water	1.000 [a]
	– wall thickness	6.3 mm	– PMMA	1.170 [b]
	– lid thickness	6.3 mm		
	– height	40.0 mm		
CTP	cylindrical body with inserts:		RSP values from [103]:	
	– outer diameter	150.0 mm	– PMP	0.883
	– insert diameter	12.0 mm	– LDPE	0.979
	– height	25.0 mm	– polystyrene	1.024
			– body/epoxy	1.144
			– PMMA	1.160
			– Delrin	1.359
			– Teflon	1.790
head	anatomical model of a five–year–old pediatric head using tissue–equivalent materials		RSP values from [140]:	
			– soft tissue	1.032
			– brain tissue	1.044
			– spinal disc	1.069
			– trabecular bone	1.111
			– cortical bone	1.331
			– tooth dentin	1.524
			– tooth enamel	1.651
ellipse	homogeneous elliptical body:		RSP value from [141]:	
	– long axis	165.0 mm	– PMMA	1.160
	– short axis	80.0 mm		
	– height	80.0 mm		

[a] by definition, [b] estimated from a pCT scan

Figure 3.6: Images of all phantoms used in this work: (a) head phantom, (b) CTP phantom, (c) water phantom, (d) ellipse phantom, and (e) calibration phantom.

CALIBRATION PHANTOM   The calibration phantom consists of a double–wedge with additional bricks that can be added. It scans the dynamic WEPL range of the scanner and is used for its calibration as described in section 3.2.3.

WATER PHANTOM   The water phantom is a polystyrene cylinder with a lid that contains purified water. It can be used to verify the homogeneity of the detector response and its RSP value is by definition equal to 1.

CTP PHANTOM   The CTP404 module of the Catphan® 600 phantom (*Phantom Laboratory*, New York, USA) is referred to as CTP phantom within this work. It consists of a cylindrical body with six cylindrical inserts of different materials as well as two air–filled inserts. The body is made from epoxy and the inserts are polymethylpentene (PMP), low–density polyethylene (LDPE), polystyrene, polymethylmethacrylate/acrylic (PMMA), Delrin and Teflon.

HEAD PHANTOM   The pediatric head phantom (ATOM®, Model 715 HN, *CIRS Inc.*, Norfolk, USA) is a commercial anthropomorphic phantom that mimics the head of a five–year old child using tissue–equivalent materials.

ELLIPSE PHANTOM   The ellipse phantom is a custom–built phantom made from PMMA that was used in section 9.4 to reduce image artifacts of the pCT scanner.

3.7 EXPERIMENTAL ACQUISITIONS

For this work, several experimental acquisitions using the pCT prototype scanner were undertaken, which are listed in table 3.2. All acquisitions took place in the Northwestern Medicine Chicago proton center in Warrenville (Illinois), USA. The table indicates the date of the data acquisition as well as the beam type and phantoms used. For fluence–modulated scans, which will be used in chapters 6 to 8, data was acquired using a grid of small pencil beams for which the intensity of each beam could be modulated. Typical pCT scans, however, employ a broad beam with a FWHM of about 65 mm, which scans across the field–of–view of the scanner. This beam type was used for the investigations in chapters 5 and 9. The beam energy typically is 200 MeV, but was reduced in some beamtimes to 187.5 MeV to study image artifacts of the scanner. The beam current is usually fine–tuned to bring the data acquisition rate below the 1 MHz, which is viable for the scanner. To generally allow

Table 3.2: List of beamtimes for the experimental acquisition of data used in this work with indication of the beam type and phantoms used. The chapter where data of the beamtimes was used is indicated. *Abbreviations:* PB – pencil beam

date	beam type & phantoms	energy in MeV	current in nA	slit settings* in mm	chapter
2016-08-13[†]	broad beam water, head, CTP	200	—	—	5
2018-09-11	PB unit fluence head	200	2.4	1.1/1.1/3.0/3.0	6
2018-09-12	PB unit fluence water, CTP	200	2.4/5.0	1.1/1.1/3.0/3.0	6
2018-09-13	broad beam	*support for other beamtime*			
2019-07-25	PB modulated	*no scans, issue with tracking detector*			
2019-07-26	PB modulated	*test scans to fix issue with tracking detector*			
2019-07-27	PB modulated water, head, CTP	200	1.3	2.0/2.0/3.0/3.0	7
2019-07-29	broad beam water	187.5/200	2.5	1.7/1.5/5.0/5.0	9.3
2019-07-30	PB unit fluence water	200	1.3	2.0/2.0/3.0/3.0	7
2020-10-29[‡]	broad beam ellipse, water, head, CTP	187.5/200	1.75	1.5/1.5/5.0/5.0	9.4

* horizontal/vertical divergence and horizontal/vertical momentum slit setting given in this order

† data used in this work came from an earlier beamtime, which the author did not attend

‡ the author attended the beamtime virtually via videoconference

for such low proton fluences, both divergence slits and momentum slits of the accelerator need to be closed to a setting beyond what is typically used for treatment. These settings are also listed in table 3.2. Eventually, the table refers the reader to the chapter of this work where the data of a given beamtime was used.

4

COMPUTATIONAL MODELS

Monte Carlo simulations are the gold standard for the in–silico investigation of the performance of imaging setups and were used extensively in this work. In section 4.1, a specific Monte Carlo code developed by Giacometti et al. [103] is discussed, that models the full geometry of the prototype pCT scanner used in this work. Giacometti et al. [140] also developed a voxelized version of the head phantom, which is presented in section 4.2. The usage of patient data in pCT simulations as it was needed in chapter 8 and Dickmann et al. [138] is detailed in section 4.3.

4.1 MONTE CARLO SIMULATIONS

Imaging setups, and in particular pCT systems, are very complex, making it challenging to model their properties under realistic conditions using analytical equations as those introduced in chapters 2 and 3. For this purpose, the gold standard is to use Monte Carlo simulations, which help to solve the strongly coupled system of possible physical interactions inside a pCT scanner. Monte Carlo codes for particle transport employ tabulated values and analytical models of randomly sampled particle–matter interactions to simulate the transport of individual particles within a user–defined geometry. The path and interaction of secondary particles that are generated during the interactions of the primary particle is subsequently simulated. The simulation is repeated for more primary particles until sufficient statistics is generated for user–defined output variables, such as the dose to a given volume or the charges deposited in it. One of the most widespread Monte Carlo codes used in particle physics is Geant4 [61], originally developed at CERN. Other common Monte Carlo codes are FLUKA [142] as well as the two applications TOPAS [143] and GATE [144], which themselves are based on Geant4.

Giacometti et al. [103] have used Geant4 to create a detailed model of the prototype pCT scanner used in this work. The code models all components of the scanner, including the tracking detectors and the energy detector. A rendering of the three–dimensional model is shown in fig. 4.1. The individual silicon strips of the tracking detector are simulated and used to digitize the position measurement. In the energy detector, the five stages are modeled as individual volumes. To improve the computation speed of the simulation, light transport in the scintillator is not simulated directly, but the energy deposited in each stage is converted to realistic measurements using an analytical model. All output by the simulation is in the same
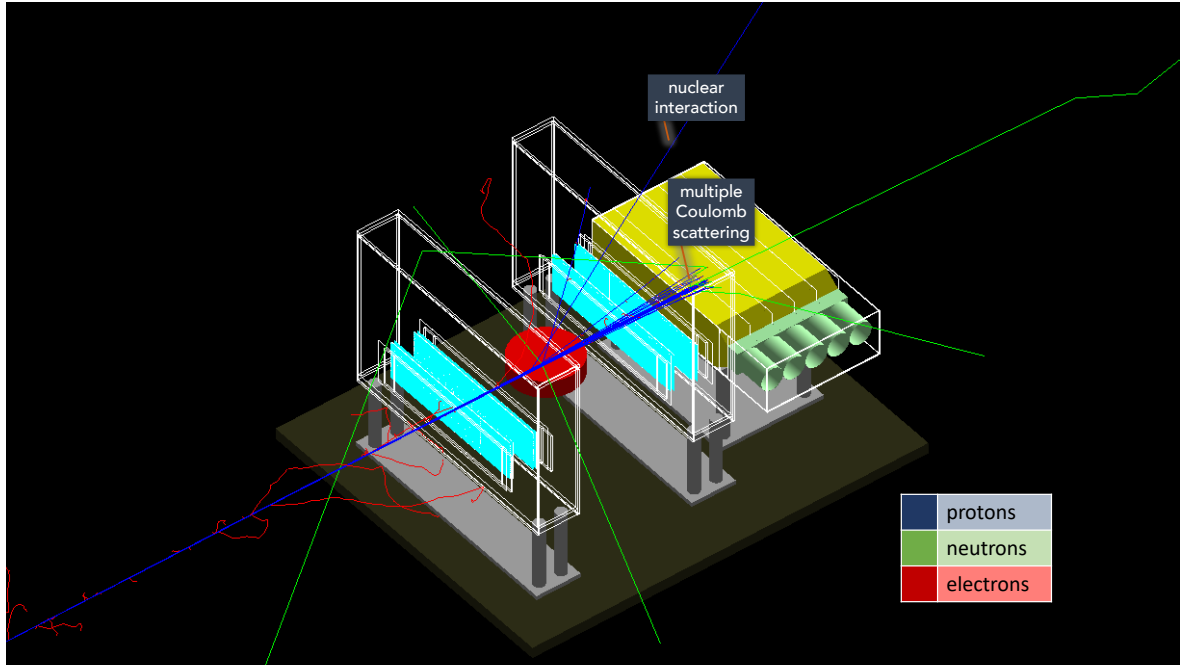
Figure 4.1: Monte Carlo model [103] of the prototype pCT scanner with a cylindrical phantom in place. One proton that underwent a nuclear interaction within the phantom is labeled. The primary proton beam reaching the detector is broadened due to MCS occurring both in the phantom and in the detector.

format as for the original scanner and is processed in the same way, including the calibration using the double–wedge phantom. This allows for a highly realistic generation of synthetic pCT data and was used in numerous studies to investigate the imaging performance of the scanner, as summarized by Dedes et al. [102]. The simulation allows to place and rotate digital phantoms modeling their physical counterparts. The phantoms are implemented based on their known geometry and composition and using analytical shapes. All phantoms in section 3.6 and more are implemented allowing to compare results to experiments. All RSP values in table 3.1 are reproduced with an accuracy of better than 0.1 % by fine–tuning the mean ionization potential $\langle I \rangle$ of the respective materials. This material constant is defined in eq. (2.2) and used to calculate the energy loss in the Monte Carlo simulation. The ionization potential of water was set to $78.0\,eV$ according to the latest recommendations of the International Commission on Radiation Units and Measurements (ICRU).

To allow for a reasonable computation speed, the simulation code only accounts for those physical processes and particle energies that are relevant for the application. The physical processes are selected in GEANT4 by pre–defined physics lists. In particular, for the simulation code of Giacometti et al. [103], the *Livermore EM* physics list is used to model electromagnetic interactions. In addition, the *G4HadronPhysicsQGSP_BIC_HP* and the *G4HadronElasticPhysicsHP* were enabled to model inelastic and elastic scattering of hadrons using tabulated cross sections. The *neutron high precision* model was used which is validated for neutron interactions from thermal neutrons and up to $20\,MeV$. The *G4IonBinaryCascadePhysics* list modeled ion hadronic interactions. To avoid calculating parti-

cles that would not impact the scoring results (i. e. that would not reach the energy detector), range cuts were set outside of the detector that prevented the production of particles with a range of 10 mm or less. In that case, the secondary's energy is deposited locally instead. This cut was reduced to 5 μm in the detector.

4.2 MODEL OF THE PEDIATRIC HEAD PHANTOM

Giacometti et al. [140] also realized a voxelized representation of the pediatric head phantom, where an implementation with analytical shapes was not feasible due to the complexity of the phantom. They scanned the phantom with a high–resolution protocol of a clinical x–ray CT scanner. The CT image was thresholded to get connected regions belonging to a list of known materials. For each material listed in table 3.1, the material composition was known from the vendor of the head phantom and it could be implemented as a material in GEANT4. After manual geometry corrections, this resulted in a three–dimensional voxelized representation of the head phantom as shown in fig. 4.2 with known materials and densities. The voxel size was 0.19 mm × 0.19 mm and the slice thickness 1.25 mm. The voxel size is considerably lower than the pCT scanner's spatial resolution, but was yet required to correctly reproduce experimental scans. The mean RSP error using the high–resolution phantom was at only 0.7 % [140].
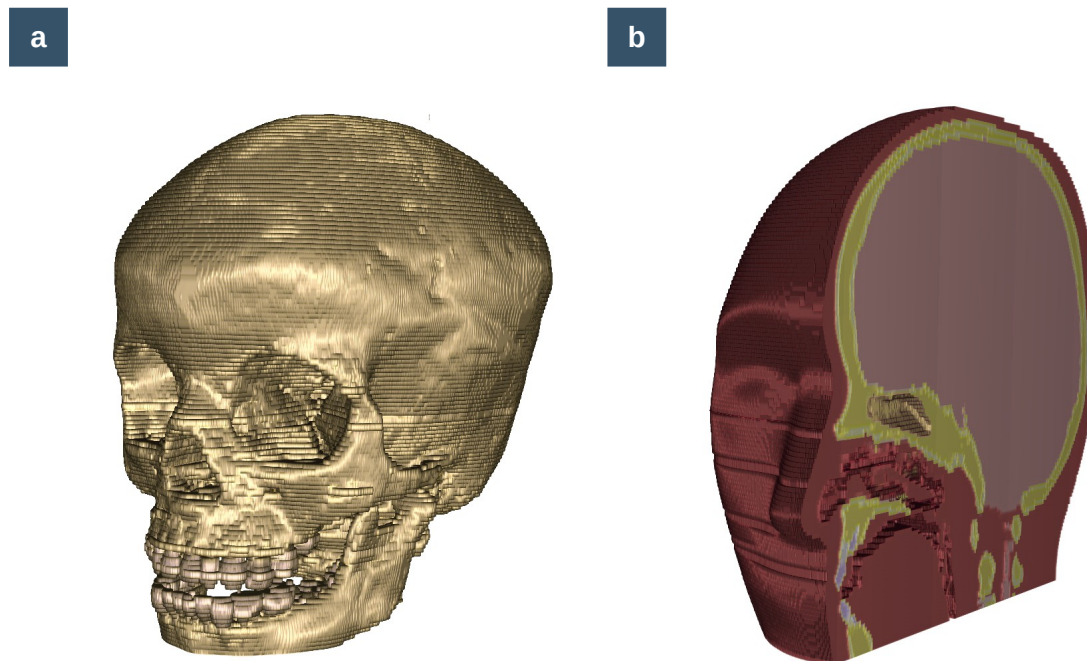
Figure 4.2: Volume rendering of the voxelized representation of the pediatric head phantom of Giacometti et al. [140] displaying (a) the bone volume and (b) a cross–section in the anterior–posterior plane.

4.3 PATIENT DATA IN SIMULATION STUDIES

In chapter 8 and in Dickmann et al. [138], a simulation study was performed based on x–ray CT images of patients that underwent radiotherapy. For those patients only CT values in the HU scale were available. A simulation of pCT data using the Monte Carlo simulation of Giacometti et al. [103], however, requires knowledge of a volumetric map of material compositions and densities. Since CT values are correlated with the material density, it is possible convert CT values to material densities using a piecewise–linear calibration curve as shown in fig. 4.3 (a). In a subsequent calibration shown in fig. 4.3 (b), discrete bins of material densities are assigned elemental compositions that are typically encountered in human tissue of that density [145, 146]. This allowed to simulate pCT scans based on a treatment planning x–ray CT.

The inverse calibration from RSP maps to CT values was required to use simulated pCT scans in a TPS to verify their accuracy for proton range calculations. This was done in
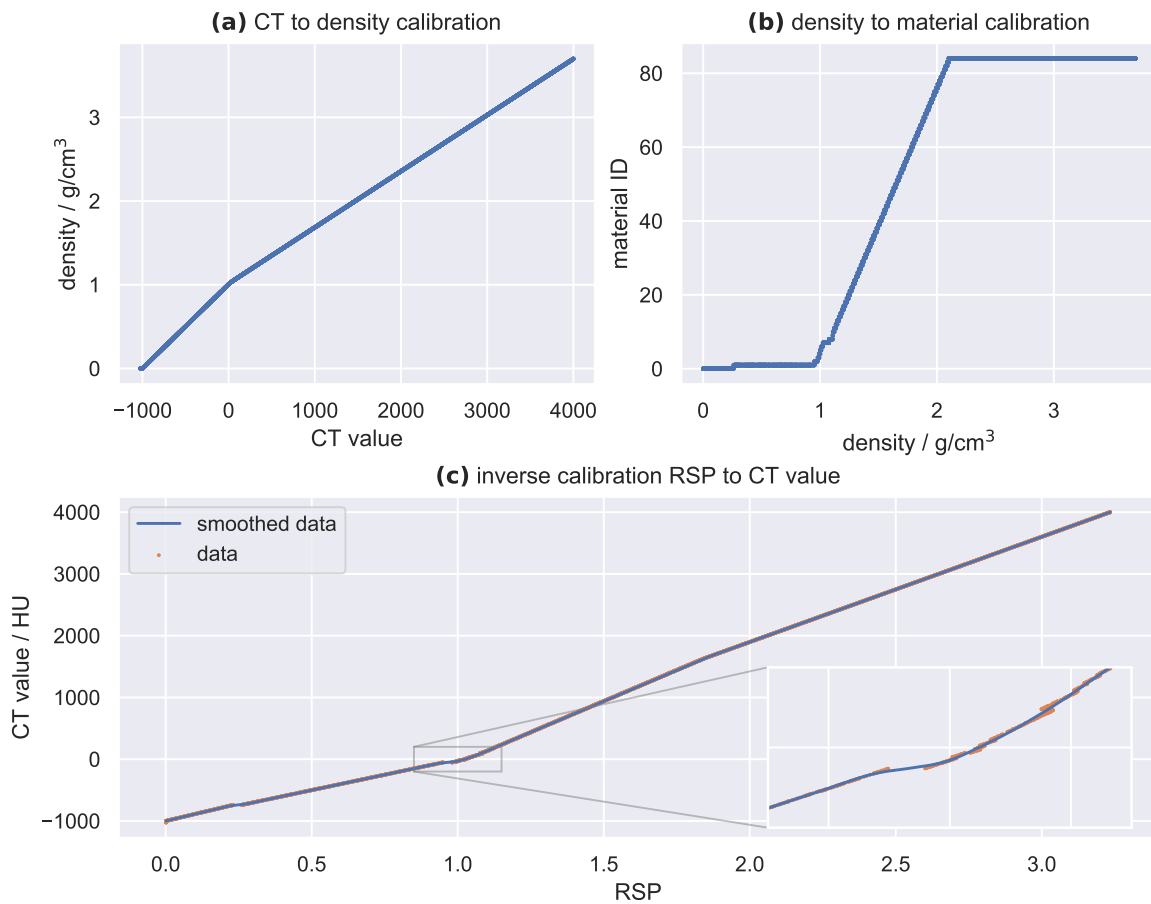
Figure 4.3: Calibration curves used to import x–ray CT data into the Monte Carlo simulation: (a) the CT to density calibration, (b) the subsequent density to material calibration, and (c) the inverse calibration of RSP to CT value including a smoothing to account for discontinuities. *Data source:* Resch et al. [145].

Dickmann et al. [138] by calculating the RSP values of a phantom containing random CT values from $-1000\,$HU to $4000\,$HU using the procedure described above. From this data, an inverse mapping from RSP to CT value was created as shown in fig. 4.3 (c). Due to the changes in elemental composition, this mapping had several discontinuities where a fitting of splines was applied enforcing the RSP to CT value relationship to be bijective (see enlarged sub–plot of fig. 4.3 (c)).

For the use of patient data in this work, please note the ethics statement on page 175.

5 | IMAGE NOISE

T<small>HE UNDERSTANDING AND PREDICTION</small> of noise in images of a specific prototype pCT scanner is a key requirement to enable the development of fluence control algorithms for FMpCT. In this chapter, the different contributions to image noise are disentangled using a realistic Monte Carlo simulation of the imaging setup. This required to improve the modeling of the energy detector to consider the effect of light quenching, as well as a careful reproduction of the proton beam profile and energy distribution. The precise modeling of all noise contributions within the imaging chain allowed to match experimental noise levels of the prototype scanner. By disentangling noise contributions, the impact of heterogeneities in the phantom on the image noise was studied and was found to be a critical contribution, which was not considered in previous works. The agreement of noise predictions and experimental scans of the same phantoms was better than 7 % for the three investigated phantoms: the water, CTP and the head phantom. This was found to be a sufficiently good agreement to allow for future experimental acquisitions of FMpCT scans optimized based on noise predictions of the improved Monte Carlo simulation.

The results presented in this chapter were published in *Physics in Medicine & Biology* [100]. Parts of the results were presented at the PTCOG Meeting 2019 in Manchester, United Kingdom [147] and at the Jagiellonian Symposium in Krakow, Poland [148].

PUBLICATIONS IN THIS CHAPTER

[100] » J. Dickmann et al.: "Prediction of image noise contributions in proton computed tomography and comparison to measurements". *Physics in Medicine & Biology* **64**(14) (2019), 145016. ISSN: 1361-6560. DOI: 10.1088/1361-6560/ab2474

CONFERENCE CONTRIBUTIONS IN THIS CHAPTER

[147] » J. Dickmann et al.: "Modelling of contributions to image variance in proton CT for application in low-dose fluence-modulated imaging". *Conference of the Particle Therapy Cooperative Group (PTCOG), Manchester, United Kingdom* (2019)

[148] » J. Dickmann et al.: "Monte Carlo study of image noise contributions of a prototype proton computed tomography scanner". *Jagiellonian Symposium on Fundamental and Applied Subatomic Physics, Krakow, Poland* (2019)

5.1 VERIFICATION OF VARIANCE RECONSTRUCTION

MATERIALS & METHODS

To verify the use of variance reconstruction described in section 3.5.4 and in particular to estimate if it is accurate enough to guide FMpCT scans, a first investigation was to compare results of the variance reconstruction according to eq. (3.62) based on a single dataset to the ground truth image noise based on independent simulations calculated according to eq. (3.43) as detailed in the next paragraph. In particular, the variance reconstruction algorithm in eq. (3.62) was used, which does not account for the pixel–by–pixel noise differences due to interpolation. This was motivated by the need of a computationally efficient algorithm that could be used in an FMpCT optimization, and considering interpolation only as a pre–factor allowed to avoid two of the three convolutions in the reconstruction formula. In Rädler et al. [124], the method was only applied to idealized data and using parallel proton beams. They also compared results only to image variance evaluated in annular rings of a single object, suggesting the need to revisit this in a verification study using ground truth noise from repeated simulations.

To generate data for the investigation, the Monte Carlo simulation described in section 4.1 was used together with the water phantom. N = 40 statistically independent tomographic acquisitions were simulated and individually reconstructed using the FDK–DDB algorithm. Subsequently, the voxel–by–voxel image variance was calculated to generate a ground truth noise map. This noise map was compared to the result of image variance reconstruction based on a single dataset. Before comparison, both noise maps were converted to standard deviations by taking the square root of the variance values.

RESULTS

Figure 5.1 (a) shows the RSP standard deviation of image voxels that was calculated as the square root of the ground truth image variance over the 40 independent reconstructions. Noise is reduced in the center and increases towards the edges. In particular, the edge of the
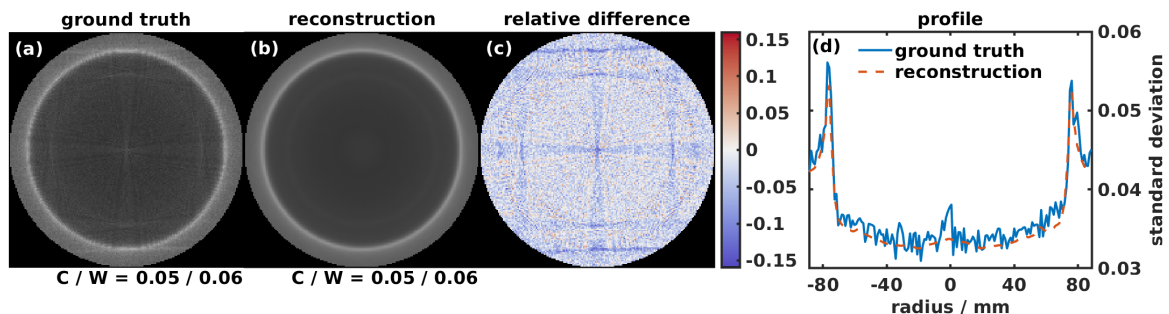
Figure 5.1: Comparison of (a) ground truth image noise maps to (b) the results of the variance reconstruction algorithm including (c) the relative difference between (a) and (b) as well as (d) diagonal profile plots at the center of the phantom. All data are shown at the central slice of the volume.

phantom is visible as a peak in image standard deviation. In addition, a fine star–shaped pattern is visible that increases image noise in the center as well as along lines aligned with the pixel grid and focused on the center. In comparison, fig. 5.1 (b) shows the square root of the image variance reconstruction, which is in good agreement with the ground truth image noise, but lacks the star–shaped pattern. In fig. 5.1 (c) the relative difference between (a) and (b) is shown, which is dominated by a star–shaped under–representation of image noise but, apart from that, shows only small errors. The mean error over the whole phantom was $-2.5\,\%$ and the root–mean–square error $4.1\,\%$. Profile plots in fig. 5.1 (d) confirm the slight under–representation, but a generally good agreement between the ground truth image noise and the results of the variance reconstruction algorithm in terms of the standard deviation is achieved.

5.2 NON–LINEARITIES OF THE ENERGY DETECTOR

MATERIALS & METHODS

The light production in scintillating detectors becomes non–linear for particles with a high stopping power, such as protons. This effect of quenching was experimentally investigated and described by Birks [149]. Light production in general, and such non–linearities in particular were not modeled in the simulation platform and instead the light yield was assumed to be proportional to the energy deposited in each stage. To allow for a direct comparison between experiments and simulations, in particular in terms of noise, quenching needed to be accounted for in the simulation. This is because quenching affects the calibration curves described in section 3.2.3, which in turn impact how noise in the energy measurement translates to noise in the WEPL. Quenching can be simulated using the theory of Birks [149], which describes the distorted energy measurement E'_n in stage n for a proton with a residual range R_n in the detector material as

$$E'_n(R_n) = S_n \cdot \int_{R_n}^{0} dx \, \frac{dE/dx}{1 + k_b \cdot dE/dx}, \tag{5.1}$$

where k_b is the material–dependent empirical Birks' factor, dE/dx is the stopping power of protons in the energy detector and S_n is an additional scaling factor of each stage introduced here to account for the calibration process of the pCT scanner and will be calculated in the following. During the calibration, data of $200\,\mathrm{MeV}$ protons in a degrader–free run is forced to produce energy readings equal to the E_n^{G4} defined in eq. (3.6), which were calculated without considering quenching. This can be expressed as

$$E_n^{G4} = S_n \cdot \int_{R_n^i}^{R_n^f} dx \, \frac{dE/dx}{1 + k_b \cdot dE/dx}, \tag{5.2}$$

where R_n^i and R_n^f are the initial range at the entrance and final range at the exit of stage n. If l_{stage} is the thickness of each stage and R_0 is the range of protons at the entrance of the energy detector without any absorber, then they can be calculated as

$$R_n^i = R_0 - (n-1) \cdot l_{stage}, \qquad (5.3)$$

$$R_n^f = R_0 - n \cdot l_{stage}. \qquad (5.4)$$

Consequently the scaling factors are dependent on k_b and calculate as

$$S_n = E_n^{G4} \cdot \left(\int_{R_n^i}^{R_n^f} dx \, \frac{dE/dx}{1 + k_b \cdot dE/dx} \right)^{-1}. \qquad (5.5)$$

The range R_n for a proton that additionally passed through a WEPL W inside the object can be calculated as

$$R_n = R_n^i - W/RSP_{det}, \qquad (5.6)$$

where $RSP_{det} = 1.038$ is the RSP of the detector. With this, the distorted energy deposit can be expressed as a function of the WEPL as

$$E_n'(W) = S_n \cdot \int_{R_n^i - W/RSP_{det}}^{0} dx \, \frac{dE/dx}{1 + k_b \cdot dE/dx}, \qquad (5.7)$$

where the stopping power in the detector dE/dx was calculated using GEANT4 for the theoretical material composition, for which the I–value was tweaked to produce the expected RSP.

To calculate the material–specific Birks' factor k_b, we inspected experimental data of the wedge–shaped calibration phantom, which was sampling the dynamic range of the detector from zero up to $254\,\text{mm}$ and where each proton's WEPL could be calculated from the hit locations and the known geometry of the phantom. From the data, a two–dimensional histogram was created of the measured energy deposit E_n^* to stopping stage n and the WEPL. The same histograms are shown in fig. 3.3 (a) and are used for the calibration of the scanner. The relationship $E_n^*(W)$ was determined for each stage by finding the most frequent energy deposit for each WEPL W.

The Birks' factor k_b as well as the quantities R_0 and l_{stage} were determined by optimization as

$$k_b, R_0, l_{stage} = \underset{k_b, R_0, l_{stage}}{\arg\min} \sum_n \left[\sum_W \left(E_n'(W) - E_n^*(W) \right) \right]^2, \qquad (5.8)$$

where the sums are over all stages and all WEPL bins of the histogram. The optimization was performed using the quasi–Newton method of Broyden [150]. While R_0 and l_{stage} could also be estimated, they were intentionally left as optimization variables to avoid that small errors in their estimation disturbed the correct estimation of k_b.

In subsequent simulations, the optimal Birks' factor k_b was then used to distort each incremental energy deposit of the simulation as

$$\frac{dADC}{dx} \propto \frac{dE'}{dx} \propto \frac{dE/dx}{1 + k_b \cdot dE/dx}, \qquad (5.9)$$

where dADC is the incremental increase of the ADC number. The scaling factors S_n were only important during the optimization and can be disregarded in the simulation due to the arbitrary scaling of ADC numbers. The stopping power dE/dx at a given proton energy can be queried directly in each calculation step of the GEANT4 simulation.

RESULTS

In fig. 5.2 (a) the histogram of the energy deposit to the stopping stage and the WEPL is shown. This histogram is equivalent to the one shown in fig. 3.3 (a), except for a slightly different arrangement of the bricks with respect to the wedge compared to what is shown in fig. 3.2. During the acquisition for fig. 5.2 (a), the bricks were aligned with one edge of the wedge, allowing protons to exit at the side of the brick. Such protons then had a smaller WEPL than if they had fully penetrated all bricks. They were therefore missing for WEPLs at multiples of the brick thickness of 52 mm and occurred elsewhere in the histogram. This is visible in fig. 5.2 as a sudden reduction of counts at around 65 MeV for each stage. In fig. 3.3 (a), the bricks were shifted by 20 mm which reduced this effect. Nevertheless, a slight change in statistics as observed here does not impact the subsequent evaluation, since each WEPL bin was evaluated independently to find the maximum energy deposit.

In fig. 5.2 (b) the same histogram is shown for a simulation that did not account for the effect of quenching. Energy deposits to the stopping stage are considerably higher for all stages except for the fifth stage at the lowest WEPLs. The transitions in between stopping stages occur at the same WEPL values as in the experiment.

Table 5.1 reports the fitting results for the three open parameters k_b, R_0 and l_{stage}. Using the Birks' coefficient allowed to repeat the simulation of the histograms considering the effect of quenching, which resulted in the data shown in fig. 5.2 (c). The energy deposits for a given WEPL here agreed much better to the experiments in (a) compared to the data
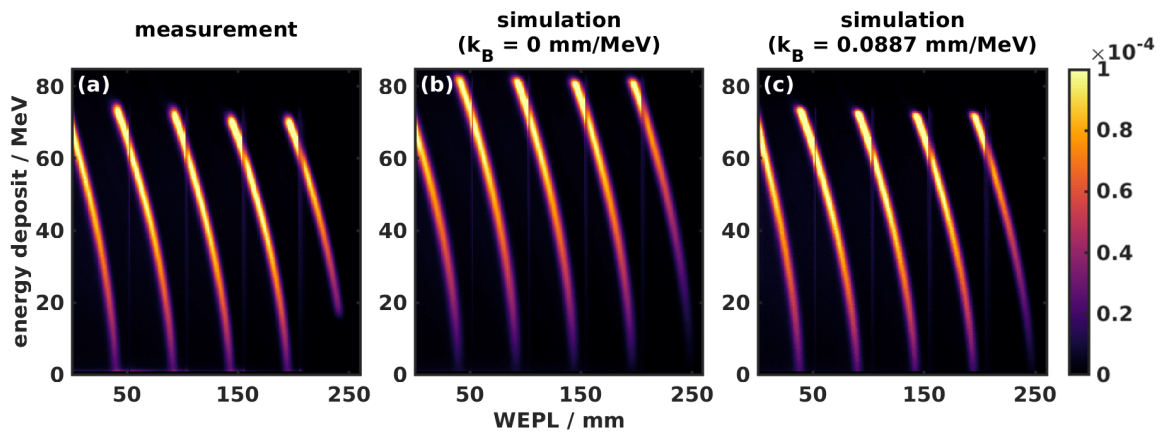
Figure 5.2: Histograms of the energy deposit to the stopping stage and the corresponding WEPL for (a) measurements, (b) simulations without the effect of quenching and (c) simulation with the optimized quenching parameter. One line in each histogram corresponds to one stopping stage, where the furthest stage is at the lowest WEPLs values.

Table 5.1: Fitting parameters for the optimization of the quenching model.

parameter		fit estimate
Birks' coefficient	k_b	$0.0887 \, \text{mm} \, \text{MeV}^{-1}$
residual range at entrance	R_0	$237.9 \, \text{mm}$
stage thickness	l_{stage}	$49.7 \, \text{mm}$
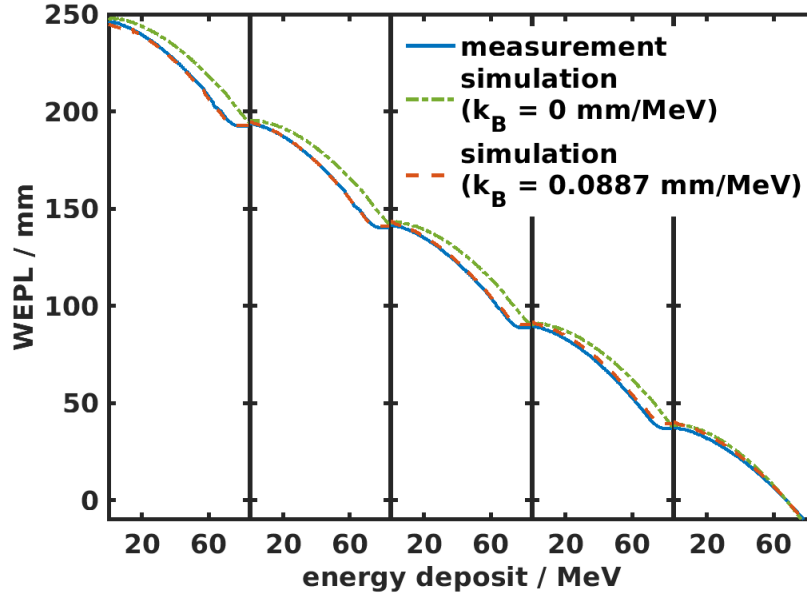
Figure 5.3: Calibration curves obtained from experiments and using simulations with and without simulating the effect of quenching. ⓒ Reproduced from Dickmann et al. [100] under the CC BY 3.0 license.

that did not consider quenching. The only clearly visible remaining difference between the histograms are WEPLs above $245 \, \text{mm}$, where the energy deposits are below $20 \, \text{MeV}$ in the first stage. In the experimental data this threshold is used to trigger the acquisition, as described in section 3.2.3. This is not needed and also not reproduced in the simulation.

In fig. 5.3, the resulting calibration curves are shown that were calculated from the data in fig. 5.2 and that are similar to those presented in section 3.2.3 and fig. 3.3. The calibration curve of the experimental data agreed well with the simulation using the optimized Birks' coefficient. The calibration curve for the simulation without the effect of quenching differed. Curves agreed at the beginning of a stage, where the energy deposit to the stopping stage was $0 \, \text{MeV}$ and quenching had little effect, and an increasing difference built up towards the end of a stage.

5.3 A MORE REALISTIC BEAM MODEL

MATERIALS & METHODS

Noise in the projection domain is inversely proportional to the number of particles incident to the detector as described by eq. (3.45). For this reason, it is important to carefully model the incident proton beam's fluence when comparing simulations to experiments both in terms of distributions of the initial positions and directions of the beam as well as for the initial energy spread of protons.

Experimental data used in this chapter employed a broad beam as described in section 3.7. The spatial and directional distribution of the incident proton beam could directly be replicated in simulations by using the tracking information of an experimental scan. For this purpose, the position of each detected proton was projected to a point at $400\,mm$ in front of the isocenter and $232.8\,mm$ upstream of the front tracker. This was done by following its direction information along a straight line and resulted in a list of positions and directions, that could be used in a subsequent simulation to place primary protons. To prevent interplay effects between the strips of the detector and the placement in the simulation, the location information was blurred with random numbers with a standard deviation equal to the distance between two strips of the detector. Since in the data available for this study the beam profile was purposefully changed after the calibration, and no degrader–free runs were available for the beam profiles used for the phantom scans, the beam model was based on a scan of the water phantom. Since this suffered from a reduced fluence in the area of the phantom due to attenuation, protons were randomly selected to be included in the beam model with a probability anti–proportional to the transmittance, thereby undoing the effect of attenuation. The transmittance was estimated from Monte Carlo simulations using a homogeneous beam by comparing the detected protons to the initial number of protons.

The energy spread σ_{beam} of the incident proton beam was determined indirectly by calculating the spread of energy deposits to the fifth stage of the detector in an absorber–free run. The spread was calculated by fitting a Gaussian function to histograms of the energy deposits. This was done to prevent outliers distorting the calculation. The standard deviation of energy deposits to the fifth stage, subsequently called σ_{E_5}, was compared to the spread of E_5 observed in a set of simulations with a varying beam energy spread σ_{beam}. To do so, σ_{beam} was set in Monte Carlo simulations to values between $0.5\,MeV$ and $1\,MeV$ in steps of $0.1\,MeV$ and the resulting σ_{E_5} as a function of σ_{beam} was fitted with a second order polynomial. The intersection of the polynomial with the experimental value of σ_{beam} allowed to determine a value of σ_{beam} that produced the same noise in the last stage of the detector as in an experiment. This of course assumes that the electronic noise is low compared to the noise caused by energy straggling and the beam energy spread, which is a valid assumption [56]. Another assumption is that the energy deposits to the stopping stage are equal between measurements and simulations for all WEPL values, which was achieved using the Birks' model described in section 5.2.
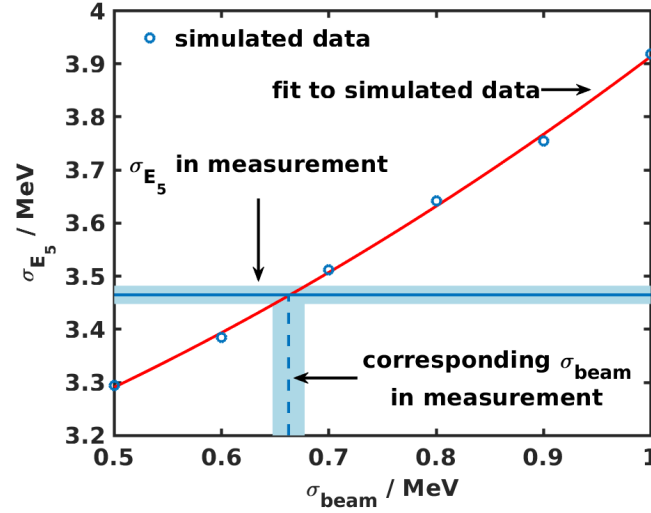
Figure 5.4: The spread of the energy deposit to the last stage σ_{E_5} as a function of the initial beam energy spread σ_{beam}. A fit on the simulated data was used to find a σ_{beam} that matches the experimental value.

RESULTS

Figure 5.4 displays the spread of energy measurements σ_{E_5} in the last stage as a function of the beam energy spread σ_{beam} for a set of simulations. The data were fitted by a quadratic model. To determine the beam energy spread that agreed with measurements, the spread of energy deposits in the last stage was determined for experimental data as $\sigma_{E_5} = (3.47 \pm 0.02)\,MeV$. According to the polynomial fit, this corresponded to a beam energy spread of $\sigma_{beam} = (0.66 \pm 0.02)\,MeV$ which was used in subsequent simulations. The uncertainty of σ_{E_5} was the uncertainty of the Gaussian histogram fit and the corresponding uncertainty for σ_{beam} was propagated through the quadratic function. Both uncertainties are displayed as shaded ares in fig. 5.4.

❯

5.4 CONTRIBUTIONS TO PROJECTION NOISE

MATERIALS & METHODS

The modeling of quenching and the beam energy spread in sections 5.2 and 5.3 allowed to directly compare simulated image noise with the noise in experimental data. This was first done based on data of the wedge–shaped calibration phantom allowing to verify the simulation model for the complete dynamic range of the detector. To allow for a comparison with the results of Bashkirov et al. [56], the per–proton WEPL standard deviation was calculated as

$$\sigma_{WEPL} = \sigma_P \cdot \sqrt{N}, \tag{5.10}$$

where σ_P is the standard deviation expected for the mean WEPL of N protons as defined in eq. (3.45) (standard error of the mean). The projection noise data was binned to WEPL intervals in multiples of 1.5 mm and in each of these bins the median value of σ_{WEPL} was calculated.

Apart from comparing results to experiments, contributions to image noise in the projection domain could be disentangled with the help of Monte Carlo simulations. The following contributions were hypothesized: scattering due to MCS, energy straggling in the object, the uncertainty of tracking, the energy detection process, and the beam energy spread. To disentangled these contributions, various scoring techniques were employed in the Monte Carlo simulation, which are described in the following. For this purpose, two ideal scoring planes were implemented that have no effect on the crossing particles, with one located before and one after the object and both just inside the two tracking detectors. The noise contributions were as follows.

WEPL SCORING   In each calculation step of the Monte Carlo simulation, the step length was multiplied with the current material's RSP and summed for each proton to calculate the exact WEPL. The exact coordinates and directions were recorded at the two scoring planes. Noise in this case was only affected by scattering of protons with different histories to the same distance–driven bin. Distance–driven binning forces protons in a given bin to have the same location at a certain depth of interaction, but their paths from and to that bin will differ due to MCS and they will experience different WEPLs, which results in a (noisy) distribution of WEPLs in the distance–driven bin.

ENERGY SCORING   At the two scoring planes, the proton's exact energy, position and direction were recorded. The energy was converted to WEPL with eq. (3.3). In addition to noise from scattering, this scoring technique also considered energy straggling in the object as a contribution.

ENERGY SCORING (REALISTIC POSITION)   With scoring of the energy as before, the location and direction were replaced by the values inferred from the simulation of the strips in the tracking detector. This technique, therefore, also included noise caused by the tracking process. While this contribution is expected to be small, a less accurate path estimate may contribute to noise in a similar way that MCS does.

REALISTIC SCORING (NO BEAM ENERGY SPREAD)   The full detector simulation was employed, including the modeling of the energy detector and the corresponding calibration process. The beam energy spread was set to $\sigma_{beam} = 0\,MeV$. This scoring technique considered all sources of noise, except for the beam energy spread and in particular also noise from energy straggling in the detector and the calibration process.

REALISTIC SCORING   The full detector simulation was used with the beam energy spread determined in section 5.3. This accounts for all hypothesized sources of noise.

––––––––––

The noise due to scattering was determined directly by the WEPL scoring technique. Noise due to energy straggling in the object was determined as a difference between noise in

the energy scoring and the WEPL scoring technique. Noise due to tracking was calculated as the difference between the two energy scoring techniques. Noise due to the energy detection process was equal to the difference of the realistic scoring technique without beam energy spread and the energy scoring technique with realistic position measurements. Finally, noise due to the beam energy spread was determined as a difference of the noise in the two realistic scoring techniques. To calculate any of these differences, the variance values needed to be considered instead of standard deviations. This is, because two sources of (uncorrelated) variances add up linearly, while two standard deviations need to be added as the square root of the sum of the squares.

RESULTS

Figure 5.5 (a) – (c) shows the standard deviation of the various scoring techniques based on the data of the wedge–shaped calibration phantom. Data are shown at the front tracker, at the isocenter and at the rear tracker. The WEPL scoring noise is negligibly small at the front tracker, but considerably increases at the rear tracker. At multiples of the brick
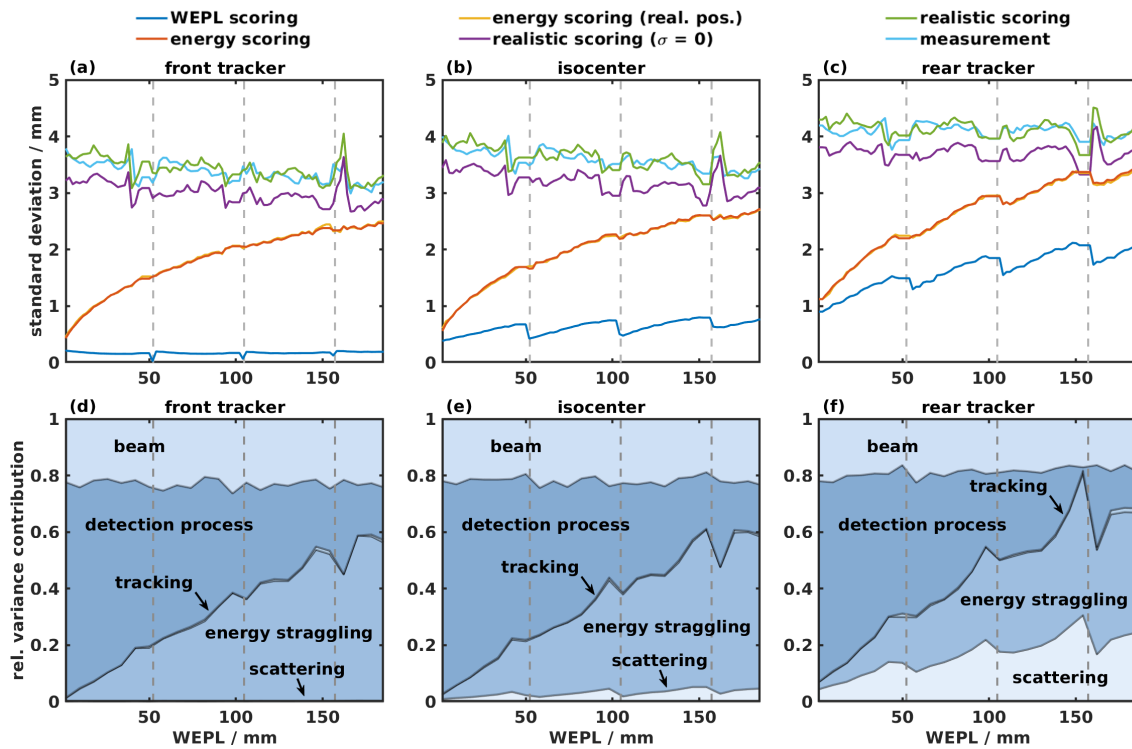
Figure 5.5: Contributions to image variance and comparison of simulated and experimental data in the projection. In (a) – (c) the standard deviation of the individual scoring techniques is shown and in (d) – (f) the relative contributions to projection variance of the five noise contributions. Data are evaluated for (a) and (d) at the front tracker, for (b) and (e) at the isocenter, and for (c) and (f) at the rear tracker. Dashed lines indicate multiples of the calibration phantom's brick thickness. The term detection process only refers to detection of the proton energy and not to tracking. ⓒ Reproduced from Dickmann et al. [100] under the CC BY 3.0 license.

thickness (indicated by dashed lines) there is a sudden decrease of projection noise due to scattering. The standard deviation in the energy scoring data is small for small WEPLs and increases with for larger WEPLs. The use of energy scoring with realistic tracking only had a minor impact on projection noise compared to energy scoring with ideal tracking. For realistic scoring, the dependence on the WEPL became less severe. The standard deviation per proton is between 3.1 mm for large WEPLs and 3.8 mm for the smallest WEPLs. The agreement with the experimental data is good, even a minor increase of noise in the center of each dataset is reproduced.

In fig. 5.5 (d) – (f) the corresponding contributions to image variance are shown and normalized to the variance using realistic scoring. Scattering in this dataset can be disregarded, except for at the rear tracker, where the contribution is about 20 % for the largest WEPLs. Variance due to tracking is not relevant for the dataset under investigation, independent of the tracking plane. The sum of energy straggling in the object and the noise caused by the energy detection process is approximately constant with energy detection noise prevailing for low WEPLs and energy straggling prevailing for large WEPLs. The contribution of the beam energy spread is about 20 % for all tracking depths and across the complete WEPL range.

5.5 COMPARISON OF SIMULATED & EXPERIMENTAL IMAGE NOISE MAPS

MATERIALS & METHODS

To assess the accuracy of the Monte Carlo noise model, pCT scans were simulated for the water phantom, the CTP phantom and the pediatric head phantom (see section 3.6) and image noise maps were calculated using the variance reconstruction algorithm. The simulation was performed using the realistic scoring technique. For the same phantoms, experimental data was obtained from scans performed at the Northwestern Medicine Chicago proton center with the prototype pCT scanner and image noise maps were calculated using variance reconstruction. The voxel size for this study was uniformly 1 mm in all dimensions. All noise maps were normalized to an average projection fluence of $f_0 = 20 \, \mathrm{mm}^{-2}$ to match experimental and simulated scans, which had slightly different particle numbers. For the phantoms that were symmetric in z–direction (water and CTP), 16 slices were averaged. The averaging of the noise maps was done after noise reconstruction, so that the map still refers to the noise in a single slice before averaging. For the head phantom, three representative slices were selected for evaluation.

In order to align the simulated and experimental noise maps, two corresponding slices in the reconstructed RSP maps were selected and both slices were rigidly registered to each other allowing a translation and a rotation. The determined registration was then applied to the image noise maps allowing for a voxel–by–voxel comparison. For this, the evaluation

only considered voxels inside the object's hull, which was determined by an RSP threshold of 0.15.

A voxel–wise map of the relative error was calculated as

$$\Delta\sigma = \frac{\sigma_{\text{RSP, sim}} - \sigma_{\text{RSP, exp}}}{\sigma_{\text{RSP, exp}}},\tag{5.11}$$

where σ_{RSP} is the square root of the variance reconstruction of simulated or experimental data.

In the analysis of the data of heterogeneous phantoms, it turned out that the contribution of scatter to image noise played a more important role compared to what could be expected from the previous results using the homogeneous calibration phantom. To quantify the noise contribution of scatter σ_{scatter}, the simulations were repeated using the WEPL scoring technique. The non–scatter contribution, which is the expected noise level in the absence of scattering, was determined as the difference

$$\sigma_{\text{non–scatter}} = \sqrt{\sigma_{\text{RSP}}^2 - \sigma_{\text{scatter}}^2}.\tag{5.12}$$

Mean imaging doses in each slice were scored using the GEANT4 simulation to put the noise levels into perspective with the required imaging dose.

RESULTS

In fig. 5.6 (a) RSP maps are shown for the three phantoms used in this investigation. For the head phantom, three representative slices were selected: slice 1 with the ears and the nasal cavity, slice 2 with the teeth and the spine and slice 3 with the eyes and the brain. In particular in the water and the CTP phantom, ring artifacts distort the image. Those ring artifacts are also seen in the head phantom, but mostly masked by the geometry of the phantom. Ring artifacts are known to distort images of this pCT scanner [81] and are subject to the works reported in this thesis in chapter 9.

Figure 5.6 (b) and (c) show the experimental and the corresponding simulated image noise maps in terms of the standard deviation with (d) showing the relative difference between both. There is a good agreement between the two noise maps with the absolute level of noise agreeing and even subtle features of the noise maps being reproduced by the simulation. For the water phantom, the image noise is low in the center and increases considerably towards the edge of the phantom, as it was observed already in section 5.1. Increased noise also occurs close to the air inserts of the CTP phantom as well as close to heterogeneities of the head phantom, in particular in the area of the nasal cavity. The relative difference seems to be slightly increased close to ring artifacts, which is visible in particular for the water phantom. Table 5.2 summarizes the relative mean error as well as the relative root mean square (RMS) error for all phantoms and slices. While the CTP and the water phantom are both underrepresented by the simulation, the head phantom both contains slices with negative and positive mean error. The RMS error is below 7 % for all phantoms and slices.

Figure 5.6 (e) displays horizontal central line profiles through the variance maps of experimental and simulated standard deviation maps. They confirm the good agreement
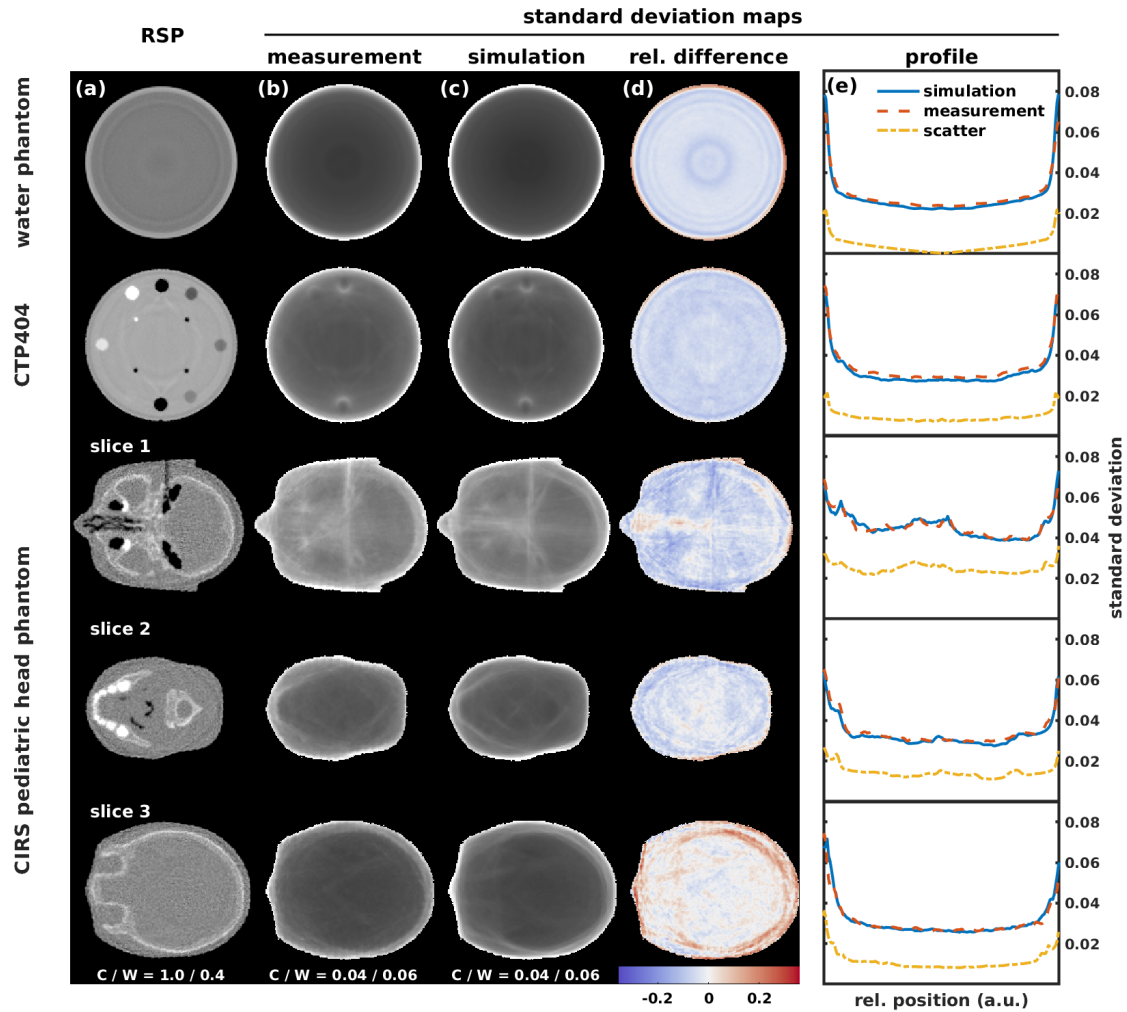
Figure 5.6: (a) RSP and image noise for three different phantoms comparing (b) simulated to (c) experimental standard deviation maps; (d) relative difference between simulation and experiment and (e) horizontal profiles along a central line of the image noise maps as well as the contribution from scattering. ⓒ Reproduced from Dickmann et al. [100] under the CC BY 3.0 license.

Table 5.2: Mean error and root mean square (RMS) error for the comparison of simulated and experimental image noise maps. The errors of the simulated versus the experimental results are reported with relative to the experimental data.

phantom	relative mean error	relative RMS error
water phantom	−2.9 %	6.7 %
CTP phantom	−5.6 %	6.4 %
head phantom (slice 1)	−4.6 %	6.8 %
head phantom (slice 2)	−3.6 %	5.3 %
head phantom (slice 3)	3.2 %	6.2 %

Table 5.3: Standard deviation levels of a central circular region (diameter 40 mm) of experimental and simulated noise reconstructions and the corresponding mean imaging doses.

phantom	noise level *simulation*	noise level *experiment*	imaging dose / mGy *simulation*
water phantom	0.022 ± 0.001	0.024 ± 0.001	0.85 ± 0.04
CTP phantom	0.027 ± 0.001	0.029 ± 0.001	0.86 ± 0.04
head phantom (slice 1)	0.044 ± 0.003	0.046 ± 0.002	0.82 ± 0.04
head phantom (slice 2)	0.030 ± 0.001	0.030 ± 0.001	0.84 ± 0.04
head phantom (slice 3)	0.026 ± 0.001	0.027 ± 0.001	0.88 ± 0.04

that was already visible from the relative difference maps, but also highlight that differences between experimental and simulated noise data are minor compared to the overall fluctuation of image noise within a slice. The lowest image noise was observed in the center of the water phantom as well as in the homogeneous region of slice 3 of the head phantom (both around 0.025). In the CTP phantom and slice 2 of the head phantom, the lowest noise was at around 0.03 while for the center of slice 1 the standard deviation increased to 0.05, which is a two–fold increase of standard deviation compared to the homogeneous slice 1 of the same phantom. A detailed analysis of noise levels in a circular central region with a diameter of 40 mm is shown in table 5.3.

The scatter contribution to image noise is shown in fig. 5.6 (e) as a third profile. For the water phantom and the homogeneous slice 3 of the head phantom the scatter contribution is close to or at zero while for slice 1 of the head phantom it is above 0.025. For all phantoms the scatter–only noise increases towards the edge, most prominently observed for the water phantom. The non–scatter noise contribution according to eq. (5.12) is approximately constant for all phantoms. In the central region of all phantoms it was $\sigma_{\mathrm{non-scatter}} = 0.028 \pm 0.004$ with the uncertainty calculated over all phantoms and slices.

Imaging doses for the simulations are also reported in table 5.3 allowing for a comparison of noise values reported here to other studies. Since pile–up in the detector was found to be a negligible effect [46], imaging doses for the experimental scans are expected to be similar.

5.6 A BOW–TIE FILTER FOR PROTON CT

MATERIALS & METHODS

For diagnostic imaging it is desirable to have a constant imaging noise level throughout the image. For this purpose, x–ray CT scanners employ a so–called *bow–tie filter*, a physical device which attenuates the primary beam and thereby flattens the fluence at the detector, which homogenizes image noise as described by Harpen [151] or Graham et al. [152]. This approach, however, is not directly applicable to pCT, since different phenomena govern the noise formation. For x–ray CT it is the attenuation of the object, which reduces the fluence

incident to the detector. For pCT, attenuation plays a minor role and instead the noise level is governed by the interplay of the heterogeneity of the object and MCS as shown in section 5.5 and by Rädler et al. [124]. As a preliminary fluence modulation study, a fluence profile was designed that would be the equivalent of a bow–tie filter for pCT, and which makes noise at the detector level flat, and thereby also results in a homogeneous image noise. This pCT–bow–tie would not be a physical device, but just a fluence profile.

The bow–tie–like fluence modulation profile was created based on simulated data of the water phantom and subsequently applied to experimental data of the same phantom. To apply the fluence modulation, protons of the experimental dataset were randomly accepted or rejected with an acceptance probability $p(u)$ which was defined as a function of the lateral detector coordinate u (along the long side of the water phantom). No modulation was required along v since the phantom is symmetric in this direction. The acceptance probability was calculated as

$$p(u) = \min\left(\frac{\sigma_{p,\text{sim}}^2(u)}{\sigma_0^2}, 1\right), \tag{5.13}$$

where $\sigma_{p,\text{sim}}^2(u)$ is the simulated variance profile along u at the isocenter and σ_0^2 is the desired constant variance level. The acceptance probability was forced to be smaller than 1 by the min operator in the case that $\sigma_{p,\text{sim}}(u) > \sigma_0$. This allowed to prescribe a desired variance that was less than the maximum variance in the reference scan (at the edges of the phantom) at the cost that the prescription would not be achieved in this region. In this study, a projection standard deviation of $\sigma_0 = 5.48\,\text{mm}$ was prescribed, which corresponds to a variance of $\sigma_0^2 = 30\,\text{mm}^2$.

RESULTS

Figure 5.7 (a) shows the image noise of the water phantom which is equal to the one in fig. 5.6 (b) and was obtained with a uniform fluence profile, which is shown in fig. 5.7 (d). Figure 5.7 (b) shows the image noise map for the same phantom acquired with the bow–tie fluence modulation profile which is also shown in (d). The image noise is flat as
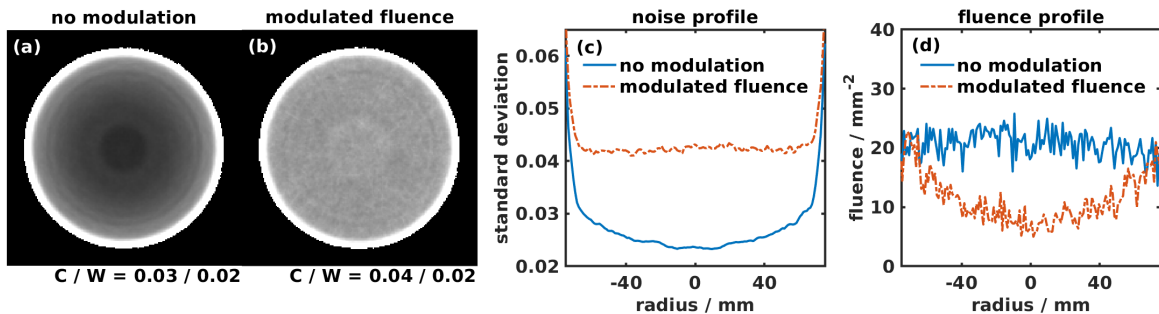
Figure 5.7: A first fluence modulation based on the image variance predictions: (a) image variance for the water phantom without and (b) with a modulation forcing the image variance to be constant; (c) horizontal profiles through the image noise maps and (d) corresponding fluence profiles. ⓒ Reproduced from Dickmann et al. [100] under the CC BY 3.0 license.

intended with only an increase of image noise at the edge of the phantom. Figure 5.7 (c) displays central profiles through the image standard deviations in (a) and (b). The noise level in the modulated scan is considerably higher since protons could only be rejected in this retrospective modulation. The profile is flat within the body of the phantom and in particular compared to the profile of the original scan. The fluence profiles in fig. 5.7 (d) follow opposite trends compared to the noise profiles: while in the original scan noise is low in the center and fluence is constant, the modulated scan has constant noise in the image, but a reduced fluence in the center. An analysis of the RSP values of the two scans showed no degradation and changes in the mean RSP were less than the image noise level.

5.7 DISCUSSION

Verification of variance reconstruction

The use of variance reconstruction to calculate image noise from just a single dataset was verified by comparing the reconstruction results to the variance of $N = 40$ independent reconstructions of the same phantom. Differences of just a few percent mainly stem from a star–shape noise increase due to effects of interpolation, which was purposefully not considered and which is a systematic error that does not depend on the object shape or dimension. In particular, for the relative comparisons performed in the rest of this chapter, the disregarded star–shaped noise increase would not impact the results as it cancels out in a relative comparison. For the purpose of guiding an FMpCT fluence optimization, the noise reconstruction without the explicit modeling of interpolation effects appears to be well suitable given that the star–shaped increase is only of a small amplitude and only affects a thin region of the image. However, this is not relevant for fluence modulation due to the finite size of proton pencil beams, which limits the spatial frequency at which image noise can be modulated. Moreover, errors of just a few percent, compared to the ground truth noise seem to be acceptable for FMpCT, given that modulation of noise just due to the object shape were observed in this study that reached a factor of two in between phantoms and within one phantom towards the edge. The error of variance reconstruction, therefore, is small compared to the expected noise modulations.

Non–linearities of the energy detector

To model non–linearities of the energy detector, experimental data of the wedge–shaped phantom was used to fit a Birks' coefficient that allows to model the effect of light quenching in subsequent simulations. For this, the initial range of protons at the detector entrance as well as the thickness of each stage were left as open fit parameters. The initial range of $200\,MeV$ protons is $250\,mm$ in polystyrene [53]. From this value, the range loss in the tracking detector and in air, as well as in the entrance window of the energy detector would need to be reduced. For this, the fit value of $R_0 = 237.9\,mm$ appears to be reasonable. Also

the stage thickness, which physically is 51 mm agrees within the expected uncertainty with the fit value of $l_{stage} = 49.7$ mm. Potential errors in the estimate for the detector's RSP and also the impact of the stage wrapping could impact that value. The fact that both fit parameters were close to the expected values, but not exactly equal, confirmed, that leaving them as open parameters was a valid choice and that fixing them at the expected values may have deteriorated the estimate for k_b.

The Birks' factor was estimated as

$$k_b \cdot \rho_{PS} = 9.4 \times 10^{-3} \frac{g}{MeVcm^3}, \tag{5.14}$$

where $\rho_{PS} = 1.06$ gcm^{-3} is the density of polystyrene. This agrees well with values from literature, which were reported to be $k_b \cdot \rho_{PS} = 9 \times 10^{-3}$ gMeV^{-1}cm^{-3} by Tretyak [153] and $k_b \cdot \rho_{PS} = 14 \times 10^{-3}$ gMeV^{-1}cm^{-3} by Reichhart et al. [154].

With the correct modeling of quenching, calibration histograms for simulations and experiments, as well as the calibration curves themselves, agreed considerably better. While this agreement may also have been achievable by simulating the photon transport in the GEANT4 code, the approach presented here is preferable since it comes at little computational cost compared to a full simulation of the photon transport. While modeling quenching was not crucial in the original work of Giacometti et al. [103], where the RSP fidelity was the objective, in this thesis, quenching is more important since it impacts the slope of the calibration curve and may, therefore, have an important impact for the prediction of image noise levels using the Monte Carlo simulation. Additionally, it allowed to estimate the beam energy spread in a subsequent investigation, which required to have equal energy deposits for the same WEPL between experiments and the simulation.

A more realistic beam model

A precise model of the incident beam in an experiment was found by exploiting the tracking information of the pCT scanner. By reading the information and correcting it for the effect of attenuation, the experimental beam could be reproduced in the Monte Carlo simulation. The beam energy spread was estimated to be about 0.3 % of the beam energy by comparing the spread of energy deposits in the detector's final stage with experiments. This value is lower than the value expected for clinical operation of the proton beam, which typically is between 0.5 % and 1 % of the beam energy [155]. However, the operation of the pCT scanner required to reduce the proton fluence considerably, so that the tracking rate was below 1 MHz and the scanner could operate without pileup. For this purpose, energy slits of the accelerator needed to be closed beyond values that are commissioned for treatment. More narrow energy slits, therefore, may explain the smaller energy spread observed in this work.

Contributions to projection noise

Using the optimized Monte Carlo code allowed to disentangle contributions to image noise in the projection of the wedge–shaped calibration phantom. The contribution of scattering

was negligible for the front tracker but considerably increased at the rear tracker because of the continuous opening of a proton beam due to MCS as it traverses the phantom. In addition, the calibration phantom was positioned close to the front tracker, leaving an air gap where the protons' drift along straight trajectories (with strongly reduced MCS) and which further increased the lateral beam spread. The scattering contribution at the rear tracker suddenly decreased at multiples of the brick thickness. This was, because at this point data with n bricks at the center of the wedge and with $n + 1$ bricks at the thin side of the wedge were merged. While in general a larger WEPL increases the noise due to MCS, an additional brick reduces the air gap and consequently decreases the drift distance. This abruptly also decreased the lateral beam spread until an even larger WEPL compensates for the decrease.

Energy straggling was the dominating contribution to projection noise and was originating both from straggling in the object and in the detector. For protons in air, straggling occurred only in the detector. For an increasing WEPL, the path length in the detector decreased and with it the noise caused by the energy detection process. At the same rate, noise due to energy straggling in the object increased. Noise due to the energy detection process may also be elevated due to the calibration curve, which is of varying slope and itself noisy as well as from electronic noise and noise of the scintillating detector. However, their contributions are small and the detector operates close to the energy straggling limit, as investigated by Bashkirov et al. [56]. Remaining contributions may be covered by the estimation of the beam energy spread. The beam energy spread contributed about 20 % to the total projection variance. This is a remarkable contribution considering that the spread was only 0.3 % of the initial beam energy. It, therefore, requires a careful modeling to predict correct noise values with simulations.

The per–proton projection noise using the full simulation agreed well with the values observed in experiments. The summed noise contributions were flat across the WEPL dynamic range of the detector with a slight decrease for larger WEPLs. The per–proton WEPL uncertainty agrees well with the values reported by Bashkirov et al. [56] for the same scanner and using eq. (3.5).

Comparison of simulated & experimental image noise maps

Application of variance reconstruction to simulated and experimental data of three phantoms allowed to quantify the accuracy of noise predictions in the image using Monte Carlo simulations. The agreement between simulations and experiments in terms of the standard deviation was better than 7 % for all phantoms. In addition to the potential error due to the use of variance reconstruction estimated in section 5.1 of a few percent, which canceled out in this comparison, the expected noise prediction error will still be less than 10 %. This is small compared to the modulation of image standard deviation that is expected for even the homogeneous water phantom and which were up to a factor of two. Performing an FMpCT optimization based on the predicted image noise should, therefore, be feasible.

An interesting observation in this study was that the noise level depended strongly on the heterogeneity of the phantom. This effect was driven by the contribution of scattering noise, which is close to zero in the center of the water phantom, but increased considerably towards

the edges of the phantom. This caused the total noise in the water phantom to be lowest in the center and increase towards the plastic hull. A considerable noise contribution of scatter was also found for heterogeneous phantoms. For this reason, noise even in between two non–adjacent slices of the same phantom, the head phantom, can change by up to a factor of two. According to eq. (3.45), a noise increase in terms of the standard deviation by a factor of two would need to be compensated by an incident fluence that is increased by a factor of four, which consequently also increases imaging dose by a factor of four. For this reason, the imaging dose advantage of pCT over x–ray CT, that was suggested by Schulte et al. [44], may need to be revisited, since this was performed at the center of the water phantom, which had the lowest noise level in this study and would not be representative of a clinical usage of pCT. At the same time, the non–scatter contribution was similar for all phantoms, which is consistent with section 5.4 and the observations of Bashkirov et al. [56], that the per–proton noise is constant with the WEPL.

A bow–tie filter for proton CT

Using the noise prediction of the Monte Carlo simulation, a bow–tie like fluence modulation was designed, which achieved flat noise in the image. At the same time, the accuracy of RSP values was maintained. This is a first application for fluence modulation and shows that it should be feasible to use noise predictions from Monte Carlo simulations to guide the optimization of FMpCT. The fluence modulation profiles required for flat image noise in pCT are fundamentally different from those employed by bow–tie filters in x–ray CT [151, 152]. While a bow–tie filter needs to attenuate less in the center to compensate for the attenuation of the object, the fluence incident to the object in pCT needs to be elevated in the center due to increased MCS and the noise related to that.

While in x–ray CT, bow–tie filters are a standard approach to reduce patient exposure, a similar approach has not been used for pCT so far. A reason for that might be, that it is not possible in pCT to design a modulation profile that is suitable for a wide range of patients or objects. As this study showed, even two slices of the same phantom experience considerably different noise levels and would, consequently, require different modulation profiles. FMpCT will, therefore, be always a patient–specific approach that carefully needs to model image noise in the specific patient to adapt the fluence accordingly. Nevertheless, calculating a virtual bow–tie filter for the water phantom demonstrated the feasibility of using Monte Carlo simulations to predict image noise and calculate fluence modulation profiles that allow for a flat image noise distribution.

5.8 CONCLUSIONS

In this chapter and originally in Dickmann et al. [100], an existing Monte Carlo code for the simulation of pCT scans of a prototype scanner was improved in order to reproduce image noise maps of experimental scans. For this purpose, a quenching model was implemented and the corresponding Birks' parameter was determined based on experimental data. Moreover, a realistic beam model was established based on the tracking data of the

scanner and the initial beam energy spread was estimated by an indirect comparison of the spread of energies of the detector's last stage measured in experiments and simulations. This allowed to match experimental noise levels over the complete dynamic WEPL range of the detector. Additionally, the Monte Carlo simulation could be used to quantify the relative contributions to WEPL noise of the energy detection process, energy straggling in the object, scattering, tracking and the initial beam energy spread. In particular the noise due to the initial beam spread was considerable with a relative variance contribution of about 20 %. For three different phantoms, the image noise was calculated with the help of a variance reconstruction algorithm, whose accuracy was verified within this study. Noise was shown to considerably increase close to heterogeneities and this increase is driven by the noise contribution caused by MCS.

In this first full simulation of all relevant contributions to image noise, an agreement to experimental data of better than 7 % in terms of the standard deviation was achieved. This error was shown to be much smaller compared to the noise fluctuations expected for pCT scans. Using the noise predictions for FMpCT, therefore, should be feasible. To further motivate that, a modulation profile for the water phantom was calculated, that successfully achieved constant noise in the image and thereby modeled the equivalent of a bow–tie filter for pCT. However, the bow–tie modulation would need to be adapted for each patient, underlining the need for patient–specific fluence modulation using FMpCT.

6 FLUENCE–MODULATED PROTON COMPUTED TOMOGRAPHY

B�y ᴇᴍᴘʟᴏʏɪɴɢ ꜱᴍᴀʟʟ ᴘʀᴏᴛᴏɴ ᴘᴇɴᴄɪʟ ʙᴇᴀᴍꜱ with modulated intensities or dwell times, fluence–modulated pCT aims to achieve a task–specific distribution of image noise. Such a spatially varying image quality is particularly meaningful in the context of particle therapy, where good image quality for treatment planning, and thus low noise, is only needed in the vicinity of the therapeutic proton beam (the ROI). Outside of this ROI, image noise can be increased and imaging dose decreased accordingly.

In this chapter, an optimization algorithm is presented, that calculates fluence modulation factors for each pencil beam, which achieve a desired image noise distribution. The algorithm is based on a uniform fluence pCT Monte Carlo simulation of the object, that predicts variance levels of a specific pCT scanner. It also makes use of a Gaussian pencil beam model, which was established based on experimental data within this work. The performance of the optimization algorithm is tested in a Monte Carlo simulation study based on three phantoms and three different noise prescriptions. Resulting imaging dose savings are compared to the uniform fluence case and to a simple, intersection–based reference method, that was used in earlier works. By prescribing a constant image noise, the imaging dose to a homogeneous water phantom could be reduced by 8.9 % compared to an acquisition using a uniform fluence and at the same peak noise level. For the heterogeneous head phantom, this dose reduction increased to 16.0 %. For two FMpCT imaging tasks with different ROI shapes, the dose saving outside of the ROI was between 25.7 % and 40.5 % at the same peak noise level inside the ROI. The imaging dose inside the ROI increased by between 9.2 % and 19.2 %. The agreement between the prescribed and the achieved noise levels was satisfactory and differences were small compared to the overall intended fluence modulation. The use of a realistic pencil beam model and simulation of the scanner paved the way towards an experimental realization reported in the next chapter.

The results presented in this chapter were published in *Medical Physics* [156]. Parts of the results were presented at the AAPM conference in San Antonio, USA [157], at the DGMP meeting in Stuttgart, Germany [158], and at the ESTRO conference in Vienna, Austria [159].

PUBLICATIONS IN THIS CHAPTER

[156] » J. Dickmann et al.: "An optimization algorithm for dose reduction with fluence-modulated proton CT". *Medical Physics* **47**(4) (2020), 1895–1906. ɪꜱꜱɴ: 0094-2405. ᴅᴏɪ: 10.1002/mp.14084

CONFERENCE CONTRIBUTIONS IN THIS CHAPTER

💬 [157] » J. Dickmann et al.: "Method for Fluence Field Optimization to Achieve Nonconvex Image Noise Prescriptions with Fluence-Modulated Proton CT". *Annual Meeting of the American Association of Physicists in Medicine (AAPM), San Antonio, USA, proceedings published in Medical Physics* **46**(6) (2019), E477–E477. ISSN: 24734209. DOI: 10.1002/mp.13589

💬 [158] » J. Dickmann et al.: "Dosiseinsparung durch fluenzmodulierte Protonen-Computertomographie mit Anwendung in der Protonentherapie". *Jahrestagung der Deutschen Gesellschaft für Medizinische Physik (DGMP), Stuttgart, Germany* (2019)

💬 [159] » J. Dickmann et al.: "PO-1706: Low dose fluence-modulated proton CT: simulation study and first experimental results". *Conference of the European Society for Radiotherapy and Oncology (ESTRO), Vienna, Austria (online), proceedings published in Radiotherapy and Oncology* **152**(S1) (2020), S941. DOI: 10.1016/S0167-8140(21)01724-2

6.1 FLUENCE–MODULATION IN X–RAY CT

A standard fluence–modulation approach in clinical x–ray CT is the use of bow–tie filters that follow the profile of an idealized patient and try to achieve constant noise in the image [151, 152]. This is done by making the summed attenuation of patient and bow–tie filter constant and thereby making the fluence incident to the patient homogeneous. The bow–tie filter consequently needs to be changed if a different region of the patient is to be imaged. Such systems can also under–perform, if the patient is not centered or exceptionally large or small and the filter does not match the patient profile anymore [160]. They also do not allow to focus the imaging dose to a non–central ROI.

To overcome these limitations the use of fluence modulation was suggested by Graham, Siewerdsen, and Jaffray [161], for which several control algorithms [161–170] and corresponding technical instrumentation [171–179] for dynamic modulation of the fluence field have been proposed in the last decade. For x–ray CT scanners in particular, the technical implementation of a non–static filter is not trivial. The system of Szczykutowicz and Mistretta [172] used wedges that were dynamically moved in front of the x–ray source for fluence modulation. While this worked on a C–arm system with slow rotation speeds, an implementation in a clinical scanner that rotates several times per second and that only has limited space to accommodate a system, seems challenging. The system of Stayman et al. [174] makes use of a more compact design that overlaps two gratings with a different spatial frequency causing a Moiré pattern. Thereby, only a small movement of one of the gratings is required for a relatively fast movement of the attenuation profile. Huck, Parodi, and Stierstorfer [175] used a similar approach where a single grating, similar to an anti–scatter grid, was centered on the source and small inclinations of the grating allowed a broadening and shift of the attenuation profile.

6.2 APPLICATION OF FLUENCE–MODULATION TO PROTON CT

For fluence modulation in x–ray CT the main challenge is the technical implementation of the modulated fluence delivery. PCT scanners, instead, rely for the fluence delivery on proton treatment systems that have been exactly designed for the precise and fast modulated delivery of proton fluence for treatment. A crucial requirement for FMpCT, therefore, is already solved. Dedes et al. [45] suggested the use of fluence modulation for pCT using simulations and soon after also demonstrated a first experimental feasibility [46] of the method. They used a proton fluence field generated by small pencil beams where central pencil beams had an increased dwell time compared to off–center pencil beams. By this, imaging dose at the periphery of a homogeneous phantom could be decreased in a central circular ROI while maintaining image quality therein.

The initial works of Dedes et al. [45, 46] did not use optimization to calculate the intended fluence field, but used a simple intersection–based approach that prescribed a high fluence to pencil beams for which the center intersected a central ROI and reduced the fluence for non–intersecting pencil beams. These fluence levels were arbitrarily chosen and not optimized for a specific imaging task. Consequently, the fluence inside the ROI was homogeneous, which would result in non–homogeneous image noise, as discussed in chapter 5. The challenge of a further development of FMpCT, therefore, lies in the use of a patient– or object–specific noise model to optimize a modulated fluence field achieving prescribed image noise levels.

6.3 MODELING OF PENCIL BEAMS

MATERIALS & METHODS

6.3.1 Analytical model

To realistically model proton pencil beams for fluence modulation, an analytical pencil beam model was established that matches the pencil beams at the Chicago proton center, where the experimental data for the rest of this work were acquired. From the experimental tracking data, so–called *counts maps* $C(u, v, d)$ were calculated, that contained in each voxel (u, v, d) the number of protons that intersected the pixel area around (u, v) at the binning depth d. The counts maps only considered protons that were actually used for image reconstruction and were, therefore, affected by data filtration and, in the case of phantom scans, by attenuation. In contrast to that, maps of the actual proton fluence (without attenuation and data filtration), are referred to as *fluence maps* $F(u, v, d)$. For the description of the pencil beam model, and also in the rest of this work, the point $u = v = d = 0\,\mathrm{mm}$ is located at the isocenter of the scanner.

To separate single pencil beams in the experimental data, the timestamp information of the proton tracking data was exploited, allowing to calculate a count rate in steps of $0.8\,\mathrm{ms}$. In between two pencil beam spots, the count rate dropped to zero for a few milliseconds. This short gap in time is used to adjust the scanning magnets of the PBS nozzle so that the next pencil beam can be delivered at a different position. This allowed to separate single pencil

beams. From this data, for each pencil beam b, counts maps $C_b(u, v, d)$ were calculated and fitted by the Gaussian model

$$G(u, v, d) = \frac{N_0}{2\pi\sigma'_u\sigma'_v} \cdot \exp\left(-\frac{(u - u'_0(d))^2}{2\sigma'^2_u} - \frac{(v - v'_0(d))^2}{2\sigma'^2_v}\right), \tag{6.1}$$

where N_0 was the total number of protons per pencil beam, $(u'_0(d), v'_0(d))$ was the pencil beam center at depth d. N_0 was expected to be equal for all pencil beams, since the fluence was not modulated in the experimental scan. The pencil beam center was assumed to diverge with the binning depths, and thus

$$u'_0(d) = u_0 \cdot (1 + \delta_u \cdot d). \tag{6.2}$$
$$v'_0(d) = v_0 \cdot (1 + \delta_v \cdot d), \tag{6.3}$$

where (u_0, v_0) is the pencil beam center at depth $d = 0\,\text{mm}$ and δ_u and δ_v are linear magnification factors, that captured the divergence of the beam. With this model, the central pencil beam with $u_0 = v_0 = 0\,\text{mm}$ would be parallel to the d–axis and would not diverge. The Gaussian spread in eq. (6.1) was modeled as

$$\sigma'_u = \sigma_u \cdot \sqrt{1 + \delta^2_u u^2_0}, \tag{6.4}$$
$$\sigma'_v = \sigma_v \cdot \sqrt{1 + \delta^2_v v^2_0}, \tag{6.5}$$

where σ'_u and σ'_v are the beam widths projected to a plane perpendicular to the d–axis, and σ_u and σ_v are the beam widths in beam direction.

For each counts map C_b, the fit with eq. (6.1) resulted in a set of seven open fit coefficients $c_{fit} = (N_0, u_0, v_0, \sigma_u, \sigma_v, \delta_u, \delta_v)$, by minimization of the squared deviation as

$$c_{fit} = \arg\min_{c_{fit}} \sum_{u,v,d} (C_b(u, v, d) - G(u, v, d))^2. \tag{6.6}$$

After optimization of parameters for all pencil beams, σ_u, σ_v, δ_u, and δ_v were averaged since they were not specific to a single pencil. N_0, v_0 and u_0 instead, were overwritten in a subsequent usage of the pencil beam as described later. The magnification factors δ_u and δ_v may have also been estimated from the position of the scanning magnets of the PBS nozzle, but it was chosen to leave them as open fit parameters to avoid that an incorrect assumption distorts the other fit estimates.

6.3.2 Simulation of pencil beams

Each fluence field employed in this chapter was composed of a regular grid of proton pencil beams interspaced by $\Delta_{PB,u} = 12\,\text{mm}$ in the u–direction and $\Delta_{PB,v} = 8\,\text{mm}$ in the v–direction. To avoid that opposing pencil beams (for which the rotation angle differs by 180°) were overlapping, the whole pencil beam grid was offset in u by $\Delta_{PB,u}/4 = 3\,\text{mm}$. With this approach, which is similar to the quarter detector shift used in x–ray CT, the number of pencil beams could be reduced without loss of modulation flexibility of the algorithm by increasing the interspace $\Delta_{PB,u}$ to a larger value compared to $\Delta_{PB,v}$. In the

simulation, protons were initially placed at a plane at $d_0 = -400\,\text{mm}$, which was located in front of the scanner's front tracker and was in agreement to section 5.3. The emission point was at

$$\vec{r}_0 = \begin{pmatrix} u_0 \cdot (1 + \delta_u \cdot d_0) \\ v_0 \cdot (1 + \delta_v \cdot d_0) \\ d_0 \end{pmatrix} + \begin{pmatrix} r_u \\ r_v \\ 0 \end{pmatrix}, \tag{6.7}$$

where r_u and r_v were normally distributed random numbers with a standard deviation of σ_u and σ_v, respectively. Protons within a single pencil beam were assumed to be parallel along the direction vector

$$\vec{\rho}_0 = \begin{pmatrix} u_0 \delta_u \\ v_0 \delta_v \\ 1. \end{pmatrix} \tag{6.8}$$

The pencil beam centers (u_0, v_0) were chosen according to the regular grid described above. Per pencil beam, $N_0 = N$ protons were simulated for unmodulated scans and $N_0 = m_b^\alpha N$ protons for pencil beams modulated with a modulation factor m_b^α (the parameters α and b will be introduced later). The initial energy of protons was set to $(200.00 \pm 0.66)\,\text{MeV}$, where the spread was according to the investigation in section 5.3.

6.3.3 Pencil beam optimization

Apart from the simulation of pencil beams, the pencil beam model in eq. (6.1) also served as basis function for the fluence modulation algorithm. This will be described in detail in section 6.4, and the usage of the pencil beam model is introduced here to simplify the notation in the following. For the pencil beam optimization, reference counts for pencil beam b were calculated as

$$F_b(u, v, d) = G(u, v, d)\Big|_{N_0 = N, u_0 = u_b, v_0 = v_b} \tag{6.9}$$

for a pencil beam at (u_0, v_0) according to the regular grid and a constant number of N protons for each pencil beam. The basis functions F_b allowed to represent any arbitrary counts field $C^\alpha(u, v, d)$ at a given rotation angle α by optimizing pencil beam weights w_b^α such that C^α is expressed as a linear combination of the F_b. By minimization of the squared deviation the weights can be found as

$$w_b^\alpha(C^\alpha) = \arg\min_{w_b^\alpha} \iint du\,dv \left(C^\alpha(u, v, 0) - \sum_b w_b^\alpha F_b(u, v, 0) \right)^2, \tag{6.10}$$

where the integration was performed over u and v and only at $d = 0$. The optimization was performed using the method of Nelder and Mead [180]. Please keep in mind that C are affected by attenuation while F are not. Consequently, the weights w_b^α are not yet the modulation factors m_b^α applied in FMpCT scans, but need to be corrected for the effect of attenuation, which will be described later in section 6.4.3.

Based on experimental tracking data, the Gaussian spread of the pencil beam model in eq. (6.1) were determined as $\sigma_u = (4.04 \pm 0.08)\,\mathrm{mm}$ and $\sigma_v = (5.24 \pm 0.09)\,\mathrm{mm}$ as an average over all pencil beam fit values. The stated uncertainty is the standard deviation of all values. The beam spread was considerably larger in the v–direction. The magnification factors were found as $\delta_u = (5.2 \pm 0.6) \times 10^{-4}\,\mathrm{mm}^{-1}$ and $\delta_v = (5.8 \pm 1.4) \times 10^{-4}\,\mathrm{mm}^{-1}$. The two magnification values agreed within the uncertainty bounds. They corresponded to a locations of the scanning magnets at $1/\delta_u = (1.9 \pm 0.2)\,\mathrm{m}$ and $1/\delta_v = (1.7 \pm 0.4)\,\mathrm{m}$ from the isocenter. These parameters were used together with eq. (6.1) in sections 6.4 and 6.5 to simulate and optimize pencil beam weights.

6.4 AN OPTIMIZATION ALGORITHM FOR VARIANCE PRESCRIPTION

An optimal fluence field that best achieves a given image variance target $V_{target}(x, y, z)$ can be described with help of the pencil beam model from section 6.3 as a set of fluence modulation factors $m_b^\alpha \in [0, 1]$ for pencil beam b at rotation angle α. In the following, an optimization algorithm is described, that finds an optimal set of modulation factors m_b^α for any given image variance target $V_{target}(x, y, z)$. It thereby needs to overcome the challenge that fluences and with them the modulation factors are defined in the projection domain
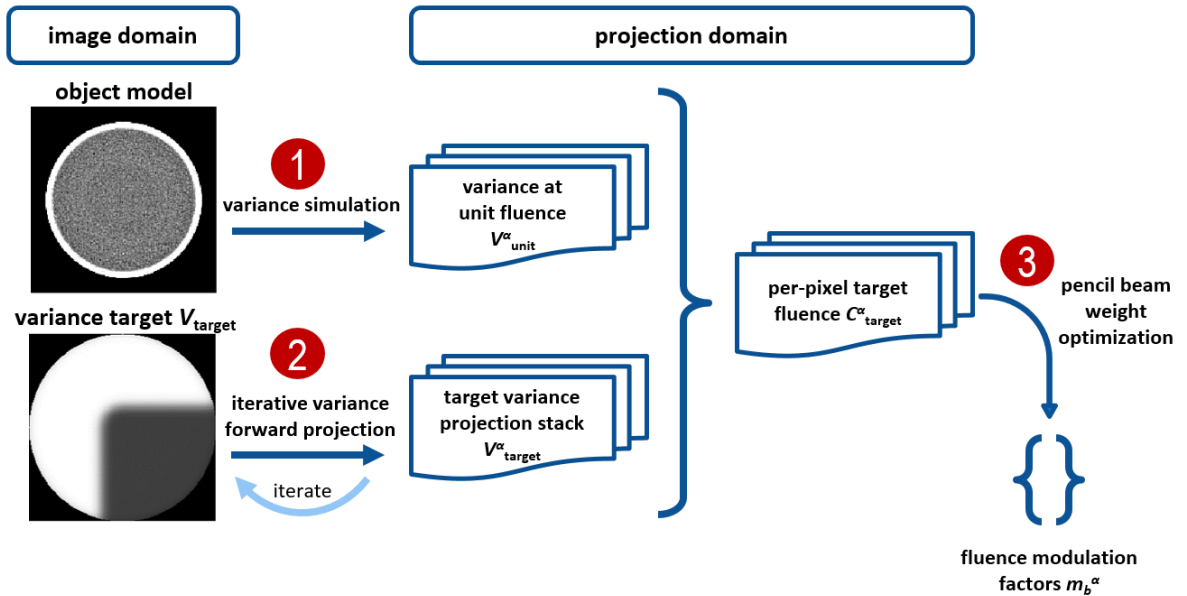
Figure 6.1: The workflow of the FMpCT optimization algorithm consists of three steps: (1) using a Monte Carlo simulation, a stack of variance projections at unit fluence is simulated, (2) an iterative approach calculates the target variance projections that yield the target image variance, and (3) a pencil beam model is fitted to the ratio of the two variance quantities.

(denoted by the coordinates (u, v, d) and the rotation angle α) while the variance target is defined in the image domain (denoted by the coordinates (x, y, z)). A general workflow of the proposed algorithm is depicted in fig. 6.1 and consists of three steps:

STEP I   Find the variance projections $V_{\text{unit}}^{\alpha}(u, v, d)$ that result from a simulation at unit fluence with $m_b^{\alpha} = 1$ for all pencil beams.

STEP II   Find the variance projections $V_{\text{target}}^{\alpha}(u, v, d)$ that, when variance reconstruction is applied to them, result in the image variance target $V_{\text{target}}(x, y, z)$.

STEP III   Using $V_{\text{unit}}^{\alpha}(u, v, d)$ and $V_{\text{target}}^{\alpha}(u, v, d)$, calculate the pixel–wise counts target $C_{\text{target}}^{\alpha}$. Then first optimize pencil beam weights w_b^{α} according to eq. (6.10), which are affected by attenuation, and subsequently calculate pencil beam modulation factors m_b^{α}, for which the effect of attenuation cancels out.

––––––––––

All three steps will be explained in detail in the following paragraphs. The proposed method builds upon ideas developed for x–ray CT by Bartolac et al. [162] and Hsieh and Pelc [164] and extends these approaches to allow for usage with pCT, by considering the three–dimensional projections due to distance–driven binning [113] and the pCT noise model [100, 124] described in chapter 5.

6.4.1   Step I: Variance at unit fluence

To calculate variance projections at unit fluence $V_{\text{unit}}^{\alpha}(u, v, d)$, the Monte Carlo simulation was employed with the Gaussian pencil beam model and all pencil beam modulation factors set to $m_b^{\alpha} = 1$, which results in a counts map of $C_{\text{unit}}^{\alpha}(u, v, d)$. This step is specific to a given phantom and the simulation needs a corresponding phantom model as described in sections 4.1 and 4.2. Were the method to be applied to patient data, a corresponding digital model could be created from a prior x–ray CT scan as described in section 4.3. If a prior pCT scan is available, $V_{\text{unit}}^{\alpha}(u, v, d)$ and $C_{\text{unit}}^{\alpha}(u, v, d)$ can be calculated from the available data and this step is omissible. Variance values from the simulated data were calculated based on the list mode WEPL data and according to eq. (3.45). The corresponding unit fluence image variance $V_{\text{unit}}(x, y, z)$ can be calculated from this data using variance reconstruction in eq. (3.62).

6.4.2   Step II: Iterative variance forward projection

Calculation of the target variance projections $V_{\text{target}}^{\alpha}(u, v, d)$ that yield the desired image noise target $V_{\text{target}}(x, y, z)$ is an inverse problem with a large set of solutions. The problem is similar to the well–understood forward projection operation, which can calculate a set of projections that yield a given image. Indeed, a first guess $V_0^{\alpha}(u, v, d)$ can be obtained by ray–tracing [181] (calculating line integrals) through $V_{\text{target}}(x, y, z)$ along straight lines of the known cone–beam geometry, which was then followed by a ramp filtration in the u–direction. The need for the additional ramp filtration can be understood since variance
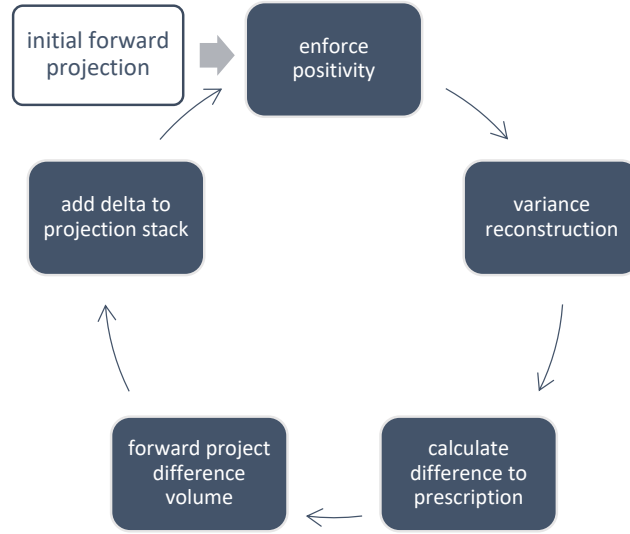
Figure 6.2: Workflow of the iterative variance forward projection in the second step of the algorithm.

reconstruction is close to an unfiltered backprojection (see eq. (3.63)), and since ray–tracing is the inverse operation of filtered backprojection, ray–tracing and ramp filtration can be an inverse operation of variance reconstruction. An additional median filtration with a kernel size of $4\,\mathrm{mm}$ was applied to $V_0^\alpha(u, v, d)$ to reduce noise introduced by the ramp filtration. Such a projection stack would yield an image variance with only small differences to the prescription. However, this stack contained negative, and thus unphysical values due to the ramp filtration and it was subsequently thresholded to positive values, introducing discrepancies with the prescription.

Starting from $V_0^\alpha(u, v, d)$, an improved estimate for $V_{\mathrm{target}}^\alpha(u, v, d)$ was found iteratively as illustrated in fig. 6.2. For this, variance reconstruction was applied to the i–th set of variance projections $V_i^\alpha(u, v, d)$, yielding the i–th variance volume $V_i(x, y, z)$. The difference to the target, $V_{\mathrm{target}}(x, y, z) - V_i(x, y, z)$ was forward–projected using ray–tracing, resulting in an update projection that was added to $V_i^\alpha(u, v, d)$ in the next iteration. Again, at every iteration, $V_i^\alpha(u, v, d)$ was forced to be strictly positive. Repeated application of these forward projection and reconstruction operations will converge to physical (i.e. positive) variance projections $V_{\mathrm{target}}^\alpha(u, v, d)$, that approach the image variance prescription $V_{\mathrm{target}}(x, y, z)$ when variance reconstruction is applied to them.

6.4.3 Step III: Fluence optimization

Variance projection values are, according to the definition in eq. (3.45), inversely proportional to the corresponding number of protons C. With this assumption, a counts target $(V_{\mathrm{unit}}^\alpha / V_{\mathrm{target}}^\alpha) \cdot C_{\mathrm{unit}}^\alpha$ could be calculated to fit pencil beam weights according to eq. (6.10). However, this linear scaling of variance values with counts only holds true, if a sufficient number of protons contribute to a pixel. If the mean proton number per pixel decreases, the variance increases over–proportionally, as shown in fig. 6.3, where the image variance
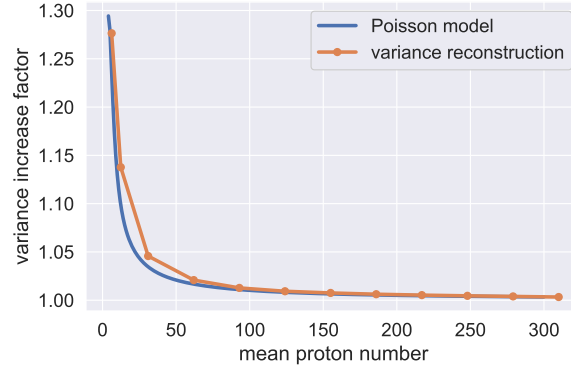
Figure 6.3: Over–proportional increase of image variance for very low proton numbers and modeling using a Poisson assumption. The variance increase factor is defined as the actual variance divided by the expected variance using the linear approximation. ⓒ Reproduced from Dickmann et al. [156] under the CC BY 4.0 license.

for a mean proton number C per pixel is divided by the assumption using eq. (3.45). The increase for low counts can be modeled by recognizing that the variance in a given pixel is calculated from C protons on average, but C can vary and is Poisson distributed. This leads to the correction function

$$k(C) = C \cdot \sum_{n=1}^{\infty} P_C(n) \cdot \beta_{n,C} = C^2 \cdot \sum_{n=1}^{\infty} \frac{P_C(n)}{n}, \tag{6.11}$$

where $P_C(n) = C^n \exp(-C)/n!$ is the Poisson probability of detecting n protons instead of the expected C, and $\beta_{n,C} = C/n$ is the relative variance change when detecting n protons instead of C. The sum in $k(C)$ converges and was calculated numerically by summing the first 1000 summands. It is also plotted in fig. 6.3. Moreover, the function was thresholded to return at least $C_{min} = 8$ protons, which avoided detector elements with missing information.

With this, the target fluence could be calculated as

$$C_{target}^{\alpha}(u, v, d) = k \left[\frac{V_{unit}^{\alpha}(u, v, d)}{V_{target}^{\alpha}(u, v, d)} \cdot C_{unit}^{\alpha}(u, v, d) \right], \tag{6.12}$$

where $V_{target}^{\alpha} > V_{unit}^{\alpha}$ was enforced by thresholding, since modulation factors were searched only in the interval $[0, 1]$, thus only allowing fluence reductions with respect to the unit fluence scan.

Both C_{target} and C_{unit} are affected by attenuation, and for both the pencil beam optimization according to eq. (6.10) would result in reduced pencil beam weights w_b^{α}. For this reason, pencil beam weights were calculated for both counts maps and the final fluence modulation factors were found as the ratio

$$m_b^{\alpha} = \frac{w_b^{\alpha}(C_{target}^{\alpha})}{w_b^{\alpha}(C_{unit}^{\alpha})}. \tag{6.13}$$

This normalization effectively canceled out the effect of attenuation and the optimized modulation factors m_b^{α} could be used to simulate FMpCT scans according to section 6.5. If the factors were outside of the interval $[0, 1]$, they were thresholded accordingly.

6.4.4 Reference approach

In addition to the algorithm above, a reference approach was implemented, that was used in previous works of FMpCT [45, 46]. In this simple approach, pencil beam modulation factors are assigned depending if their central axis intersected the ROI or not. This leads to the expression

$$m_b^\alpha = \begin{cases} 1 & \text{if intersecting} \\ \gamma & \text{otherwise} \end{cases} , \tag{6.14}$$

where $0 < \gamma < 1$ is the modulation strength, which in the following was chosen to be equal to the contrast in $V_{target}(x, y, z)$ of the proposed method.

6.5 EVALUATION OF OPTIMIZED FMPCT SCANS

MATERIALS & METHODS

To test the performance of the proposed algorithm, the Monte Carlo simulation of the pCT scanner was used, where pencil beams could be simulated according to section 6.3 with the optimized modulation factors of section 6.4. For this purpose, modulation plans for three different variance targets, as shown in fig. 6.4, were calculated and in each prescription, a ROI region with low variance and, in two cases, a non–ROI region with high variance were identified. The three prescriptions were

- a constant variance V_{ROI} throughout the entire volume,

- FMpCT prescription A with variance V_{ROI} in one quadrant of the image and $4 \cdot V_{ROI}$ outside, and

- FMpCT prescription B with variance V_{ROI} in a rectangular central region of the image and $4 \cdot V_{ROI}$ outside.

The variance contrast of four was chosen in agreement to chapter 5 to be higher than the expected variations of a uniform fluence scan, but to still be reasonably low to avoid distortions of the RSP values inside the ROI, which was observed by Dedes et al. [45] for high binary modulations. The targets were slightly blurred to allow for a smooth transition between the low– and high–noise regions.

The prescription of constant image variance was motivated by the fact that uniform fluence pCT scans do not show uniform noise, as investigated in chapter 5 and by Rädler et al. [124]. With a constant noise prescription, imaging dose could be saved at the same peak noise level, if the entire image is used diagnostically. Instead, prescriptions A and B model assume cases of treatment planning, where the treatment beams could come from 90 and 180 degrees in A and from 90 and 270 degrees in B. In the following, the terms *constant*, *A*, and *B* are used to refer to the three prescriptions.

Simulations were performed for the water phantom, the CTP phantom and the head phantom. Initially, a unit fluence scan with modulation factors $m_b^\alpha = 1$ was generated, that served as a reference and also to guide the FMpCT optimization. The number of protons per

constant variance prescription — FMpCT prescription A — FMpCT prescription B

C/W = (2/4) V_{ROI}
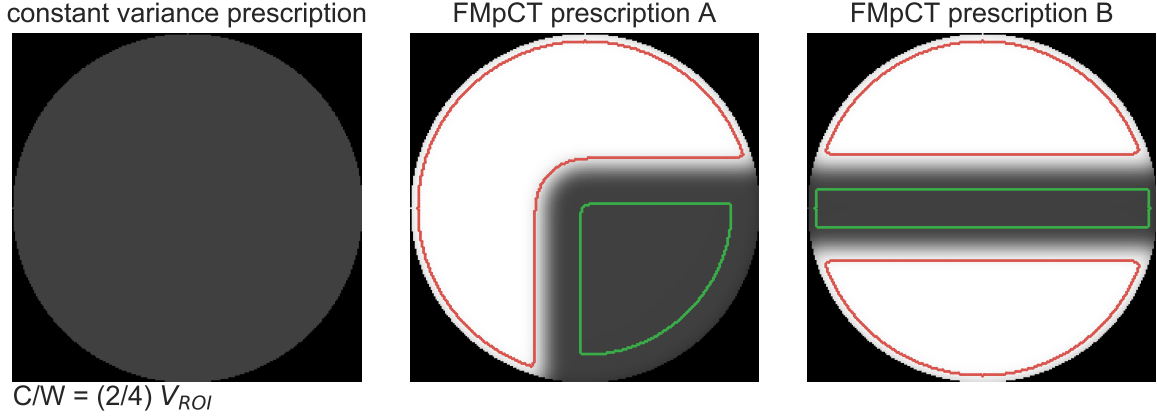
Figure 6.4: Image variance prescriptions and indications of the ROI (green) and non–ROI (red) volume for the three variance targets in this study. For the constant variance prescription, the whole volume was the ROI.

pencil beam N was chosen to yield an imaging dose of 1.4 mGy in the unit fluence scans, which is a typical value for pCT [82]. From these scans, the value of V_{ROI} was determined for each phantom as the 95–th percentile value of the variance inside the object. This was

- $V_{ROI} = 4.61 \times 10^{-4}$ for the water phantom (standard deviation 2.1 %),

- $V_{ROI} = 5.89 \times 10^{-4}$ for the CTP phantom (standard deviation 2.4 %), and

- $V_{ROI} = 11.96 \times 10^{-4}$ for the head phantom (standard deviation 3.4 %),

which is consistent with chapter 5.

To allow for a fair comparison of image variance and imaging doses, the resulting pCT and FMpCT scans were normalized based on the 95–th percentile variance value v_{95}^{ROI} inside the ROI. Subsequently, for every acquisition, the imaging doses were multiplied by $\eta = v_{95}^{ROI}/V_{ROI}$ and image variances were multiplied by $1/\eta$. The choice of the 95th–percentile value was chosen as a compromise between tolerating outliers and the requirement of achieving the prescription dose V_{ROI} (or lower) inside the ROI. For the water phantom and the CTP phantom only the central slice was evaluated, while for the head phantom the entire volume was considered for the calculation of v_{95}^{ROI} and V_{ROI}. For each phantom, the hull was determined by a threshold on RSP of 0.5 which was subsequently eroded by 7 mm and values outside of the hull were disregarded. The erosion was used to avoid that increased variance values at the edge of the phantom would cause an unreasonably high skin dose in this region.

RESULTS

Figure 6.5 (a) shows the cost functions as a function of the iteration number for the iterative variance forward projection in step II of the algorithm. It is shown for all three optimization targets and is valid for all three phantoms, since step II does not depend on the phantom
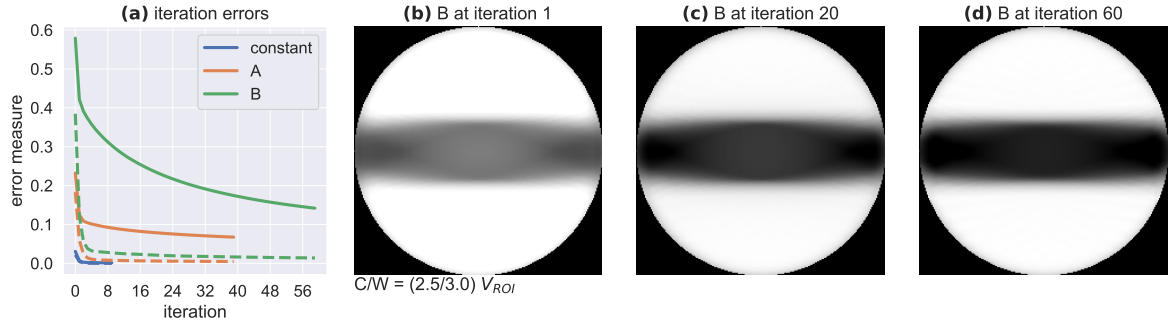
Figure 6.5: Iterative variance optimization with (a) root–mean–square error (solid) and mean error (dashed) values as a function of the iteration and for the three variance prescription, and (b) – (d) intermediate image variance maps for three different iterations and for variance prescription B. ⓒ Reproduced from Dickmann et al. [156] under the CC BY 4.0 license.
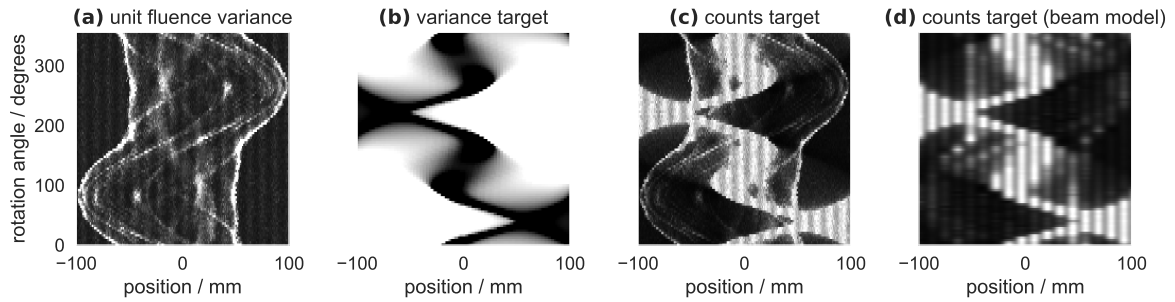
Figure 6.6: Sinograms as intermediate results of the fluence optimization with prescription A: (a) variance at unit fluence $V^{\alpha}_{\text{unit}}(u, 0, 0)$, (b) projection variance target $V^{\alpha}_{\text{target}}(u, 0, 0)$, (c) counts target $C^{\alpha}_{\text{target}}$, and (d) the counts target as fitted by the pencil beam model. The grayscale display was $C = 0.4\,\text{mm}^2, W = 0.8\,\text{mm}^2$ for variances and $C = 80, W = 160$ for counts. ⓒ Reproduced from Dickmann et al. [156] under the CC BY 4.0 license.

geometry. For the constant noise target, the cost function immediately drops to a minimum close to zero. Instead, for variance prescriptions A and B the optimization takes more iterations to converge. For both, the mean error reduces much more quickly than the RMS error. In particular, prescription B required 60 iterations to bring the cost function down. The RMS error was changed by less than 1 % per iteration, when the iterative algorithm was stopped for all prescriptions. Figure 6.5 (b) – (d) shows the intermediate variance estimate $V_i(x, y, z)$ for prescription B and at iterations 1, 20 and 60. While remaining errors persist, the achieved contrast and gradient is satisfactory. There is a considerable improvement in between the displayed iterations.

In fig. 6.6, fluence and variance sinograms are shown for the fluence optimization for prescription A and the head phantom. The sinograms display the quantity at the central slice and binning depth $v = d = 0\,\text{mm}$ and for all rotation angles. Figure 6.6 (a) displays the variance at unit fluence $V^{\alpha}_{\text{unit}}(u, v, d)$, which is calculated in step I of the algorithm and is elevated at the edge of the phantom as well as close to internal heterogeneities. A slight modulation with the interspaced pencil beams is visible as vertical strips of increased variance throughout the sinogram. In fig. 6.6 (b), the variance target $V^{\alpha}_{\text{target}}(u, v, d)$ is shown,

which is the output of step II of the optimization algorithm. Regions centrally intersecting with the ROI generally require a low variance and regions not intersecting generally require a high variance, but there is a smooth transition in between those regions. In fig. 6.6 (c), the counts target $C_{target}^{\alpha}(u,v,d)$ from eq. (6.12) is displayed, which calculates from (a) and (b). Some parts of the sinogram are assigned a value of zero in (b) and consequently required the unit fluence counts in (c). Eventually, fig. 6.6 (d) is the counts target in (c) as fitted by the Gaussian pencil beam model in step III. The values in (d) can be calculated as

$$\sum_b w_b^{\alpha} F_b(u,v,d). \tag{6.15}$$

Slight differences to (c) are observed, apart from subtle features at the edge that could not be resolved using the pencil beams which are larger than the counts grid pixel size.
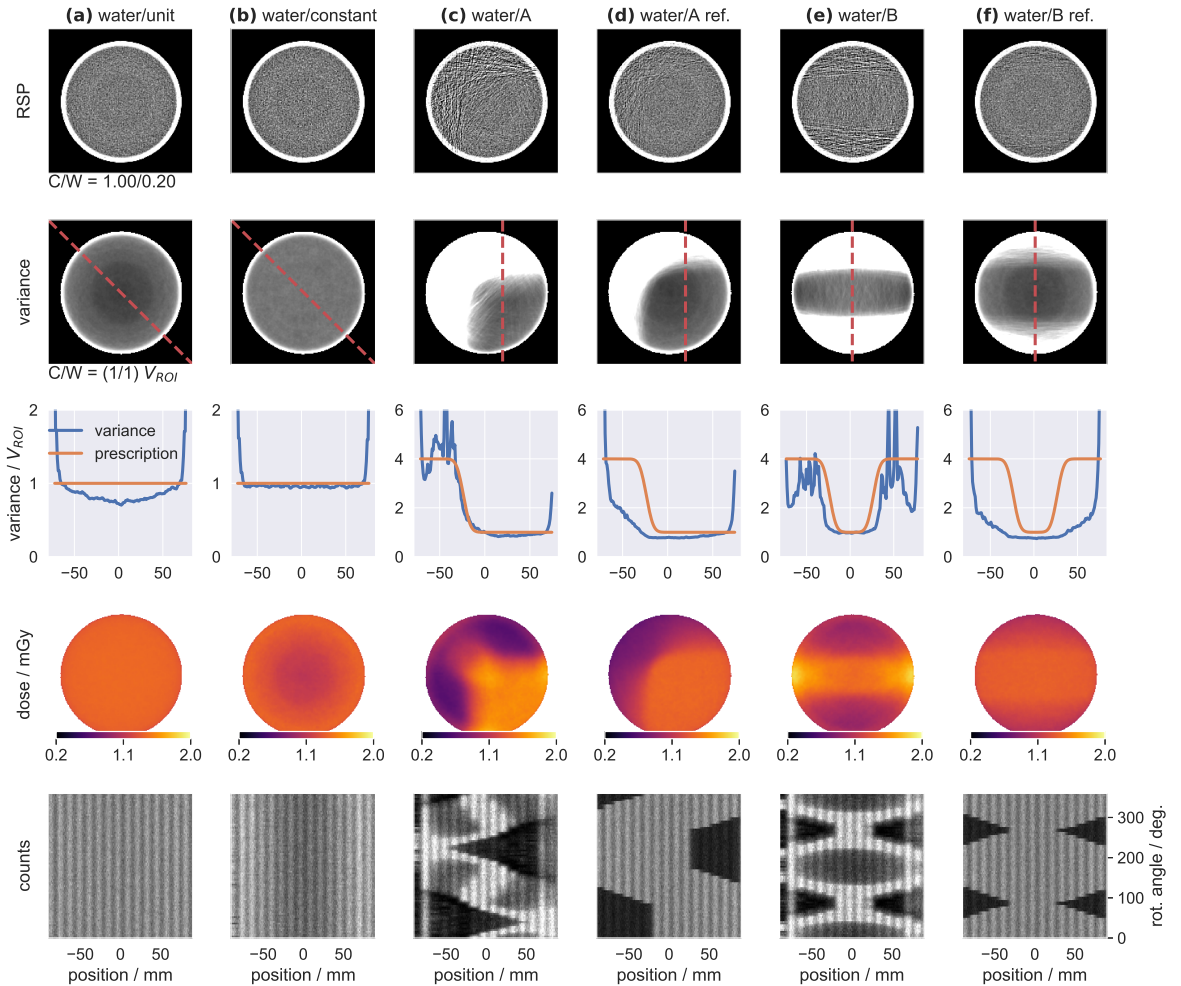
Figure 6.7: Simulation study for the fluence optimization algorithm and for the reference method with the water phantom. In the first three rows, RSP and variance maps are shown with a profile through the variance along the red dashed lines. In the last two rows the imaging dose and counts sinograms are shown. For the counts sinograms the display setting was $C = 120, W = 240$. © Reproduced from Dickmann et al. [156] under the CC BY 4.0 license.
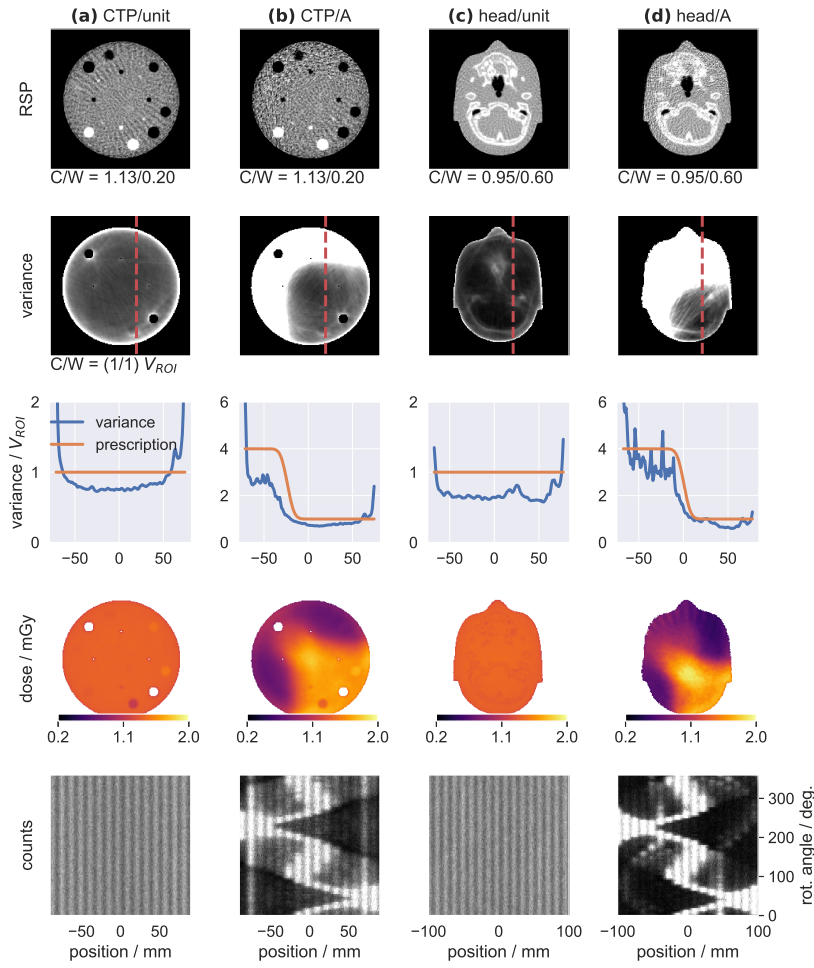
Figure 6.8: Simulation study for the fluence optimization algorithm and for the reference method with the CTP and the head phantom. In the first three rows, RSP and variance maps are shown with a profile through the variance along the red dashed lines. In the last two rows the imaging dose and counts sinograms are shown. For the counts sinograms the display setting was $C = 120, W = 240$. ⓒ Reproduced from Dickmann et al. [156] under the CC BY 4.0 license.

Figures 6.7 and 6.8 show the results of the simulation study for the water phantom (fig. 6.7) and the CTP and the head phantom (fig. 6.8). They display, row by row, the RSP maps, image variance maps, profiles through the image variance maps along the red dashed line, the imaging dose and, in the last row, fluence sinograms. Figure 6.7 (a) shows the water phantom at the uniform unit fluence. The dose consequently is homogeneous, but variance is reduced in the center of the phantom, as was observed before. The pencil beams are visible in the fluence sinogram, but, summed over all projection angles, the fluence was constant, and so was the imaging dose, since the pencil beam pattern was shifted by a quarter interspace. Figure 6.7 (b) displays the water phantom optimized for a constant image noise. While in the RSP map only minor differences can be observed, the variance map and the corresponding profiles are flat as intended. At the same time, fluence and dose are reduced in the center.

The results for the fluence optimization for FMpCT task A are shown in fig. 6.7 (c). Already the RSP maps features an increased noise level outside of the ROI. This is confirmed in the variance map and the corresponding profile, which shows that the target variance agreed well with the prescription inside the ROI, while outside variance fluctuated around the desired value. In contrast to that, fig. 6.7 (d) shows the results of the reference approach, which also resulted in a variance increase outside of the ROI, but failed to achieve the prescription. In particular, noise is not flat inside the ROI and the variance increase is less steep outside of it. The binary modulation of the reference approach can also be seen in the counts sinogram. A similar trend can be seen in fig. 6.7 (e) and (f), where the fluence modulation using the proposed method and the reference approach for FMpCT target B are shown. For the proposed method, the variance increase is steeper and agreement with the prescription is better. The imaging dose outside of the ROI is also lower. At the same time, imaging doses using the proposed method inside the ROI are increased compared to the unit fluence case and also compared to the reference method. In particular in fig. 6.7 (c) fine streaks of high variance are seen that enter the otherwise flat noise level inside the ROI.

In fig. 6.8 the corresponding evaluations are shown for the CTP and the head phantom. The CTP phantom in fig. 6.8 (a) and (b) has a sharp increase of variance also at the air insert, which was included in the ROI. The variance contrast observed using the proposed method is less steep as for the water phantom and variance in the ROI is slightly below the prescription. Nevertheless, imaging dose is decreased considerably outside of the ROI and slightly increased inside. The same can be observed for the head phantom in fig. 6.8 (c) and (d), where the variance in the unit fluence scan was even more heterogeneous. Variance is increased considerably outside of the ROI and the prescription is met inside. Imaging dose again is decreased outside of the ROI, but increased inside. The fluence sinogram of the head does not only follow the ROI structure, as for water and CTP phantom, but also features clearly visible fluence increases that compensate for heterogeneities of the phantom.

For the CTP phantom, the RSP values of the two inserts inside the ROI and the body were determined in the fluence modulated scan to be 1.776 (−0.8 % error) for Teflon, 0.881 (−0.2 % error) for PMP, and 1.143 (−0.1 % error) for the body made from epoxy. In the unit fluence scan, this was 1.776 (−0.8 % error), 0.879 (−0.4 % error), and 1.143 (−0.1 % error). These differences are small compared to the difference of both scans to the ground truth values from table 3.1, which is reported as the error value in brackets.

Imaging doses for all evaluations are summarized in fig. 6.9. The dashed line indicates the mean imaging dose over the whole phantom of the unit fluence scan, which was 1.4 mGy for the water phantom and which was slightly reduced in the two other phantoms due to different material compositions. The unit fluence was equal in all scans. Prescribing a constant image variance to the water phantom saved 8.9 % of the imaging dose at equal peak noise level. For both FMpCT targets, the mean imaging dose reduced when using the proposed method. Inside the ROI, the imaging dose increased by between 9.2 % and up to 19.2 %. At the same time, the dose reduction outside of the ROI was between 35.4 % and up to 40.5 % using the proposed method. The reference method did not increase the dose inside the ROI, but dose reductions on average and in particular outside of the ROI were less. Outside the ROI, the dose reduction was only between 13.2 % and 29.2 %.
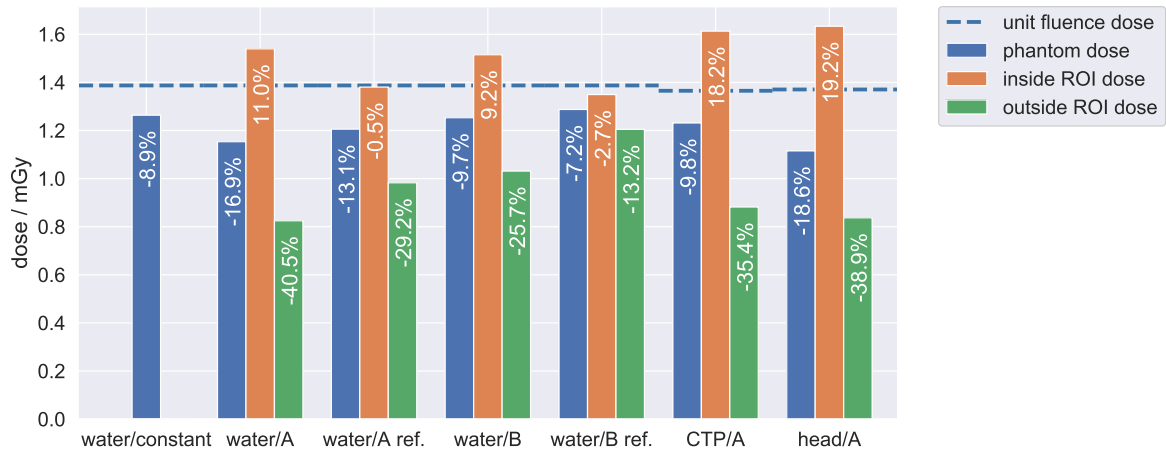
Figure 6.9: Imaging doses of the simulation study testing the performance of the optimization algorithm. The dashed line indicates the unit fluence dose which changed slightly for the three phantoms even though the corresponding fluence was equal. ⓒ Reproduced from Dickmann et al. [156] under the CC BY 4.0 license.
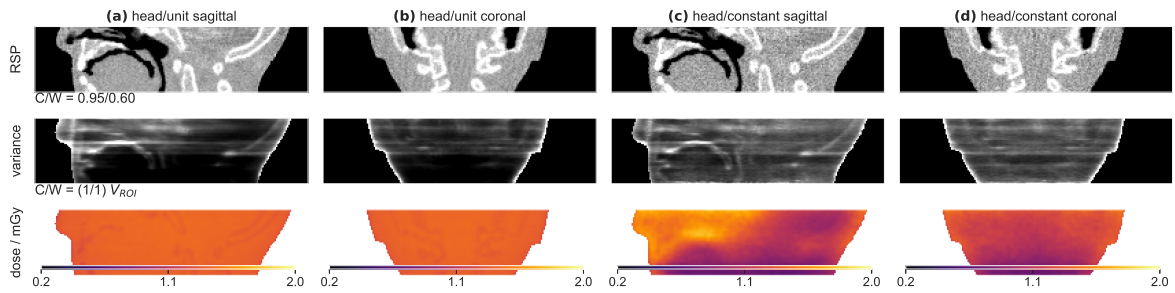
Figure 6.10: Imaging dose for the head phantom with (a), (b) uniform fluence, and (c), (d) the fluence optimized to achieve constant noise in the image. ⓒ Reproduced from Dickmann et al. [156] under the CC BY 4.0 license.

Figure 6.10 shows the head phantom in a (a), (c) sagittal and (b), (d) coronal view, where (a) and (b) are the RSP, image variance and imaging dose for the unit fluence acquisition and (c) and (d) were optimized for a constant variance target. Again, variances and doses were normalized to the peak variance. The variance of the head phantom in the unit fluence scan changes dramatically between the nasal cavity and the back of the head. This is partially compensated in the fluence modulated scan, which allowed to decrease the imaging dose in the region were image variance was low in the unit fluence scan. The dose over the whole phantom changed from 1.37 mGy to 1.15 mGy, which is a reduction by 16 %.

6.6 DISCUSSION

Modeling of pencil beams

Based on experimental tracking data, a Gaussian pencil beam model could be established that allows to describe realistic pencil beams with arbitrary fluences and at any location within the scanner's field of view. This was a prerequisite to allow for an optimization of pencil beam modulation factors in step III of the proposed method, where the pencil beams served as basis functions to produce the required fluence. They also allowed to perform Monte Carlo simulations of scans with arbitrary fluence modulations.

The uncertainty of fits in the u–direction was smaller compared to the uncertainty in the v–direction, since fit parameters were found as an average over all pencil beams and the scanner's aperture is larger in this direction. Interestingly, the beam spread σ_u was considerably smaller than σ_v. Such an anisotropy is not expected for clinical operation of the treatment system, but could be caused by the fact that energy and momentum slits needed to be closed beyond what is commissioned for clinical operation in order to keep the fluence low and viable for the scanner.

Evaluation of optimized FMpCT scans

The iterative variance forward projection in step II of the proposed algorithm was employed to calculate a stack of variance projections that best reproduce the target image variance. Depending on the prescription, this required a different amount of iterations and in particular many, if the initial forward calculation step resulted in negative variance values, which needed to be corrected in the following iterations. The constant variance prescription was the easiest case, since it did not require negative variance values (a constant stack of variance projections would result in a constant image variance). The two FMpCT prescriptions A and B required 40 and 60 iterations to reach a satisfactory agreement with the variance target. Remaining discrepancies between the final variance projections stack and the target in image space may impact the achievable variance contrast. They do not impact the accuracy of the achieved variance in the ROI, since they could be linearly rescaled such that V_{ROI} was achieved inside the ROI.

The fluence optimization in step III of the algorithm allowed to calculate pencil beam modulation factors. For this step, only $m_b^\alpha \leqslant 1$, and consequently fluences less than or equal to unit fluence, were allowed to avoid unreasonably high doses to the skin, which would otherwise result from the elevated noise at the edge of phantoms as expected. An increased noise at the very edge may be acceptable, since this is a limited area traversed by a therapeutic proton beam. To avoid missing data, proton fluences were also forced to be at least C_{min}. Both thresholds may impact the achievable variance contrast, but again not the accuracy of the resulting variances in the ROI due to the rescaling with the correction factor η.

The proposed optimization algorithm was subsequently used in a simulation study with three phantoms and and three different image noise targets. The first optimization was to achieve constant image variance across the phantom. This assumes, that the entire volume is

required and must be imaged with low image noise. When compared at equal peak noise level to a uniform fluence scan, where noise is elevated at heterogeneities, this can result in a dose saving. For the water phantom, this dose saving was estimated to be 8.9 %. This completes the first modulation attempt in fig. 5.7. For the head phantom, which contains more heterogeneities, the dose saving was increased and estimated to be 16.0 % compared to a uniform fluence scan. This is equivalent to an improvement of the signal–to–noise ratio by 35 % at the same dose.

For the two FMpCT prescriptions, considerable dose savings between 25.7 % and 40.5 % could be demonstrated outside of the ROI using three different phantoms. The dose saving primarily depended on the shape of the variance prescription, with noise target A achieving better dose savings compared to target B. At the same time, imaging dose was increased inside the ROI. This could be caused by high variance streaks that entered the ROI and, through the normalization with η, increased the dose. However, assuming that the ROI agrees with the treatment beam path, this dose increase in the ROI, may be negligible, since doses from treatment will be orders of magnitude higher compared to the imaging doses. The dose saving outside of the ROI is more important, since particle therapy allows for minimal doses outside of the therapeutic beam.

Using the sensitometric CTP phantom, the RSP accuracy of pCT and FMpCT scans was shown to be comparable and differences were all below 1 %, which is within the magnitude expected for experimental pCT scans [41].

The performance of the proposed method was compared, in terms of achievable dose reduction, to a simple intersection–based approach that was used in earlier works [45, 46]. For the reference method, the achieved variance was less conformal to the ROI and dose savings were considerably lower. The prescription of constant image variance is, by construction, not possible with this approach.

The proposed algorithm suffered from minor streaks of variance that entered the ROI as well as a dose increase inside the ROI. This may have impaired an optimal usage of the imaging fluence and could be related to the fact that the optimization just considered variance values and only implicitly reduced imaging dose by increasing noise. In addition, the optimization was projection–based and a violation of the target variance in one projection could not be compensated by fluence from another projection. In chapter 8, an alternative FMpCT optimization algorithm is presented, that optimizes pencil beam modulation factors for both variance and dose objectives and thereby improves results. It also investigates the dependency of the achievable dose saving on the shape of the ROI and the object by using patient data and deriving the ROI shape from actual treatment plans.

6.7 CONCLUSIONS

In this chapter and originally in Dickmann et al. [156], a three–step optimization algorithm for FMpCT was developed. It makes use of an iterative variance forward projection technique to transform the variance prescription from image space to projection space, where the modulation factors are defined. With this, the subsequent optimization of pencil beams was comparably simple and fast. The dose savings achievable with the proposed method were

demonstrated in a simulation study, where the performance was superior to an intersection–based reference method. In particular, the prescription of constant image variance allowed to save dose while maintaining the peak variance level in the entire volume. The dose saving achievable with specific FMpCT imaging tasks depends on the shape of the ROI and the phantom, but was within 25.7 % and 40.5 % for two representative ROI shapes.

7 EXPERIMENTAL REALIZATION OF FMPCT

THE OPTIMIZATION ALGORITHM PRESENTED IN THE LAST CHAPTER paved the way towards an experimental delivery of optimized fluence patterns for an experimental realization of FMpCT. The pencil beam model, the realistic scanner simulation, and the use of digital phantoms with physical counterparts allowed to apply the patterns optimized with simulations in an acquisition with the prototype pCT scanner.

In this chapter, the beamline of the Chicago proton center was interfaced to deliver arbitrary fluence patterns with small, modulated pencil beams. The resulting acquisitions were then compared to the planned scans in terms of achieved variance levels and RSP accuracy. A spatial variation of the incident energy within a single pencil beam initially distorted the RSP accuracy, but could be corrected in post–processing by fitting and subsequently subtracting the energy dependence. Additionally, a slight misalignment of the fluence patterns, which were purposefully shifted by a quarter pencil beam, impaired the fluence delivery and could not be corrected in postprocessing. However, distortions were small compared to the overall planned modulations. The performance of the optimization algorithm on experimental scans was tested on three phantoms and by prescribing constant noise as well as a more complex FMpCT image noise target. Overall, the agreement between the simulated and the experimentally acquired data was satisfactory both in terms of image variance and RSP accuracy. While improvements were suggested for future experimental FMpCT scans, the feasibility of using modulated pencil beams to achieve task–specific image noise distributions with FMpCT can be confirmed.

The results presented in this chapter were published in *Physics in Medicine & Biology* [99]. Parts of the results were presented at the joint meeting of AAPM & COMP in Vancouver, Canada [182], at the CT–Meeting in Regensburg, Germany [183], and at the ESTRO conference in Vienna, Austria [159].

PUBLICATIONS IN THIS CHAPTER

[99] » J. Dickmann et al.: "Experimental realization of dynamic fluence field optimization for proton computed tomography". *Physics in Medicine & Biology* **65**(19) (2020), 195001. DOI: 10.1088/1361-6560/ab9f5f

CONFERENCE CONTRIBUTIONS IN THIS CHAPTER

[182] » J. Dickmann et al.: "Prescribing Image Noise Using Dynamic Fluence Field Optimization: Experimental Results Using a Pre-Clinical Proton CT Scanner". *Joint*

Annual Meeting of the American Association of Physicists in Medicine and the Canadian Organization of Medical Physics (AAPM/COMP), Vancouver, Canada (online), proceedings published in Medical Physics **47**(6) (2020), 2580–2674. ISSN: 24734209. DOI: 10.1002/mp.14315

💬 [183] » J. Dickmann et al.: "Dynamic Fluence Modulation using Proton CT for Low-dose Imaging in Particle Therapy". *International Conference on Image Formation in X-Ray Computed Tomography (CT Meeting), Regensburg, Germany (online)* (2020)

💬 [159] » J. Dickmann et al.: "PO-1706: Low dose fluence-modulated proton CT: simulation study and first experimental results". *Conference of the European Society for Radiotherapy and Oncology (ESTRO), Vienna, Austria (online), proceedings published in Radiotherapy and Oncology* **152**(S1) (2020), S941. DOI: 10.1016/S0167-8140(21)01724-2

7.1 DELIVERING FLUENCE–MODULATED SCANS AT THE PROTON CENTER

MATERIALS & METHODS

The goal of this work was to experimentally employ the fluence patterns that were optimized in chapter 6 using the prototype pCT scanner described in section 3.2 for the same phantoms and image noise targets. In the following, these targets will be referred to as *noise plans* in analogy to treatment plans that describe the delivery of the therapeutic dose using a computational model of the patient. To deliver the plans, the control system of the PBS beamline at the Northwestern Medicine Chicago proton center, where the scanner is located, was interfaced. For this purpose, the coordinates of the pencil beam grid as well as their relative weights needed to be transformed to machine instructions. A script, which was specifically developed for this study, converted pencil beam coordinates to corresponding currents of the beam deflecting magnets. This required proprietary beam line information and was done by an employee of *Ion Beam Applications SA* (IBA, Louvain–la–Neuve, Belgium). The relative spot intensities were realized by keeping the beam current constant at $1.3\,\mathrm{nA}$ and modulating the dwell time of each pencil beam spot. This strategy, which was already successfully used by Dedes et al. [46], allows for the fastest delivery of the fluence pattern and corresponds to the way the system is used during treatment.

The maximum dwell time, which corresponds to a relative weight of one, was chosen to yield approximately four times the number of protons used in the corresponding simulations. While the correct dwell time could have been fine–tuned on the spot, this conservative approach of acquiring more data than necessary was chosen due to the limited available beam time. To select the correct number of protons in postprocessing, the number of hits in a central region of the front tracker in a scan without phantom was compared between experiment and simulation and an acceptance ratio r_{accept} was calculated. In subsequent experimental scans, protons were randomly accepted with a probability equal to r_{accept} to match statistics between the experiment and the corresponding simulations.

A direct communication between the PBS delivery system and the scanner was not possible and yet the fluence patterns needed to be delivered in synchrony with the rotation angle and acquisition of the pCT scanner. For this reason, the phantom was not rotated continuously, but at fixed time intervals large enough to allow for a manual initiation of the corresponding fluence pattern. This increased the acquisition time per tomography to around 90 min, which is considerably longer compared to the 6 to 10 min, which are typically needed with this scanner when the phantom is rotated continuously and the (broad) beam is always on.

A retrospective analysis of the acquired data showed a pattern of regularly spaced rings in the reconstructed RSP images, which coincided with the spacing of the employed pencil beam grid. These rings were attributed to a spatially varying energy distribution within a single pencil beam. The proton energy incident to the object is not measured by the scanner and assumed to be 200 MeV. A spatially varying shift of the mean energy consequently results in distorted WEPL measurements and can result in rings in the image, if pencil beams are small and regularly spaced, as in this work.

The energy shift was isolated and subtracted in subsequent acquisitions. For this, data without any absorber, for which an average WEPL of 0 mm is expected, was split into contributions from single pencil beams at the drop in count rate as described in section 6.3. For each pencil beam p, the center of mass coordinate (u_p, v_p) was calculated at the front tracker. Then, for each proton within that pencil beam the coordinates $(\tilde{u}, \tilde{v}) = (u - u_p, v - v_p)$ relative to the pencil beam center were determined. The spatial distribution of WEPLs of all protons of all pencil beams was fitted in coordinates relative to the respective pencil beam center using the quadratic correction function

$$\Delta_{\text{WEPL}}(\tilde{u}, \tilde{v}) = a + b_u \tilde{u} + b_v \tilde{v} + c_u \tilde{u}^2 + c_v \tilde{v}^2, \tag{7.1}$$

where $c_{\text{fit}} = (a, b_u, b_v, c_u, c_v)$ were the free fit parameters, which were determined as

$$c_{\text{fit}} = \arg\min_{c_{\text{fit}}} \sum_{n=1}^{N} (\Delta_{\text{WEPL}}(\tilde{u}_n, \tilde{v}_n) - W_n)^2, \tag{7.2}$$

where the sum is over all N protons, $(\tilde{u}_n, \tilde{v}_n)$ is the coordinate of proton n relative to the corresponding pencil beam center, and W_n is the WEPL measured for proton n.

To correct subsequent scans, again pencil beams were isolated and center of mass coordinates were calculated for each pencil beam. Subsequently, the WEPL W of a proton with relative coordinates $(\tilde{u}, \tilde{v})$ was replaced by $W' = W - \Delta_{\text{WEPL}}(\tilde{u}, \tilde{v})$.

To investigate the effect of the pencil beam correction, protons of all pencil beams were binned to their relative coordinates $(\tilde{u}, \tilde{v})$ to calculate a two–dimensional map of the proton number and the WEPL. From the maps of the proton number, the FWHM was calculated along both axes and is referred to as $\text{FWHM}_{u,\text{exp}}$ and $\text{FWHM}_{v,\text{exp}}$.

RESULTS

Figure 7.1 (a) shows the number of hits at the front tracker in relative coordinates $(\tilde{u}, \tilde{v})$ summed for all pencil beams in an acquisition without absorber. The expectation value, which is marked with a red cross is, by construction of the relative coordinates, at the center.
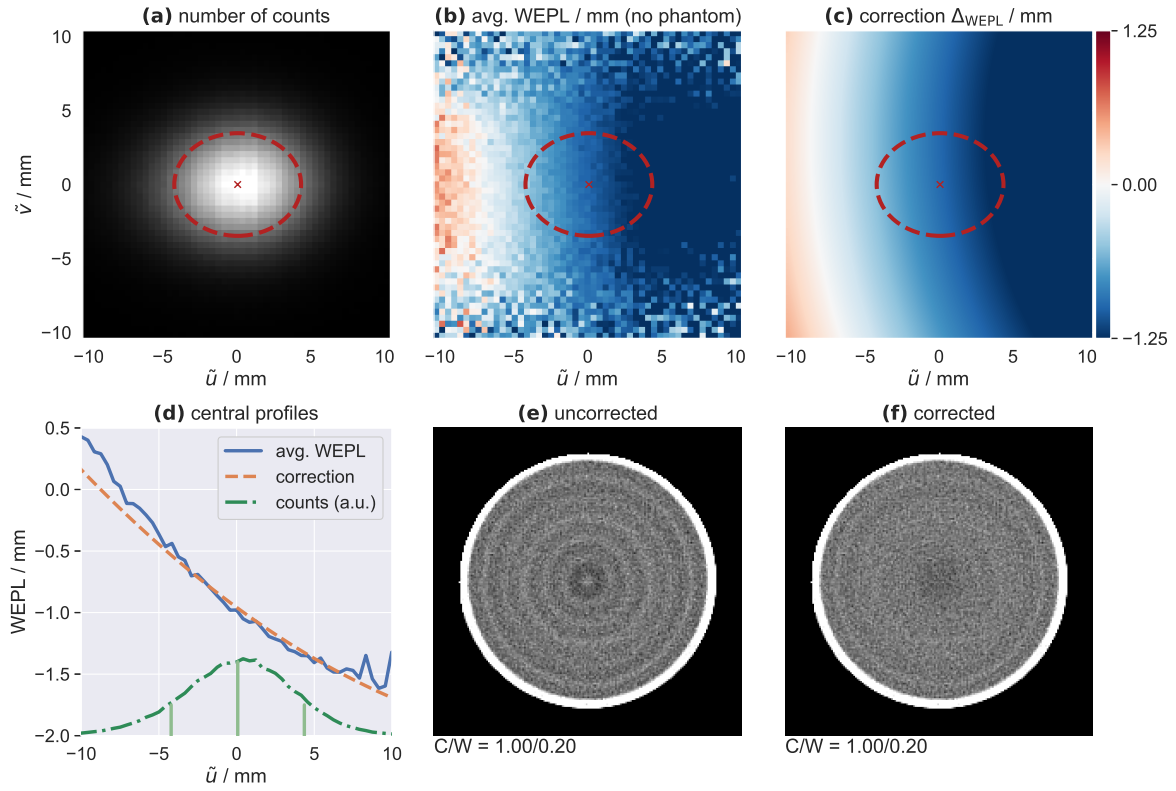
Figure 7.1: Data used for the pencil beam spot WEPL correction: (a) proton counts in relative coordinates summed for all pencil beams, (b) the average WEPL in a scan without phantom, (c) the WEPL correction function, (d) profiles through the average WEPL and the correction function as well as corresponding counts in arbitrary units, (e) an uncorrected and (f) a corrected scan of the water phantom. In (a) – (c) the elliptical beam spot is indicated as its full–width at half maximum in both directions and in (d) with the counts profile as a projection to the ũ–axis. (b) and (c) share the same colorscale.

The beam spot is not circular, but elliptical with the FWHM as indicated by the red dashed line. The FWHM along both axes was determined as

$$\text{FWHM}_{u,\text{exp}} = 8.6\,\text{mm} \tag{7.3}$$

$$\text{FWHM}_{v,\text{exp}} = 6.9\,\text{mm} \tag{7.4}$$

and, therefore, is smaller in the ũ–direction compared to the ṽ–direction. Such an anisotropy was already found in section 6.3. However, in these previous data which were used for the fluence plan optimization, the FWHM was determined as $\text{FWHM}_u = 9.5\,\text{mm}$ and $\text{FWHM}_v = 12.3\,\text{mm}$. This is comparable in the ũ–direction, but considerably larger in ṽ. Please note, that the standard deviations from section 6.3 were converted to FWHM here.

Figure 7.1 (b) and (c) show the WEPL map of the pencil beam spot as well as the corresponding correction function Δ_WEPL. WEPLs were larger for negative values of ũ and decreased for positive values. The WEPL in the center of the beam spot was slightly negative. The dependency of the WEPL in the ṽ–direction was weaker. The correction function describes the spatial distribution of WEPL values well, with a good agreement

in particular inside of the beam spot. The fit parameters of the correction functions were determined as

$$
\begin{aligned}
a &= 0.956\,\text{mm}, \\
b_u &= 9.630 \times 10^{-2}, \\
b_v &= -0.884 \times 10^{-2}, \\
c_u &= -2.080 \times 10^{-3}\,\text{mm}^{-1}, \\
c_v &= -1.520 \times 10^{-3}\,\text{mm}^{-1}.
\end{aligned}
\tag{7.5}
$$

In fig. 7.1 (d) profiles through the WEPL map and the WEPL correction function are shown at the center of the beam spot and along the $\tilde{u}$–axis. Additionally, the corresponding profile through the number of counts is plotted in arbitrary units. The agreement between the correction function and the average WEPL was satisfactory and in particular good inside the FWHM margins, which are indicated as lines with the counts profile. The WEPL in the center of the beam spot was $-1.0\,\text{mm}$. The beam spot profile was slightly left–skewed: the mode of the distribution was at $\tilde{u} = 1\,\text{mm}$ while the expectation value was at $\tilde{u} = 0\,\text{mm}$.

Figure 7.1 (e) and (f) display two reconstructions of the water phantom at uniform fluence before and after application of the beam spot WEPL correction. The rings visible in (e) occur at a constant spacing, but are completely removed in by the correction in (f). Remaining artifacts in (f) are known and caused by the calibration [41]. They will be discussed in chapter 9. The average RSP value in the water region, where a RSP of 1.0 is expected, changed from 0.990 ± 0.024 without correction to 0.993 ± 0.023 with correction. All reconstructions shown in the subsequent investigations of this chapter used the correction function with the parameters determined above.

The acceptance probability to match the number of counts in experimental acquisitions and the corresponding plans using simulations was determined as

$$
r_{\text{accept}} = 25.62\,\%
\tag{7.6}
$$

by comparing an acquisition without phantom to the corresponding simulation. This value is reasonable considering that it was planned to acquire about four times the data necessary. All subsequent evaluations in this chapter only considered a random sample of measured protons with an acceptance probability r_{accept}.

7.2 EXPERIMENTAL REALIZATION OF FMPCT

MATERIALS & METHODS

To demonstrate the feasibility of using FMpCT imaging plans optimized based on simulations in experimental scans, the fluence patterns in section 6.5 for a constant noise prescription as well as FMpCT target A were used for scans of the water and the head

phantom. The image noise target A mimics a potential application of FMpCT to imaging in particle therapy, where the ROI follows two treatment beams at 90 and 180 degrees. As a reference, uniform fluence scans were also done for both phantoms. All acquisitions were done during the same beam time at the Northwestern Medicine Chicago proton center, except for the unit fluence acquisition of the water phantom, which was performed three days later (see table 3.2).

For the water phantom and the head phantom, RSP and image noise maps were reconstructed and compared to the corresponding plans based on simulations as well as to the prescriptions. All scans used the pencil beam spot WEPL correction described in section 7.1. A detailed analysis of the RSP accuracy using data acquired with the CTP phantom was also performed and is presented later in section 7.4.

RESULTS

Figure 7.2 displays the experimental fluence–modulated acquisitions for the water phantom in (a) to (c) and the corresponding results from the simulation in (d). The uniform fluence acquisitions in fig. 7.2 (a) shows an increased image noise close to the phantom's edge, as it was observed already in chapters 5 and 6. In fact, the variance profile across the entire diameter agrees well with the expectation. The counts sinogram in the last row shows the single pencil beam spots and also agrees with the results observed in section 6.5.

The scan applying a fluence pattern optimized for a constant variance target is shown in fig. 7.2 (b) and, just as in (a), the RSP map showed no strong artifacts. The image variance map of the scan, however, showed rings of increasing and decreasing image noise, which seemed to coincide with the RSP rings seen in fig. 7.1 and which had been corrected. The variance profile confirmed this fluctuation of the variance map. In a later analysis, these rings could be linked to the smaller pencil beam size combined with a slight misalignment of the fluence pattern, which will be discussed in section 7.6. Apart from the minor rings, a good agreement with the planned image variance and the prescription was observed. The fluence sinogram showed decreased fluence in the center of the phantom, as intended in the planning study.

The RSP maps of the FMpCT target, for which experimental and simulated planned scans are shown in fig. 7.2 (c) and (d), agreed well and a clear increase of image noise could already be observed in the RSP image. The increased variance outside of the ROI is confirmed in the image variance maps. Profiles along the image variance show a good agreement between experiments and planned scans, in particular inside the ROI. Both were slightly below the prescription, which was intended since the optimizer tried to achieve the prescription or less inside the ROI. Outside of the ROI, the plan had a noise level above the prescription while the experimental scan was mostly below the prescription. The fluence sinograms in the last row reached good agreement between the planning study and what was delivered experimentally.

Figure 7.3 shows the acquisitions using the head phantom, with the unit fluence acquisition in (a) and the FMpCT scan together with the simulated planned acquisition in (b) and (c). The agreement in terms of noise between the expected and the achieved variance was satisfactory for the unit fluence scan. Again, minor rings can be observed in the variance
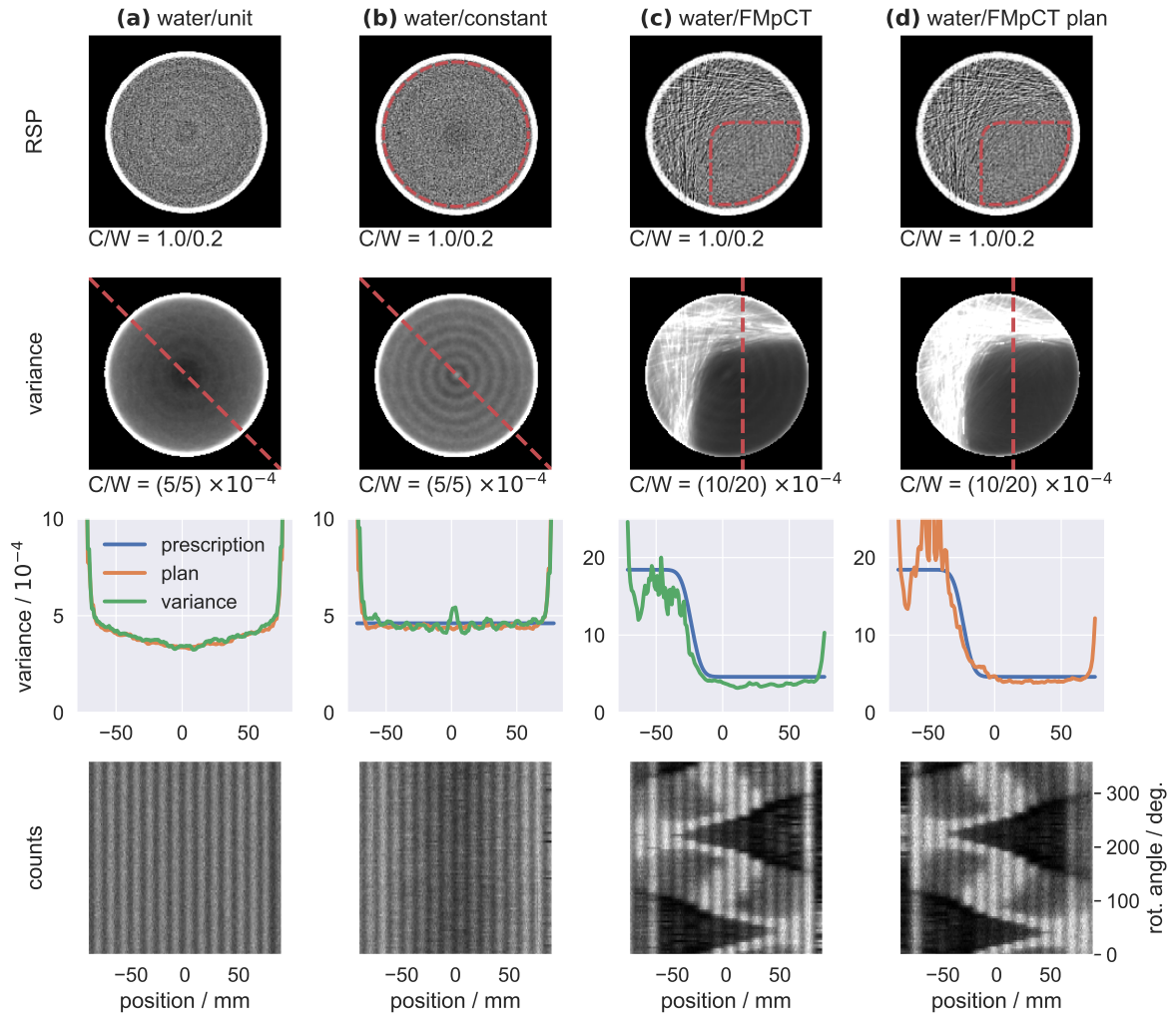
Figure 7.2: Experimental fluence–modulated acquisitions with the water phantom, (a) at unit fluence, (b) optimized for a constant noise target, (c) for the FMpCT noise target, and (d) the planned scan for the FMpCT noise target. The rows of each dataset show the RSP map, the image variance map, profiles through the image variance, the planned variance and the prescription, and counts sinograms at the central slice and central binning depth. The ROI of the noise prescription is displayed as a red dashed line in the RSP map. The profile shown in the third column is taken along the line indicated in the second column. The display settings for the counts sinogram were $C = 140$ protons, $W = 280$ protons, the other display settings are given below the plots.

map of the unit fluence acquisition, but they are small compared to the overall modulations seen across the phantom. The FMpCT scans also showed a good agreement in terms of variance between plan and experiment. Noise inside the ROI conformed well and was slightly below the prescription, as intended. Outside of the ROI, variance was elevated and agreed between experiment and plan.

❯ A discussion of these results can be found in section 7.6 on page 108.
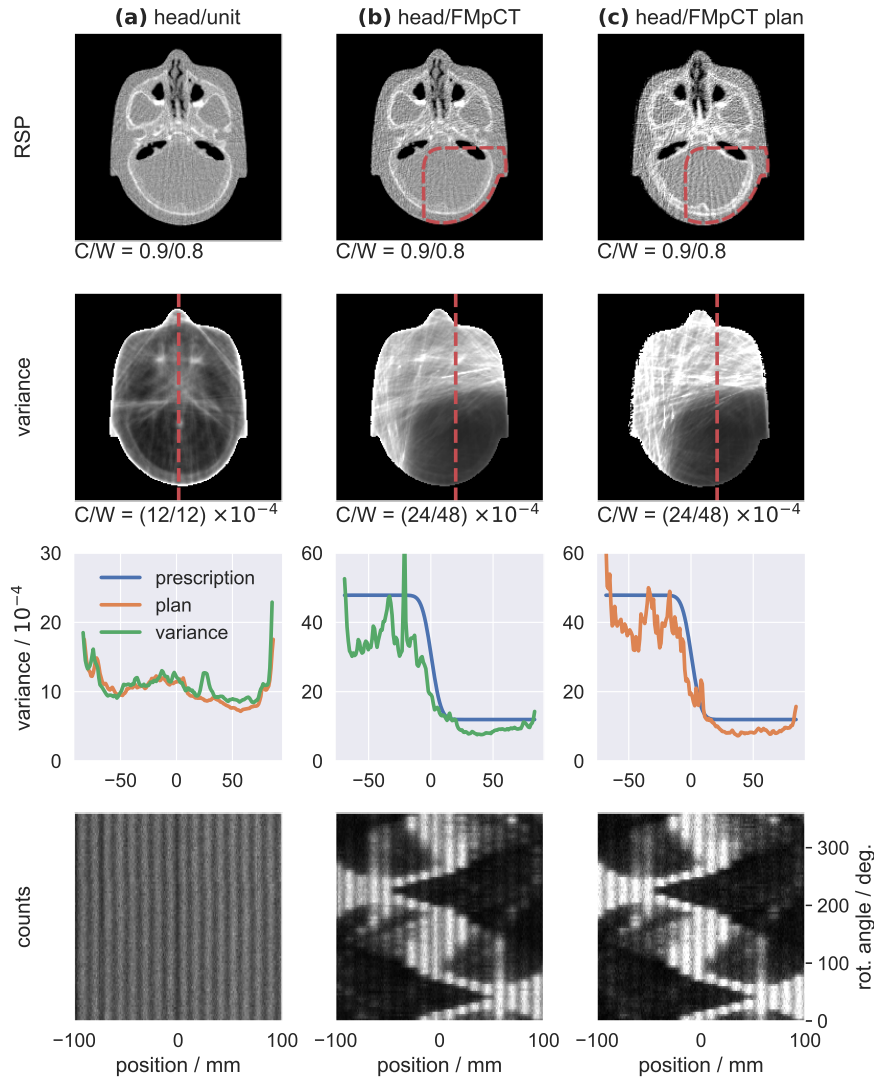
Figure 7.3: Experimental fluence–modulated acquisitions with the head phantom, (a) at unit fluence, (b) for the FMpCT noise target, and (c) the planned scan for the FMpCT noise target. The rows of each dataset show the RSP map, the image variance map, profiles through the image variance, the planned variance and the prescription, and counts sinograms at the central slice and central binning depth. The ROI of the noise prescription is displayed as a red dashed line in the RSP map. The profile shown in the third column is taken along the line indicated in the second column. The display settings for the counts sinogram were $C = 140$ protons, $W = 280$ protons, the other display settings are given below the plots.

7.3 PRECISION OF THE FLUENCE DELIVERY

MATERIALS & METHODS

The fluence delivery using pencil beams with a pattern shifted by a quarter interspace $\Delta_{\text{PB,u}}/4 = 3\,\text{mm}$ as described in section 6.3.2 relies on a precise delivery of this pattern in experimental scans. A change in spot size as observed in section 7.1 and even small shifts of

the pattern with respect to the rotation axis of the scanner can result in a distorted fluence delivery which manifests in distorted variance maps. In the worst case, if the pattern were misaligned by yet another quarter of the pencil beam interspace (and thus in total by half a pencil beam interspace), opposing beams would overlap (and so would opposing gaps in between beams), causing severe rings in the variance maps.

To further investigate the origin of the rings in image variance in figs. 7.2 and 7.3 "fluence sums" were calculated, which quantify the number of protons that contribute to the reconstruction of each pixel. Such maps can be calculated from the counts maps $C_n(u, v, d)$ for the n–th projection at rotation angle α_n, which were introduced in section 6.3. From this, the three–dimensional fluence sums $F(x, y, z)$ can be calculated by summing all projection counts maps after rotating them according to their respective rotation angle as

$$F(x, y, z) = \sum_{n=1}^{N} C_n(x \cos \alpha_n + y \sin \alpha_n, -x \sin \alpha_n + y \cos \alpha_n, z), \tag{7.7}$$

where N is the number of projections. To query $C_n(u, v, d)$, a linear interpolation was performed. Fluence sums were calculated both for experimental and simulated data.

To investigate the impact of the changed spot size, which was reported in eqs. (7.3) and (7.4) and to test the robustness of the fluence patterns to small shifts, a Monte Carlo simulation was employed using $FWHM_{u,exp}$ and $FWHM_{v,exp}$ in the beam model instead of the values determined in section 6.3. From this data, fluence sums were calculated using a modified version of eq. (7.7) that considers a small shift s_u of the pattern in the u–direction and that calculate as

$$F_{s_u}(x, y, z) = \sum_{n=1}^{N} C_n(x \cos \alpha_n + y \sin \alpha_n + s_u, -x \sin \alpha_n + y \cos \alpha_n, z), \tag{7.8}$$

where s_u was varied in five steps in between $s_u = -1.5 \, \text{mm}$ and $s_u = 1.5 \, \text{mm}$ including $s_u = 0 \, \text{mm}$, for which eq. (7.8) becomes equal to eq. (7.7). The calculation of shifted fluence sums was preferable since it did not require running Monte Carlo simulations for each setting.

RESULTS

In fig. 7.4 (a) and (b) the fluence sums for the experimental acquisitions and the corresponding plans are shown for all acquisitions. On the left of each column one slice of the fluence sum is shown within the xy–plane. On the right of each column a perpendicular view is shown along the white dashed line, which shows the z–coordinate along with a linear combination of x and y, which is denoted as x'. In general, summed fluences are at the same level and modulations seen in the plans are reproduced in the experimental scans. Discrepancies between the experiment and the plan occurred due to a smaller than planned pencil beam size, in particular in the v–direction, which corresponds to the z–coordinate here, and due to a shift of the fluence pattern in the u–direction with respect to the rotation axis of the scanner. The smaller spot size can be best observed in the perpendicular view, where there is a modulation in the z–direction for all experimental acquisitions where the
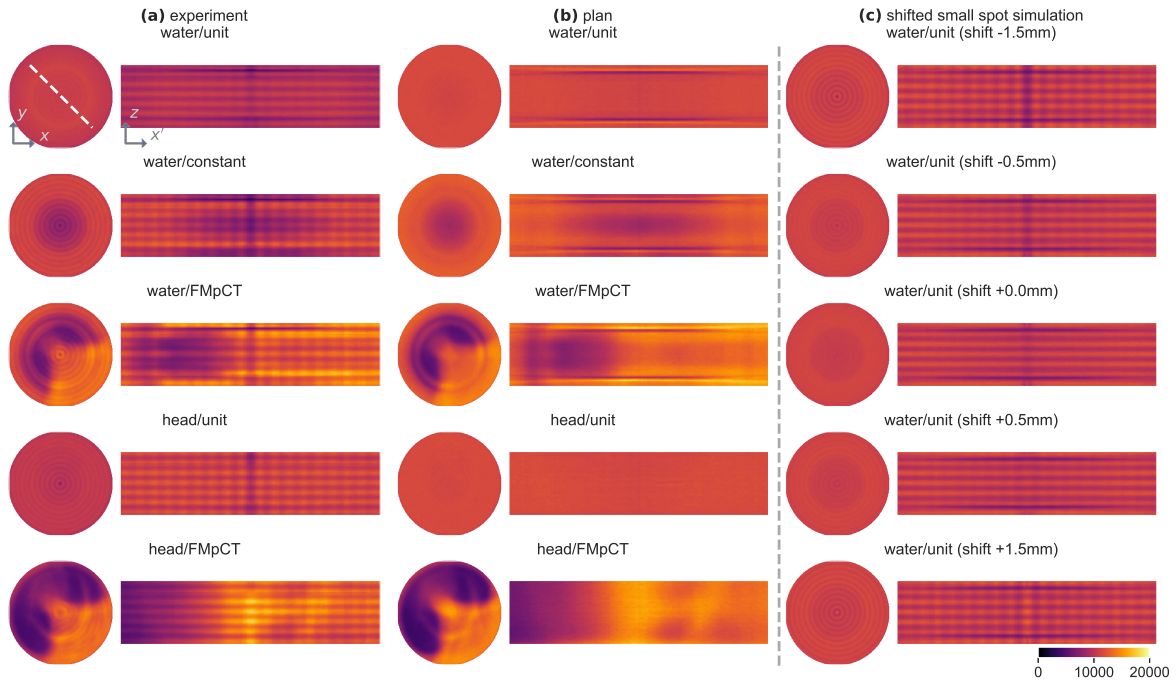
Figure 7.4: Three–dimensional fluence sums shown for one slice and as a perpendicular view along the indicated dashed line for (a) all experimental scans of the water and the head phantom and (b) for the corresponding planned scans using the simulation. Fluence sums from a Monte Carlo simulation replicating the smaller spot size and a variable shift of the fluence pattern are shown in (c). ⓘ An interactive version of this figure is available on the journal's website at https://stacks.iop.org/PMB/65/195001/mmedia. ⓒ Reproduced from Dickmann et al. [99] under the CC BY 3.0 license.

simulated plans show smooth fluence sums. A shift of the fluence pattern with respect to its intended location would cause a modulation in the x' direction and a modulation of the fluence sum in the xy–plane. The experimental unit fluence acquisition showed a smooth fluence sum in the x'–direction, which suggests that for this acquisitions the pattern was well aligned. All other scans, however, show modulations in the x'–direction as well as fluence rings within a slice of the fluence sum. The constant scans of the water and head phantom as well as the FMpCT scan of the water phantom exhibited a clear reduction of the fluence at the center close to the rotation axis, while the FMpCT acquisition of the head phantom showed an increase of the fluence, which suggests that the shift of the fluence pattern was in the opposite direction.

These observations were confirmed by the simulated fluence sums in fig. 7.4 (c), where the Monte Carlo simulation used the same spot size as in the experiment and an artificial shift s_u was added to the fluence sums. Indeed, the smaller spot size results in a modulation in the z–direction, which can be appreciated best for $s_u = 0\,\text{mm}$. A negative shift causes a modulation in the x'–direction, rings in the xy–plane, as well as a drop of fluence in close to the rotation axis. A positive shift, however, causes similar modulations, but an increase of the summed fluence close to the rotation axis. The extent of the modulations due to the shift seen in the experimental scan is comparable to the shifts of $s_u = \pm 0.5\,\text{mm}$.

❯ A discussion of these results can be found in section 7.6 on page 107.

7.4 RSP ACCURACY OF EXPERIMENTAL FMPCT SCANS

MATERIALS & METHODS

The RSP accuracy of a fluence–modulated scan compared to the corresponding unit fluence scan was evaluated using data acquired for the CTP phantom. Both datasets were reconstructed and the distribution of RSP values as well as the mean value was calculated for all inserts and the body. In the fluence–modulated acquisition a part of the body and the PMP and the Teflon insert were inside the ROI. The LDPE, polystyrene, acrylic and Delrin inserts were outside of the ROI and were imaged with an increased noise level. The diameter of the evaluation mask that was used to calculated distributions and mean values was chosen in agreement to the work of Dedes et al. [41] as a cylinder with a radius of 5 mm. The inserts inside the ROI had the lowest and highest RSP values and with that the largest expected RSP error. They were chosen such that this evaluation would serve as an upper limit estimate for the expected magnitude of RSP errors in FMpCT scans. The RSP error was quantified by the mean absolute percentage error (MAPE), which is defined as

$$\text{MAPE} = \frac{100\,\%}{M} \cdot \sum_{m=1}^{M} \frac{|\text{RSP}_{\text{mean},m} - \text{RSP}_{\text{ref.},m}|}{\text{RSP}_{\text{ref.},m}}, \tag{7.9}$$

and which averages the absolute error of all M inserts comparing their measured mean value $\text{RSP}_{\text{mean},m}$ to their ground truth value $\text{RSP}_{\text{ref.},m}$ as listed in table 3.1.

RESULTS

The RSP maps of the CTP phantom acquisition with unit fluence and the fluence modulation are shown in fig. 7.5 (a) and (b). As indicated by the sketch in fig. 7.5 (c) only the PMP and the Teflon insert were inside the ROI and were imaged with a low image variance. The noise level outside the ROI is visibly elevated in the FMpCT scan.

Figure 7.5 (d) shows histograms of RSP values in the inserts and the body of the CTP phantom both for the unit fluence and the FMpCT scan. The histograms for inserts outside of the ROI are shown on the left and inserts inside the ROI are on the right. The body was partially in and outside of the ROI. Inserts outside of the ROI experience a clear broadening of the RSP distribution in the FMpCT compared to the unit fluence scan. For the inserts in the ROI the width of the distribution stayed equal. The body only experienced a small broadening of the distribution. A clear change of the mean value could not be observed for any insert or the body.

The RSP accuracy is analyzed in detail in table 7.1 where the relative RSP error of all inserts and the body of the CTP phantom is reported for the unit fluence scan and for the fluence modulation. Results for both scans are reported both without and with the pencil beam WEPL correction resulting in four values per insert. In addition, these results are compared to the corresponding errors reported by Dedes et al. [41] in the last column of the table. Uncertainties are given as the standard error of the mean of all RSP values. In the last rows of table 7.1 the MAPE is calculated as the mean of all absolute relative values reported in the table (MAPE–ALL) as well as only for the two inserts in the ROI (MAPE–ROI).
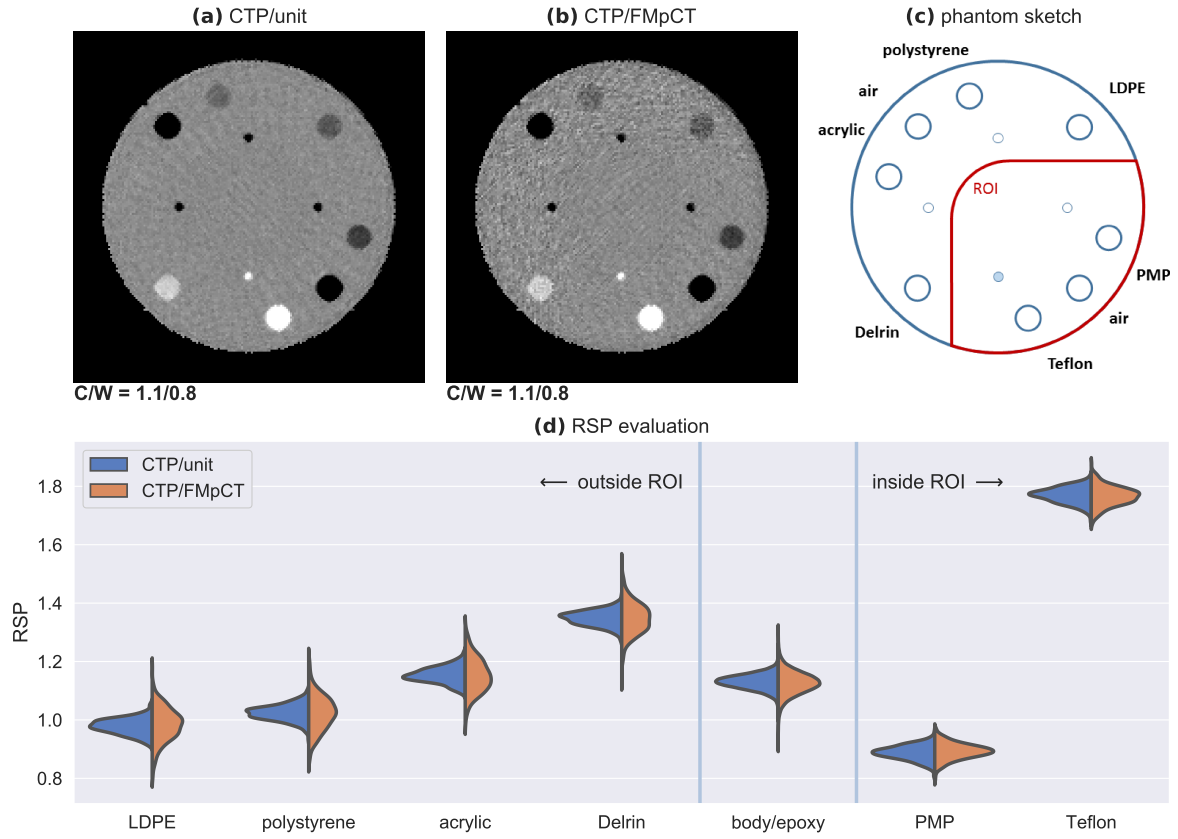
Figure 7.5: Evaluation of the accuracy of experimental FMpCT scans using the CTP phantom with (a) a unit fluence and a (b) FMpCT scan. The phantom's inserts are sketched in (c) and distributions of RSP values in the inserts and the body of the phantom are shown in (d) for both the unit fluence and the FMpCT scan. ㉢ Reproduced from Dickmann et al. [99] under the CC BY 3.0 license.

For the data acquired at unit fluence, all RSP errors improved when applying the pencil beam WEPL correction, except for the PMP insert. With that, also the MAPE–ALL improved, but the change was smaller than the associated uncertainty interval. For the FMpCT scan, the WEPL correction improved the MAPE–ALL and deteriorated the MAPE–ROI, but both changes were only a small fraction of the corresponding standard error. In particular, the PMP insert in the ROI and the LDPE insert outside of the ROI increased their relative error while all other inserts improved. For all inserts and both datasets, the RSP change caused by the correction was less than the associated standard error of the mean, which can also be seen in fig. 7.5 (d).

The fluence–modulated scan had a slightly increased MAPE–ROI compared to the corresponding value of the unit fluence scan. However, this change was again covered by the uncertainty interval. Values for the corrected FMpCT scan inside the ROI agreed well with the values reported by Dedes et al. [41]:

- for the PMP insert the relative error was 1.08 % and for the corrected FMpCT 1.06 %,

- for the Teflon insert the relative error was −1.31 % and for the corrected FMpCT −1.32 %.

Table 7.1: Evaluation of the RSP accuracy of unit fluence and fluence–modulated scans in experiments with the CTP phantom: relative RSP errors are reported for all inserts and the body for both scans and both with and without the pencil beam WEPL correction. In the last column corresponding results from Dedes et al. [41] are presented. The mean absolute percentage error is calculated for the two inserts inside the ROI (MAPE–ROI) and for all inserts (MAPE–ALL). Uncertainties are given as the standard error of the mean in each insert. *Abbreviations:* PMMA – polymethylmethacrylate/acrylic, LDPE – low–density polyethylene, PMP – polymethylpentene.

insert	RSP	uncorrected error in %		corrected error in %		error in %
		unit	FMpCT	unit	FMpCT	Dedes [41]
inside ROI						
PMP	0.883	0.18 ± 0.31	0.79 ± 0.36	0.51 ± 0.31	1.06 ± 0.35	1.08 ± 0.11
Teflon	1.790	-1.31 ± 0.18	-1.49 ± 0.21	-1.16 ± 0.17	-1.32 ± 0.21	-1.31 ± 0.05
outside ROI						
LDPE	0.979	-0.33 ± 0.32	0.24 ± 0.64	-0.12 ± 0.31	0.52 ± 0.65	-0.49 ± 0.11
polyst.	1.024	-0.12 ± 0.30	-0.25 ± 0.66	0.06 ± 0.29	-0.11 ± 0.67	-0.04 ± 0.10
body	1.144	-1.39 ± 0.02	-1.66 ± 0.03	-1.20 ± 0.02	-1.54 ± 0.03	—
PMMA	1.160	-0.80 ± 0.27	-0.80 ± 0.57	-0.54 ± 0.27	-0.63 ± 0.57	-0.30 ± 0.07
Delrin	1.359	-0.93 ± 0.21	-1.02 ± 0.45	-0.78 ± 0.21	-0.83 ± 0.45	-1.32 ± 0.21
MAPE–ALL		0.72 ± 0.09	0.89 ± 0.18	0.63 ± 0.09	0.86 ± 0.18	0.76 ± 0.05
MAPE–ROI		0.74 ± 0.18	1.14 ± 0.21	0.84 ± 0.18	1.19 ± 0.21	1.20 ± 0.06

Moreover, the MAPE–ROI agreed well between the corrected FMpCT scan and the previous study.

❯ A discussion of these results can be found in section 7.6 on page 108.

7.5 ACQUISITION TIME OF FMPCT SCANS

Data in this chapter were acquired in step–and–shoot mode which allowed enough time in between the acquisition of two projections for a manual initiation of the next modulated fluence pattern. An automated acquisition was not possible since there was no link between the scanner and the beam control system. For this reason, the acquisition of a full tomography took about 90 min, which would not be feasible for the acquisition of patient data. However, most of this time, the beam was off and the beam–on time was only 35.8 min for unit fluence tomographies and 20.3 min for the fluence–modulated acquisition of the water phantom. Considering that this study acquired four times the necessary data, the beam–on time required to produce one of the acquisitions shown in figs. 7.2 and 7.3 was just 9.2 min for the unit fluence scan and 5.2 min for the FMpCT scan. This is comparable to the typical scan time of pCT using this scanner [82]. If a synchronization were established between the scanner and the beam delivery system, the phantom could be rotated near–continuously allowing to reduce the acquisition time considerably. Moreover, the acquisition rate was

purposefully kept low in this study to avoid local pile–up at the tracker due the usage of pencil beams. Dedes et al. [45] suggested, that it may be possible to further increase the data acquisition rate by increasing the beam current without a loss of RSP accuracy, which in turn would also reduce the acquisition time to a few minutes. Eventually, acquisition rates in general need to be reduced for pCT scanners should the technology be used with patients. This problem is currently addressed by the development of new prototype scanners as discussed in section 3.2.4.

7.6 DISCUSSION

Delivering fluence–modulated scans at the proton center

In this chapter, fluence–modulated pCT data was acquired by using a grid of small proton pencil beams and modulating the fluence of each beam. Due to the extreme settings of divergence and momentum slits, which are required to keep the data acquisition rate viable for the scanner, these pencil beams are not circular as in clinical routine, but have an elliptical profile. This elliptical profile was modeled carefully in chapter 6 from previous experimental data acquired in the scope of this thesis. However, in the acquisition for this chapter, the pencil beam size was considerably smaller, in particular in the $\tilde{v}$–direction, where it was almost halved. The change in the spot size is likely to have occurred due to maintenance work and an upgrade, which was performed on the beam line in between the two acquisitions. In general, a smaller pencil beam spot size is favorable for FMpCT, since it would allow for a finer modulation of the fluence. However, since the smaller pencil beam size was not taken into account for the planning of the fluence patterns, this resulted in an unintended modulation of the fluence across slices. A detailed analysis of this modulation showed, that the modulation was small compared to the overall fluence modulation and a repetition of the experiments did not seem justified given the overall satisfactory results.

In addition to the smaller spot size, a spatial dependency of the mean energy within a single pencil beam was identified, which caused ring artifacts in the RSP maps of reconstructed images. While this is not expected to have a relevant impact on the corresponding image variance, it distorted RSP maps. A spatial dependency of the pencil beam is not expected for clinical operation and is likely to occur due to the closing of the momentum slits beyond settings commissioned for clinical use, which also degraded the beam profile. Based on data without a phantom, single pencil beams were isolated and the spatial energy dependency was fitted with an analytical function. This correction function could then be used in subsequent acquisitions to restore the expected quality of RSP maps. Previous studies using a static fluence modulation with pencil beams [46] also experienced a spatial dependency of the incident energy in a pencil beam, but noise masked this effect, as could be shown in a retrospective analysis of the data using de–noising shown in fig. 7.6, where a reconstruction of the water phantom from that dataset is shown together with line profiles and where the ring–shaped pencil beam artifacts can be seen. The application of the correction may have a minor impact on image variance due to a compression of the WEPL distribution in a pencil beam. However, it is important to note that the correction brings back the noise level to what
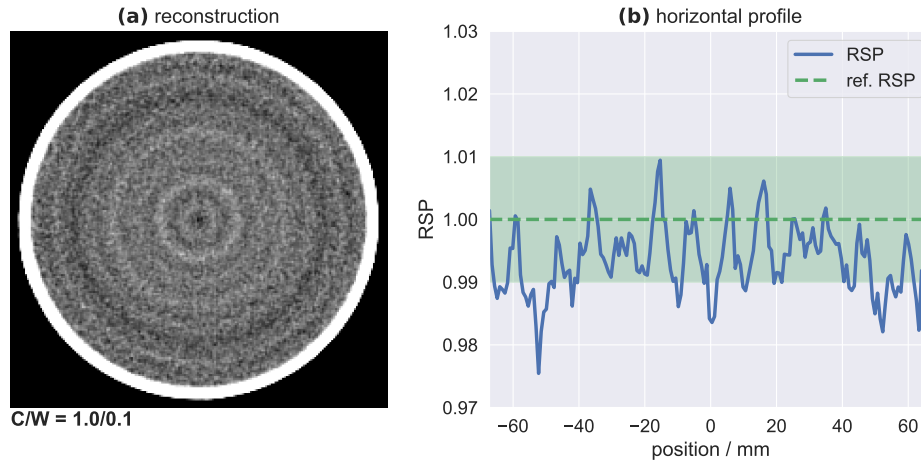
Figure 7.6: Evaluation of pencil beam artifacts present in the first experimental FMpCT study with a static fluence field: (a) the reconstruction of the water phantom with a Gaussian blur with $\sigma = 0.6\,\mathrm{mm}$ and (b) horizontal line profiles through the center of the phantom showing ring–shaped pencil beam artifacts.

was initially assumed during planning and is therefore mandatory for a better agreement with the simulation, where no energy shift was assumed.

Precision of the fluence delivery

Calculation of the number of protons intersecting a given voxel summed over all rotation angles, which was referred to as fluence sum, allowed to investigate the effect of the smaller pencil beam spot size as well as a misalignment of the fluence pattern with respect to the rotation axis. The pencil beams were mostly smaller in the $\tilde{v}$–direction, which caused a modulation of the fluence sums in the corresponding z direction (across the slices). This modulation is expected to be present also in the resulting variance maps.

In addition, misalignment of the fluence pattern with respect to the rotation axis, together with the quarter–shifted design of the grid with a larger interspace in the u–direction, caused rings of the fluence sums in the xy–plane (within one slice) for most of the scans. Only the water phantom at unit fluence did not suffer from this and probably had a better alignment of the pattern. This is plausible since the unit fluence data were acquired on a subsequent day (see table 3.2), for which the scanner was re–mounted and the alignment process needed to be repeated. By comparison of simulated data with an artificial shift, the shift present in the other phantoms was estimated to be in the order of $\pm 0.5\,\mathrm{mm}$. The rings in the fluence sums were also visible in the variance maps, in particular for the water phantom optimized for constant noise, but were much smaller than the modulations of variance achieved by FMpCT.

Given that fluence sums agreed well with the simulated counterparts, apart from the differences discussed above, the dose saving achieved in the experimental scans acquired for this study, is expected to be comparable to the values reported in chapter 6, which were up to 40.5 % outside of the ROI.

Experimental realization of FMpCT

Experimental unit fluence and FMpCT scans were reconstructed after correcting for an incident energy dependence of the pencil beams. Modulated scans were performed for three phantoms and for a constant noise target as well as an inhomogeneous image noise distribution with an imaging ROI. All scans showed a satisfactory agreement to the corresponding planning study. Distortions occurred only outside of the ROI, where noise is elevated and both the plan and the experiment did not yield the prescription. This was because the plan tried to achieve the variance target or less inside of the ROI at the cost of a potentially decreased noise level (and decreased dose saving) outside of the ROI. These effects were already discussed in section 6.6 and were found to be acceptable for FMpCT scans.

Due to the smaller pencil beam shape and a slight misalignment, rings were visible in the image noise maps, which were of a lesser extent compared to the overall modulation of image variance, but in particular distorted the constant noise target, for which no modulation of the image variance was intended. While this should be avoided in the future, it did not impair the good agreement between achieved and planned image noise distributions and the experimental realization of FMpCT scans was successful. In particular the use of a quarter–shifted pattern with a larger interspace in the u–direction should not be followed further, since it makes the correct fluence delivery more susceptible to even small movements. Instead, each projection should have pencil beams close enough to uniformly cover the field of view of the scanner. Small misalignments would then only shift the high and low noise regions and not distort the variance maps with rings.

RSP accuracy of experimental FMpCT scans

The RSP accuracy of PBS pCT scans in general and FMpCT scans in particular was investigated using the sensitometric CTP phantom. All relative RSP errors, with and without the pencil beam correction, were within the order of magnitude expected for pCT scans with broad beams, which is around 1 % [41]. With the pencil beam correction, a slight degradation compared to the unit fluence scan was observed for the two inserts that were within the ROI of the FMpCT scan. However, the deterioration was small and covered by the uncertainty intervals. Errors in the FMpCT scan also agreed well to a previous study using a broad uniform beam [41]. The two inserts inside the FMpCT ROI were also the inserts with the most extreme RSP values for which the largest errors were expected. The error estimates found in table 7.1, therefore, serve as an upper limit. In general, previous studies [41, 45] using idealized simulations and experiments showed that there is no degradation of the RSP accuracy using FMpCT, even when much stronger modulation factors were applied compared to this work. Since the errors found in this study were overall comparable to previous experimental acquisitions using the same scanner, no general deterioration of the RSP accuracy using FMpCT can be observed.

7.7 CONCLUSIONS

In this chapter, and originally in Dickmann et al. [99], fluence–modulated pCT scans were employed experimentally. The dynamic fluence patterns were optimized based on simulations with a detailed model of the phantoms and the pCT scanner. The fluence was modulated by employing a grid of small pencil beams and modulating the relative weight of each beam according to a previous optimization. The acquisition with pencil beams was impaired by a spatial dependency of the mean energy within each pencil beam. This distortion could be corrected by fitting the dependency and subtracting it in subsequent scans. In addition, the pencil beam shape was smaller compared to the values assumed in the plans and the pencil beam grid was slightly shifted. Both resulted in minor distortions of the fluence delivery and showed up as modulations and rings in the achieved variance maps. However, distortions were minor compared to the overall modulation of image variance and can be avoided in the future by not employing a quarter–shifted pencil beam pattern. The agreement between planned and experimentally achieved image variance was satisfactory and no relevant distortion of the RSP accuracy was found to be caused by using FMpCT. Hence, this is the first successful experimental implementation of FMpCT using optimized object– and task–specific fluence patterns.

8 | FLUENCE MODULATION WITH DOSE AND VARIANCE OBJECTIVES

F LUENCE–MODULATED PROTON CT tries to achieve an optimal trade–off between having sufficient image quality for treatment planning inside the ROI and reducing imaging dose elsewhere. In chapters 6 and 7, this trade–off was found via the image variance by prescribing low and high image variance levels inside and outside of the ROI. Imaging dose was treated only implicitly since an increased variance will also come with a reduced dose. In addition, this approach was projection–based and violations of the target in one projection could not be compensated by another projection.

In this chapter, a new optimization algorithm is described, that considers both dose and variance objectives in a joint cost function. With this, an image variance target can be achieved in the ROI while directly minimizing imaging dose outside. Stronger minimization penalties can be enforced in OARs, which allows to further spare critical structures. To test the performance of the new algorithm, FMpCT scans were optimized and simulated for three pediatric cases with tumors in the head region, for which the clinical delineations of OARs and the PTV were available. The ROI was set as the 10 % iso–dose volume of a proton treatment plan, which was optimized based on a ground truth RSP map using a research version of a commercial TPS. Compared to uniform fluence pCT scans, which were matched in terms of peak variance inside the ROI, the FMpCT scans allowed to reduce the imaging dose outside of the ROI by on average 74 % (from 1.2 mGy to 0.3 mGy). Doses to OARs were reduced compared to their local vicinity. The optimized proton plans were then re–calculated based on the pCT and FMpCT scans to evaluate their dosimetric accuracy. In a direct comparison, the passing rate of FMpCT scans for a 1 % criterion on the therapeutic dose difference was above 98.3 % (FMpCT/pCT) or above 88.5 % (FMpCT/ground truth) for all patients. For the corresponding proton ranges, the passing rate for a 1 mm criterion was better than 97.5 % (FMpCT/pCT) and better than 95.3 % (FMpCT/ground truth). This suggests that the FMpCT scans optimized using the proposed algorithm maintain the accuracy for proton dose calculation with only minor differences to uniform fluence pCT scans, but with a considerably reduced imaging dose in healthy tissue.

The results presented in this chapter were published in *Physics in Medicine & Biology* [138]. An early version of the algorithm was presented at the IEEE Medical Imaging Conference in Boston, USA [184].

8.1 SIMULATION AND IMAGE RECONSTRUCTION

In chapters 6 and 7, a grid of regularly spaced pencil beams was used to modulate the imaging fluence. The interspace of the grid was increased in the u–direction with the whole pattern shifted by a quarter interspace, allowing to reduce the number of pencil beams in the optimization. Two problems with this approach were identified in section 7.1: a change in the pencil beam spot size as well as even small shifts of the pattern with respect to the scanner can deteriorate the fluence delivery. In this chapter, the spot size was adapted to the values determined in chapter 7 and eqs. (7.3) and (7.4). A quarter–shifting of the pattern was avoided, which required to reduce the interspace in the u–direction. To simulate FMpCT scans, the pencil beam model in eq. (6.1) employed the standard deviations

$$\sigma_u = 3.7\,\text{mm}, \tag{8.1}$$

$$\sigma_v = 2.9\,\text{mm}, \tag{8.2}$$

and the grid consequently consisted of $N_{PB,u} = 51$ columns interspaced by $5.6\,\text{mm}$ and $N_{PB,v} = 21$ rows interspaced by $4.5\,\text{mm}$, which covered an area of $285.6\,\text{mm} \times 94.5\,\text{mm}$ with in total 1071 pencil beams per projection. The interspace between pencil beams was chosen as $1.5 \cdot \sigma$ to ensure that the summed fluence for equal pencil beam weights was homogeneous. The divergence parameters δ_u and δ_v were equal to the values reported in section 6.3.

Using this pencil beam grid, unit fluence pCT scans were simulated with $N_0 = 933$ protons per pencil beam, which resulted in 360 million primary protons per tomography with $N_P = 360$ projections, distributed uniformly across the scanner's aperture. This is an incident fluence of $37\,\text{mm}^{-2}$ ($26\,\text{mm}^{-2}$ after data cuts), which corresponds to an imaging dose of $1.2\,\text{mGy}$, which is comparable to the previous investigations and a standard pCT scan [82]. Unit fluence scans will be referred to as *pCT* and modulated scans as *FMpCT*.

In contrast to section 3.5.5, the slice thickness was increased in this study to $3\,\text{mm}$ to match the clinical x–ray CT scans that served as a patient model in the simulation study. This applied both to the z–direction of the reconstruction grid as well as to the v–direction of the distance–driven binned projections.

To facilitate the comparison of the algorithm proposed in this study and algorithms for treatment plan optimization, the notation changed slightly in this chapter. Variance projections at rotation angle α are labeled $v^\alpha(u, v, d)$ instead of $V^\alpha(u, v, d)$. Counts projections

are labeled as $f^\alpha(u, v, d)$ instead of $C^\alpha(u, v, d)$. In addition to those two quantities, the imaging dose to a given voxel (x, y, z) for a single projection with rotation angle α is labeled $d^\alpha(x, y, z)$.

8.2 PATIENT DATA & TREATMENT PLANNING

For this work, pCT and FMpCT scans were simulated based on the geometry of three pediatric patients that underwent (photon) radiotherapy of the head. For each patient, proton therapy plans were optimized based on the prescriptions and delineations of the corresponding photon plans. Pediatric tumors of the head were chosen due to the limited field–of–view and WEPL range of the prototype pCT scanner. Patients were also selected such that the full PTV could be covered with a single pCT scan. Pediatric cases in particular are most likely to benefit from a reduction of the imaging dose since they are more susceptible to radiation induced damage due to their age. Table 8.1 lists clinical parameters, tumor sites and dose prescriptions for the three patients used in this retrospective study. The tumor sites cover typical cases and different tumor locations within the head were considered. All plans employed 6 MV photon beams. For patient 1 the prescription included a boost to the primary tumor, which is why two dose levels are reported. This was realized as an integrated boost for the proton plan.

To generate ground truth RSP maps for each patient, the x–ray CT scans used for radio-therapy treatment planning were imported to the GEANT4 code as described in section 4.3 and exported as a voxelized map of RSP values for a reference proton energy of 150 MeV. These maps served as a ground truth to compare pCT and FMpCT scans to. The ground

Table 8.1: Clinical parameters of the photon radiotherapy for the three patients in the simulation study. For patient 1, a sequential boost to the primary tumor was prescribed, which is why two dose levels are reported. The last two rows list gantry angles used for the proton therapy plans and the angle used for the range evaluation.

	patient 1	patient 2	patient 3
photon plans			
age at first radiotherapy	4.0 years	5.8 years	4.4 years
tumor site	left orbit	pons	parotid gland
tumor type	embryonal rhabdomyosarcoma	glioma	acute lymphatic leukemia
prescription/Gy	50.4* / 36.0	54.0	20.0
fractions	28	30	8
fraction dose/Gy	1.8	1.8	2.5
radiotherapy intent	curative	curative	palliative
proton plans			
gantry angles/deg	30, 90	90, 180, 270	270, 315
gantry angle (range)/deg	90	180	270

* boost to primary tumor

truth RSP map was then converted back to CT values using the piece–wise linear function of section 4.3 to be able to import them to the TPS RayStation (*RaySearch Laboratories AB*, Stockholm, Sweden). The conversion from CT values to material compositions and RSP and back to CT values resulted in small differences between the original treatment planning CT and the imported ground truth CT. However, this allowed to subsequently import pCT and FMpCT scans to the TPS (using the same piece–wise linear function) and probe their accuracy for proton dose calculations.
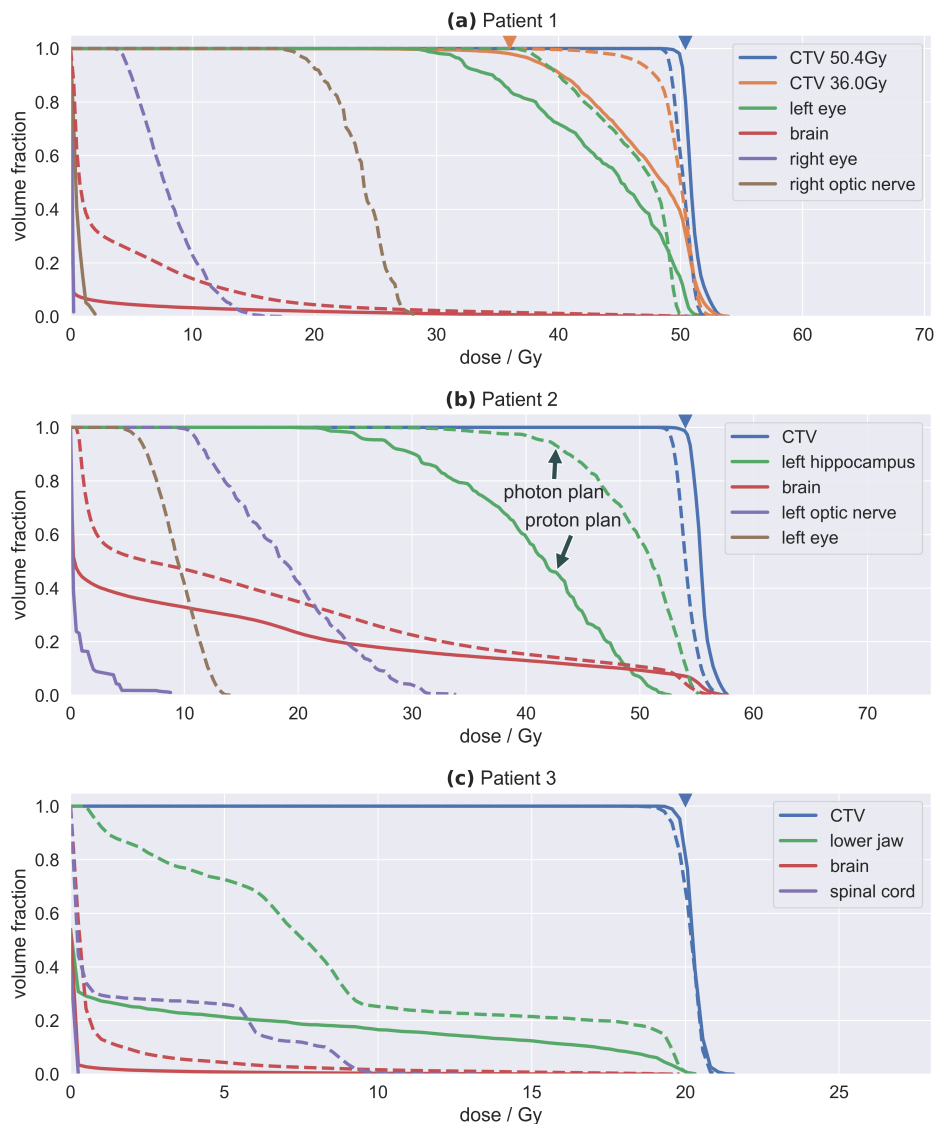
Figure 8.1: Dose–volume histograms for the photon and proton treatment plans for the three patients used in this study. The prescription dose is indicated by a small triangle. Patient 1 had a boost to the primary tumor and thus two prescriptions for the clinical target volume (CTV).

Based on the ground truth RSP maps, that were imported as CT values to the TPS, proton treatment plans were optimized, following the delineations and dose prescriptions of the original photon plans. For the proton plans, the CTV was used as the prescription volume and two or three gantry angles as listed in table 8.1 were employed. In the TPS, a generic treatment machine (RSL_IBA_DED) was used with an air gap of 100 mm and a PMMA range shifter with a water–equivalent thickness of 40 mm. The grid size for dose calculations was uniformly 3 mm. Beam spots were interspaced laterally by 3 mm and the adaptive energy spacing of the TPS was used. For proton therapy, typically the CTV instead of the PTV is used as the prescription volume for treatment planning and uncertainties of the proton delivery are considered using a robust optimization. For this, a positional uncertainty of 3 mm and a range uncertainty of 3 % was used. Figure 8.1 shows dose–volume histograms for all three patients and for both the delivered photon plans as well as the proton plans generated for this study. For all OARs as well as the CTV a similar or better distribution was achieved using protons with respect to the photon plans.

8.3 AN ALGORITHM FOR JOINT DOSE MINIMIZATION AND VARIANCE OPTIMIZATION

MATERIALS & METHODS

In this chapter, a novel optimization algorithm for FMpCT is proposed, that takes into account both imaging dose and image variance objectives. It calculates pencil beam weights that achieve a given variance target in the ROI while minimizing imaging dose outside. Non–zero imaging doses can be penalized with a spatially varying strength, allowing to further reduce imaging dose in so–called imaging OARs. A workflow of the algorithm summarizing all relevant quantities is shown in fig. 8.2. Details of the optimization will be given in the following sections.

A first part of the optimization, depicted in fig. 8.2 (a) is based on an initial Monte Carlo simulation at unit fluence. It assumes that the imaging fluence can be modulated within small bixels. A bixel is a virtual detector element, which corresponds to a ray from the detector and through the image volume. A modulation of the fluence within that bixel would then only affect image voxels crossed by that ray. The parallel bixel rays are smaller than pencil beams and do not overlap.

Using the concept of bixels, a forward model is created, that predicts image variance and imaging dose distribution for a given set of bixel weights. With this, a bixel–wise fluence optimization is possible, which uses concepts from early intensity–modulated radiotherapy [185–187]. In a second step, the bixel–wise fluence modulation is translated to pencil beam weights using the Gaussian pencil beam model of section 6.3 and a second optimization depicted in fig. 8.2 (b). All equations in this section are formulated in parallel–beam geometry and are used as such without generalization to a cone–beam geometry. This makes the computation faster and is a fair assumption, given that pencil beams originate from a virtual source (the position of the scanning magnets), which is approximately 1.8 m from the isocenter [156].
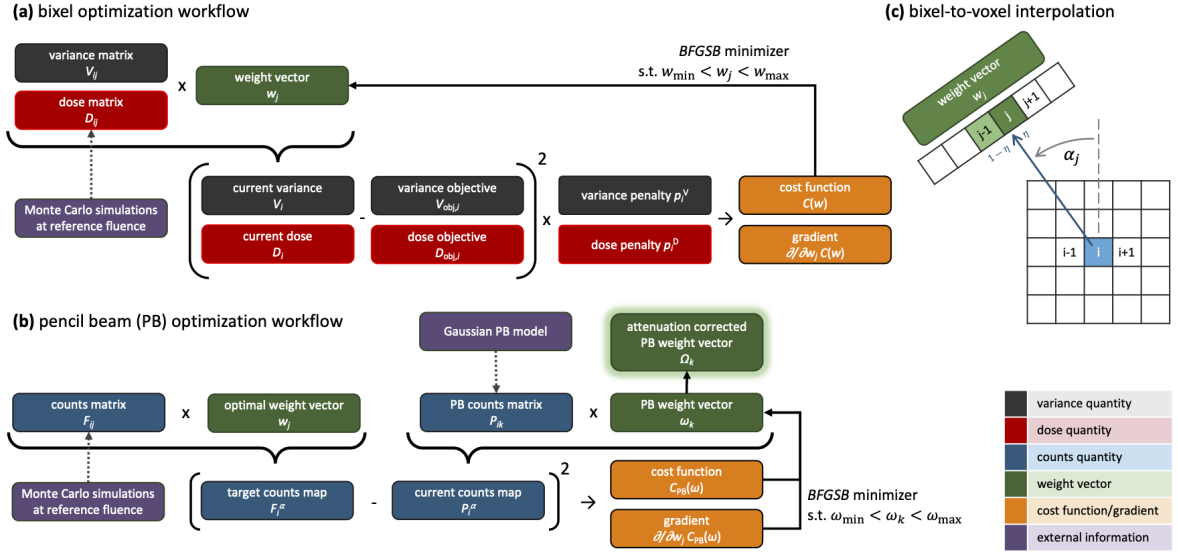
Figure 8.2: Optimization workflow for the proposed method with an overview of all required quantities: (a) the bixel–wise optimization, (b) the pencil beam optimization, and (c) the relationship between bixels and voxels. The color coding of all quantities is given in the legend. © Reproduced from Dickmann et al. [138] under the CC BY 4.0 license.

The dose and variance forward models employ a bixel grid as illustrated in fig. 8.2 (c) with each bixel corresponding to a virtual detector element of the size $4\,\mathrm{mm} \times 4\,\mathrm{mm}$. All bixels of all projections are numbered consecutively with the bixel index $j \in \{1, \ldots, M\}$. This amounts to $M = N_P \cdot N_{b,u} \cdot N_{b,v} = 518\,400$ bixels, where $N_P = 360$ is the number of projections and $N_{b,u} = 60$ and $N_{b,v} = 24$ are the number of bixels in u– and v–direction respectively. Each bixel is associated to a rotation angle α_j, where $N_{b,u} \cdot N_{b,v} = 1440$ pencil beams belong to one projection and share the same projection angle. The voxels of the image reconstruction volume are also numbered consecutively with the voxel index $i \in \{1, \ldots, N\}$ with $N = N_x \cdot N_y \cdot N_z = 86\,400$. Each bixel's center is denoted as (u_j, v_j) and each voxel's center is denoted as (x_i, y_i, z_i).

8.3.1 Forward models for dose & variance

IMAGING DOSE The first forward model used in the optimization calculates the imaging dose D_i to voxel i. It can be formulated as a matrix multiplication of the dose matrix D_{ij} with a bixel weight vector w_j as

$$D_i = \sum_{j=1}^{M} D_{ij} \cdot w_j, \tag{8.3}$$

where the weight vector w_j describes the relative fluence modulation of bixel j compared to a reference fluence. A similar approach has been proposed by Scholz, Nill, and Oelfke [8] for treatment planning. Equation (8.3) could have been written as $D = \mathbf{D} \cdot w$, but the explicit notation of the sum is chosen in the following for clarity. The dose matrix D_{ij} can be calculated from the dose $d_i^{\alpha_j}$ scored in a unit fluence Monte Carlo simulation for rotation

angle α_j and voxel i. To achieve this, the dose of the full projection is sliced into contribution of individual bixels and the dose matrix becomes

$$D_{ij} = d_i^{\alpha_j} \cdot \delta_{ij}, \tag{8.4}$$

where δ_{ij} defines the relationship between bixels of the fluence modulation and voxels in the volume. In the simplest case, δ_{ij} could be one if bixel j has an intersection with voxel i and zero otherwise. For this, the non–zero elements of D_{ij} for a given voxel i would be the unit fluence doses (in Gray) of all bixels j intersecting the voxel. For a more accurate forward calculation, an interpolation between the bixel and voxel grids is taken into account, and thus

$$\delta_{ij} = \begin{cases} \eta_{ij} & \text{if } \lfloor (x_i \cos \alpha_j - y_i \sin \alpha_j - u_j)/u_j \rfloor = 0 \text{ and } z_i = v_j \\ 1 - \eta_{ij} & \text{if } \lceil (x_i \cos \alpha_j - y_i \sin \alpha_j - u_j)/u_j \rceil = 0 \text{ and } z_i = v_j \\ 0 & \text{else} \end{cases}, \tag{8.5}$$

where $\lfloor \cdot \rfloor$ is the floor operator, $\lceil \cdot \rceil$ is the ceil operator. The interpolation fraction η_{ij} calculates as

$$\eta_{ij} = (x_i \cos \alpha_j - y_i \sin \alpha_j - u_j)/u_j - \lfloor (x_i \cos \alpha_j - y_i \sin \alpha_j - u_j)/u_j \rfloor . \tag{8.6}$$

This interpolation is is illustrated in fig. 8.2 (c) with two bixels corresponding to one voxel. The matrix D_{ij} is sparse and has only two non–zero elements for every row or column.

IMAGE VARIANCE Similarly a forward model for image variance can be established that calculates the image variance V_i in voxel i for a given bixel–wise fluence modulation w_j as

$$V_i = c \cdot \sum_{j=1}^{M} V_{ij} \cdot \tilde{w}_j, \tag{8.7}$$

where c is an additional constant needed for variance reconstruction, which will be defined later, V_{ij} is the variance matrix and $\tilde{w}_j = 1/w_j$ are the inverse of the fluence weights, which are required since variance is inversely proportional to the fluence. The variance matrix V_{ij} can be calculated from a unit fluence dataset, from which distance–driven binned variance projections in coordinates (u, v, d) can be calculated and which can be transformed to the corresponding coordinates (x_i, y_i, y_i) by a rotation with the corresponding rotation angle, which also requires an interpolation. The rotated variance projections will be called $v_i^{\alpha_j}$ and with them the variance matrix is defined as

$$V_{ij} = v_i^{\alpha_j} \cdot \delta_{ij}, \tag{8.8}$$

where δ_{ij} is defined in eq. (8.5). The additional constant c can be calculated by comparing eqs. (3.63) and (8.7). To make both equations equal, the constant needs to be

$$c = f_{\text{interp}} \cdot f_{\text{filter}} \cdot \left(\frac{\pi \Delta u}{N_P} \right)^2, \tag{8.9}$$

where the two factors are defined in eqs. (3.60) and (3.64). With this, eq. (8.7) becomes equivalent to a variance reconstruction of the modulated variances $v_i^{\alpha_j} \cdot \tilde{w}_j$ according to the simplified variance reconstruction formula in eq. (3.63), which ignores the pixel–by–pixel effect of interpolation and also avoids the convolution with the filter kernel.

8.3.2 Bixel–wise optimization

With eqs. (8.3) and (8.7) it is possible to calculate imaging dose and image variance estimates for any arbitrary fluence modulation w_j from which the corresponding $\tilde{w}_j$ can be calculated. In the optimization the inverse problem of finding a suitable set of bixel weights w_j was solved such that a given imaging dose objective $D_{obj,i}$ and a variance objective $V_{obj,i}$ are achieved in voxel i. The competing targets were assigned a dose penalty p_i^D and a variance penalty p_i^V, which describe the relative strength with which a violation of the objective is to be avoided. The objectives and penalties will be defined later in section 8.3.4 and all quantities are illustrated in fig. 8.2 (a).

The bixel weights w_j were optimized by minimizing the cost function

$$C(w) = \sum_{i=1}^{N} p_i^D (D_i(w) - D_{obj,i})^2 + \sum_{i=1}^{N} p_i^V (V_i(w) - V_{obj,i})^2, \tag{8.10}$$

where D_i and V_i depend on the weights w_j. The gradient with respect to a single weight w_j then calculates as

$$\frac{\partial}{\partial w_j} C(w) = 2 \sum_{i=1}^{N} p_i^D (D_i(w) - D_{obj,i}) D_{ij} - 2 \sum_{i=1}^{N} p_i^V (V_i(w) - V_{obj,i}) \frac{V_{ij}}{w_j^2}. \tag{8.11}$$

With this, the optimized weights $\hat{w}_j$ can be found as

$$\hat{w} = \arg\min_{w} C(w) \text{ s.t. } 0 < w_{min} \leqslant w_j \leqslant w_{max} \, \forall \, j, \tag{8.12}$$

where the weights are forced to stay within the limits w_{min} and w_{max} to ensure that only positive and finite weights are employed. Negative weights would correspond to negative fluences which are unphysical and too extreme modulations are avoided to prevent unreasonably high local imaging doses. The optimization was performed with the limited–memory BFGSB algorithm by Zhu et al. [188] using the implementation of the *Insight Toolkit (ITK)* [123]. Sparse matrix multiplication needed for the evaluation of the cost function were efficiently implemented in C++ using the *Eigen3* library [189].

8.3.3 Pencil beam optimization

While with the optimized $\hat{w}_j$ a fluence modulation is already found, this cannot be used experimentally, since fluence can only be modulated in small pencil beams and not with bixels. To solve this, the bixel weights are first converted to a target counts map, which is subsequently achieved with modulated pencil beam weights as depicted in fig. 8.2 (b). The counts map for projection angle α, which results from a given fluence modulation w_j, can be calculated as

$$F_i^\alpha = \sum_{j \in \{j_\alpha\}} F_{ij} \cdot w_j, \tag{8.13}$$

where $\{j_\alpha\}$ is the set containing all j for which $\alpha_j = \alpha$. The three–dimensional counts map $f_i^{\alpha_j}$ describes the number of protons at unit fluence from projection angle α_j contributing to voxel i and with it the fluence matrix F_{ij} can be defined as

$$F_{ij} = f_i^{\alpha_j} \cdot \delta_{ij}, \tag{8.14}$$

where δ_{ij} is defined in eq. (8.5). With this, a target projection counts map $\hat{F}_i^\alpha$ can be calculated that corresponds to the optimized weights $\hat{w}_j$. The counts map $f_i^{\alpha_j}$, and consequently F_{ij}, is subject to attenuation and only describes the number of protons contributing to image reconstruction. This will be considered later.

To calculate a fluence modulation that takes into account the shape of the $K = N_P \cdot N_{PB,u} \cdot N_{PB,v} = 385\,560$ pencil beams, a second optimization is required, that makes use of the pencil beam model described in eq. (6.1). Using the model, the counts map P_{ik} of pencil beam k in voxel i can be calculated, where the normalization was such that the maximum value was forced to be one. Each pencil beam was associated with a corresponding rotation angle α_k and a weight ω_k. With that, a modulated counts projection for rotation angle α could be calculated as

$$P_i^\alpha = \sum_{k \in \{k_\alpha\}} P_{ik} \cdot \omega_k, \tag{8.15}$$

where $\{k_\alpha\}$ is the set containing all k for which $\alpha_k = \alpha$. The Greek notation of the pencil beam weights ω_k is intended to differentiate them from bixel weights w_j. P_{ik} was calculated by querying the Gaussian pencil beam model on a fine grid of $1\,\text{mm} \times 1\,\text{mm} \times 1\,\text{mm}$, rotating the volume by α_k and subsequently resampling it to the coarse grid with voxels i. Due to this, the interpolation with δ_{ij} as in eqs. (8.3), (8.7) and (8.13) was omissible. Unnecessary calls to the Gaussian function were avoided by setting P_{ik} to zero outside of three standard deviations of the pencil beam.

Pencil beam weights could then be optimized using the cost function

$$C_{PB}(\omega) = \sum_{i=1}^N (P_i^\alpha(\omega) - F_i^\alpha)^2 \tag{8.16}$$

and the corresponding gradient

$$\frac{\partial}{\partial \omega_k} C_{PB}(\omega) = 2 \sum_{i=1}^N (P_i^\alpha(\omega) - F_i^\alpha) P_{ik}, \tag{8.17}$$

resulting in pencil beam weights ω_k of

$$\hat{\omega} = \arg \min_\omega C_{PB}(\omega) \ \text{ s.t. } \ 0 < \omega_{min} \leqslant \omega_k \leqslant \omega_{max} \ \forall\, k \in \{k_\alpha\}, \tag{8.18}$$

and using the bounds ω_{min} and ω_{max} which were equal to the bixel bounds w_{min} and w_{max}. Weights for each rotation angle α are independent and their optimization can be parallelized.

Similar to eq. (6.13) in section 6.4, the optimization of pencil beam weights was subject to attenuation due to the usage of $f_i^{\alpha_j}$ in eq. (8.13). For this reason the optimization in eq. (8.18) must be performed twice. First, for the modulated target projection counts map $\hat{F}_i^\alpha$ and then for the reference counts map $\dot{F}_i^\alpha$. The latter can be calculated using eq. (8.13) with all weights $w_j = 1$. This will result in two optimized sets of pencil beam weights $\hat{\omega}$ and $\dot{\omega}$, which are subject to attenuation. The attenuation–corrected pencil beam weights Ω_k can then be found as the ratio

$$\Omega_k = \frac{\hat{\omega}_k}{\dot{\omega}_k}, \tag{8.19}$$

for which the effect of attenuation cancels out. Ω_k is again forced to be in the interval between ω_{min} and ω_{max} and thresholded otherwise. Fluence–modulated simulations can then be performed using $\Omega_k \cdot N_0$ instead of N_0 particles for pencil beam k.

8.3.4 Optimization objectives

The ROI for FMpCT imaging was defined based on the ground truth treatment dose distribution of each patient. All voxels receiving at least 10 % of the prescription dose were assumed to belong to the ROI. This threshold is in agreement to the recommendation of the AAPM task group 119 for comparing treatment doses using gamma analysis [190, 191]. The resulting ROI volumes are listed in table 8.2.

Since imaging dose was to be minimized, the imaging dose objective was set to $D_{obj,i} = D_{obj} = 0 \, \text{mGy}$ throughout the volume. Also the variance was set to a constant value $V_{obj,i} = V_{obj}$, which was defined as the 95th variance percentile of the corresponding unit fluence scan inside the ROI. It was different for each patient and values are listed in table 8.2.

The dose and variance penalties p_i^D and p_i^V were both set to zero outside of the patient's skin, since both quantities are irrelevant there. The variance penalty p_i^V was also set to zero outside of the ROI, and to $p_i^V = 100$ inside. The dose penalty p_i^D instead was set to $p_i^D = 0.01$ inside the ROI and to $p_i^D = 1$ outside. Furthermore, each patient had between two and four imaging OARs as listed in table 8.2 for which the dose penalty was increased to $p_i^D = 20$. These choices were made empirically to ensure that $V_{obj,i}$ was achieved in the ROI and was not impaired by the dose minimization and that pencil beams intersecting only with the ROI would not result in unreasonably high doses.

The minimum weights were $w_{min} = \omega_{min} = 0.05$, which made sure that at least four protons per pixel were available for image reconstruction for an incident unit fluence of $26 \, \text{mm}^{-2}$ and a pixel area of $3 \, \text{mm}^2$. The maximum weights were $w_{max} = \omega_{max} = 10$, which again avoided unreasonably high doses in the ROI.

Table 8.2: ROI volumes and objectives for fluence–modulation and imaging OARs for all patients. For the variance objective, the corresponding RSP standard deviation σ_{obj} is also given. *Abbreviations:* ROI – region–of–interest, OAR – organ–at–risk, r – right, l – left.

	patient 1	patient 2	patient 3
ROI volume / cm^3	223	785	321
dose objective D_{obj} / mGy	0.0	0.0	0.0
variance objective V_{obj}	5.41×10^{-4}	6.09×10^{-4}	6.72×10^{-4}
σ_{obj}	0.023	0.025	0.026
imaging OARs	r. eye, brainstem	r./l. eye, r./l. optical nerve	r./l. eye, spinal cord

8.3.5 Reference approach

To evaluate the performance of the proposed algorithm, results were compared to a simple intersection–based approach for fluence modulation [46], which was also used in chapter 6 as a reference. Within the notation used in this chapter, this binary modulation can be expressed using a target volume T_i, which was zero if p_i^V was zero and one everywhere else (i. e. in the ROI). Then, the binary weights can be calculated as a scalar product

$$\Omega_k = \begin{cases} 1 & \text{for } \sum_{i=1}^{N} P_{ik} T_i > 0 \\ \omega_{\min} & \text{else} \end{cases} . \tag{8.20}$$

Since the Gaussian pencil beam model was zero outside a three standard deviation interval, this will result in the full fluence, if a pencil beam intersects with the ROI within that interval, and a minimum fluence equal to $\omega_{\min}$, if the pencil beam does not intersect. Such scans will be labeled *binary* in the following.

RESULTS

In fig. 8.3 (a) – (c) the cost function values of the bixel–wise fluence optimization are shown for all three patients. The cost function is split into contributions from the dose term
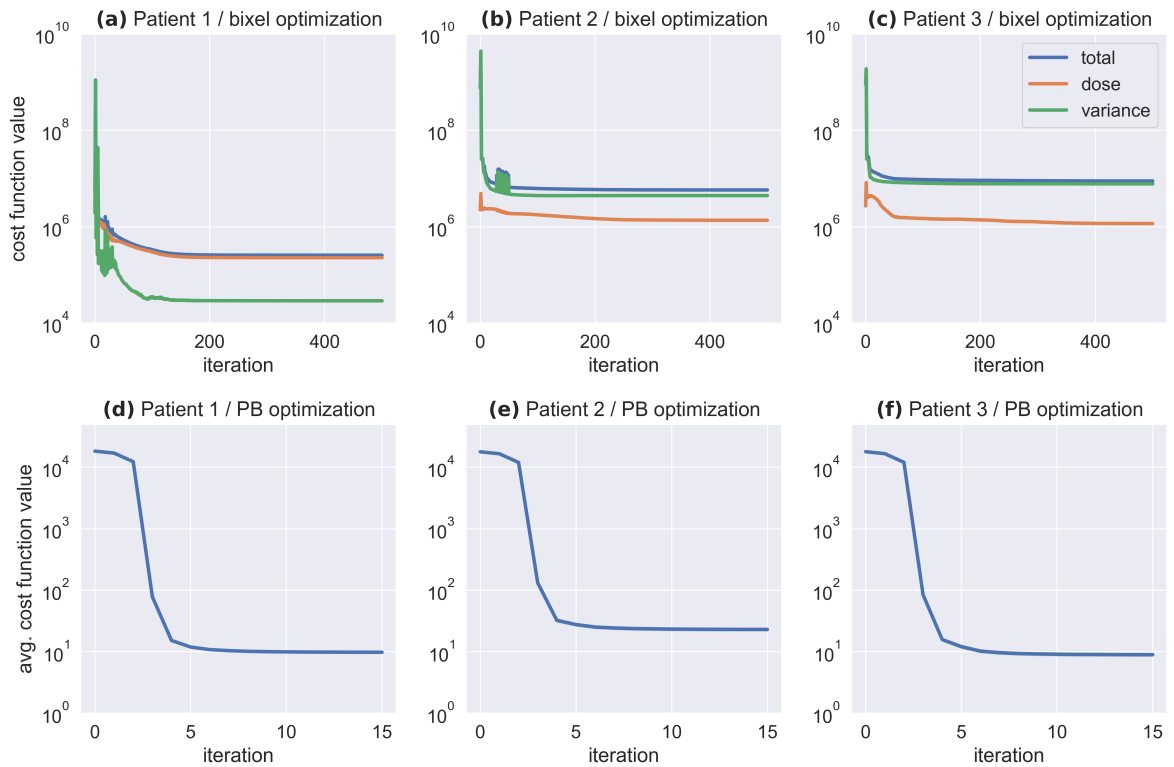
Figure 8.3: Cost function values for all patients and for each iteration of (a) – (c) the bixel–wise fluence optimization and (d) – (f) the pencil beam (PB) optimization as an average over all projections.

(first summand in eq. (8.10)) and the variance term (second summand). The optimization ran for 500 iterations, but reached a plateau already after 100 to 200 iterations. For patient 1, the dose term is the largest contribution to the cost function while for patient 2 and patient 3 the variance term dominates. Despite differences in magnitude between the two contributions, both reduce considerably within the first 200 iterations.

Figure 8.3 (d) – (f) shows cost function values for the subsequent pencil beam optimization, which are displayed as an average over all 360 projection angles. The convergence is much quicker and a convergence plateau is reached within less than 10 iterations.

Fluence modulations were optimized on a computer with two Intel Xeon E5–2667 v4 processors (*Intel Corporation*, Santa Clara, CA, USA) with a clock frequency of 3.2 GHz and with in total 16 physical cores and 252 GB of memory, which was required to store the dose
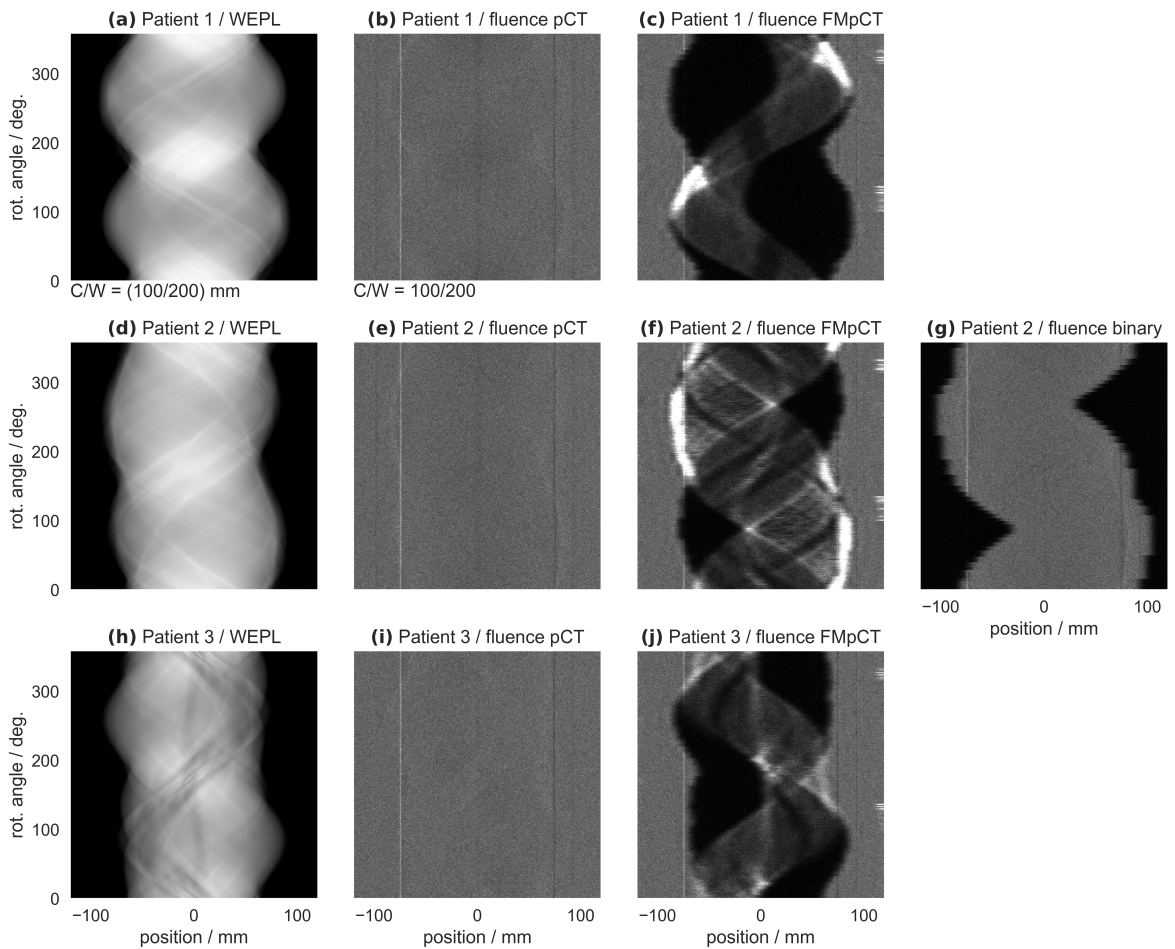
Figure 8.4: Sinograms for the three patients: (a,d,f) WEPL sinograms, (b,e,i) uniform fluence sinograms, (c,f,j) fluence–modulated fluence sinograms, and (g) a fluence sinogram for a simple intersection–based binary modulation. Sinograms are shown at a v–coordinate at the center of the ROI and at the central binning depth. The display settings are $C = 100\,\text{mm}$, $W = 200\,\text{mm}$ for WEPL sinograms and $C = 100$ protons, $W = 200$ protons for fluence sinograms.

and variance matrices. For the bixel–wise optimization, a single evaluation of the optimizer's cost function took 54 ms and one iteration on average 302 ms. In total, the bixel–wise optimization for patient 1 over all 500 iterations took 151 s. The pencil beam optimization time was 238 ms per projection and 86 s in total. The full optimization, including the creation of the optimization matrices was 19 min.

Figure 8.4 shows WEPL sinograms and fluence sinograms for all three patients and for unit fluence as well as for the optimized fluence–modulations. The fluence sinograms only consider protons used for image reconstruction and are therefore subject to attenuation. Compared to the corresponding figures in chapters 6 and 7, the unit fluence sinograms are more homogeneous and contributions of single pencil beams are not visible, since the pencil beam grid was more dense in this chapter and no quarter–shift was employed. At two distinct locations at the detector at ±70 mm, increased or decreased counts were observed, which is caused by small gaps in the tracking detector for which the proton coordinates are interpolated. A careful inspection shows a slight effect of attenuation with reduced counts for regions with larger WEPL. In comparison to that, the fluence–modulated scans show a strong fluence modulation with regions outside of the ROI receiving a minimum fluence and subtle modulations in the ROI that show similar features to the WEPL sinogram. In particular at the edge of the patient, fluence is elevated over the level of the unit fluence–scan. Figure 8.4 (g) shows the sinogram of a binary fluence–modulation, which is equal to the unit fluence for pencil beams intersecting the ROI and reduced elsewhere.

> A discussion of these results can be found in section 8.5 on page 132.

8.4 EVALUATION OF OPTIMIZED FMPCT SCANS

MATERIALS & METHODS

For all three patients, uniform fluence pCT scans were simulated and compared to the corresponding optimized FMpCT scans. From the data, RSP maps were reconstructed using the FDK–DDB algorithm and the mean RSP error inside the ROI with respect to the ground truth RSP maps was calculated both for the pCT and the FMpCT scan. To assess the accuracy of achieving $V_{obj,i}$ inside the ROI, variance maps were reconstructed as described in section 3.5.4 and compared to the desired value.

The dosimetric accuracy of pCT and FMpCT scans was evaluated by importing the two RSP maps to the TPS as described in section 8.2. The therapeutic dose, which was optimized on the ground truth RSP map was then re–calculated based on the two simulated images. The resulting dose distribution was visually assessed for a slice at the center of the CTV. Dose–volume histograms calculated by the TPS were also compared between the ground truth dose and the dose re–calculated on the pCT and FMpCT scans. To quantify differences, a passing rate for a relative dose difference criterion of 1 % was calculated. For this, the difference between the pCT or FMpCT dose distribution and the ground truth was divided by the prescription dose. Dose voxels receiving less than 10 % of the dose prescription were disregarded.

The dosimetric accuracy was also evaluated in terms of the proton range. For this, a second treatment plan with a uniform dose target was optimized based on the ground

truth RSP map with just a single field and with a gantry angle as listed in table 8.1. This single field dose was then re–calculated on the pCT and FMpCT scans with a finer dose grid of $1\,\text{mm} \times 1\,\text{mm} \times 3\,\text{mm}$. This allowed to calculate the range for each voxel in beam–eye–view as the $80\,\%$ dose–falloff. Linear interpolation between voxels was used to achieve sub–millimeter resolution. To compare the resulting ranges to the ground truth, a passing rate was calculated for a $1\,\text{mm}$ criterion on the range difference between pCT or FMpCT and ground truth. For this, all rays intersecting the CTV (for patient 1 the low dose CTV) were considered.

To assess the imaging dose saving, the simulation code was used to score and sum the dose of each projection for both pCT and FMpCT scans. From this data, imaging dose–volume histograms were calculated. Median doses were evaluated for all OARs as well as the ROI and non–ROI volume. In addition to pCT and FMpCT scans, for patient 2, imaging doses were also calculated for the binary fluence–modulation.

The accuracy of the forward models for variance and imaging dose was evaluated by comparing the output of eqs. (8.3) and (8.7) for the converged set of bixel–weights to variances and imaging doses calculated with the Monte Carlo simulation with the optimized pencil beam weights. This quantifies inaccuracies due to assumptions of the forward model, such as the modulation in small bixels instead of pencil beams, including the assumption that parallel bixels are independent from each other, which is not strictly true for both dose and variance, and the simplified variance reconstruction formula, which does not account for distance–driven binning and the convolution.

RESULTS

Figure 8.5 shows image variance maps for pCT and the corresponding FMpCT scans. For the unit fluence pCT scans in the first column, variance is not homogeneous: noise is increased at the skin as well as close to heterogeneities such as the nasal cavity of patient 2 or the pharynx of patient 3. This is in agreement to expectations from chapters 5 to 7. In the FMpCT scans in the second column the noise level is homogeneous inside the ROI, even in regions that had elevated noise levels in the pCT scan. Outside of the ROI, image variance increases sharply with just a small transition region of a few millimeters around the ROI. These observations are confirmed in the third column, where line profiles along the white dashed lines show a strong modulation inside the ROI, which is indicated as a green shaded area. Variance for FMpCT scans is flat in the ROI and increases sharply outside. The variance objective is met for all three patients with only small deviations. Peak noise for pCT and FMpCT agrees at the hull of the patient, which can be outside of the slice for which the profiles are displayed (in particular for patient 3). In most regions, pCT noise is intentionally below the FMpCT noise level.

Table 8.3 lists the mean RSP values inside the ROI for the ground truth RSP and the two pCT images. Absolute relative errors are all below $0.6\,\%$ for the pCT scans and below $0.8\,\%$ for the FMpCT scans. This is within the magnitude expected for typical pCT scans and agrees with results from previous chapters.

Figure 8.6 displays the treatment doses for all three patients: (a,e,i) as they were optimized based on the ground truth RSP map, and re–calculated for (b,f,j) the pCT or (c,g,k) the
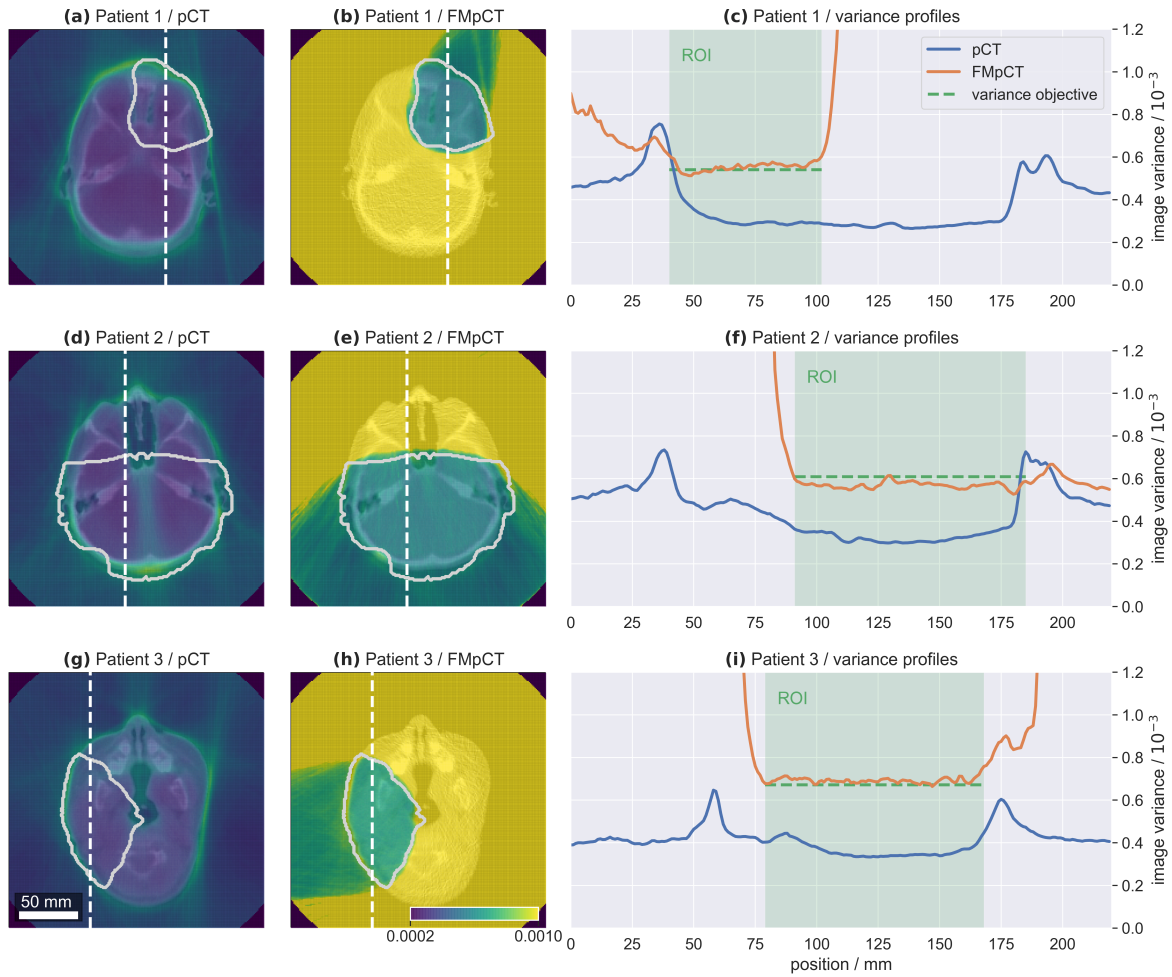
Figure 8.5: Image variance maps for (a, d, g) the unit fluence pCT scans, (b, e, h) the fluence modulated scans, and (c, f, i) profiles along the white dashed lines. The ROI is indicated by a solid white line in the variance maps and as a green shaded area in the profile plots. The variance maps are overlaid on top of the corresponding RSP maps to make the anatomy visible. ⓒ Reproduced from Dickmann et al. [138] under the CC BY 4.0 license.

FMpCT scan. The treatment doses are shown on top of the corresponding RSP maps, where for the FMpCT scan a clear increase of noise could be observed outside of the region covered by the treatment dose. Treatment doses were not distinguishable visually. The dose volume histograms in fig. 8.6 (d,h,l) were also in agreement and only for the left hippocampus of patient 2, which was located close to the end of the range of one treatment beam, differences between the ground truth dose and the dose calculated on the two pCT scans, were visible. Even organs like the brain of patient 1, which were imaged with high image noise in the FMpCT scan, showed a good agreement in the dose–volume histogram.

In table 8.4 passing rates are shown, which compare the re–calculated dose maps to the dose maps that was optimized on the ground truth RSP. Evaluations are shown for a strict 1 % criterion on the absolute difference. The largest discrepancy, and thus the lowest passing rate was found for the comparison of FMpCT to the ground truth for patient 3, which was

Table 8.3: Evaluation of the mean RSP value inside the ROI and calculation of the relative error with respect to the ground truth mean RSP value.

RSP map	patient 1	patient 2	patient 3
ground truth	1.018	1.057	0.966
pCT	1.017 (error −0.1 %)	1.051 (error −0.6 %)	0.966 (error 0.0 %)
FMpCT	1.026 (error 0.8 %)	1.050 (error −0.7 %)	0.962 (error −0.4 %)

Table 8.4: Passing rates for a 1 % criterion comparing the ground truth (GT) dose maps to those re–calculated on the pCT and FMpCT scans.

comparison	patient 1	patient 2	patient 3
pCT vs. GT passing rate / %	90.6	91.8	90.6
FMpCT vs. GT passing rate / %	91.3	93.1	88.5
FMpCT vs. pCT passing rate / %	98.3	99.7	98.9

Table 8.5: Passing rates for a 1 mm, mean absolute difference, and mean difference for the range analysis comparing the ground truth (GT) ranges to those evaluated on the pCT and FMpCT scans.

quantity	comparison	patient 1	patient 2	patient 3
passing rate / %	pCT vs. GT	97.4	100.0	95.3
	FMpCT vs. GT	96.5	96.8	95.0
	FMpCT vs. pCT	100.0	97.5	99.9
mean abs. diff. / mm	pCT vs. GT	0.29	0.25	0.28
	FMpCT vs. GT	0.31	0.32	0.32
	FMpCT vs. pCT	0.15	0.17	0.05
mean diff. / mm	pCT vs. GT	0.00	0.04	0.27
	FMpCT vs. GT	0.15	0.20	0.31
	FMpCT vs. pCT	−0.15	−0.16	−0.05
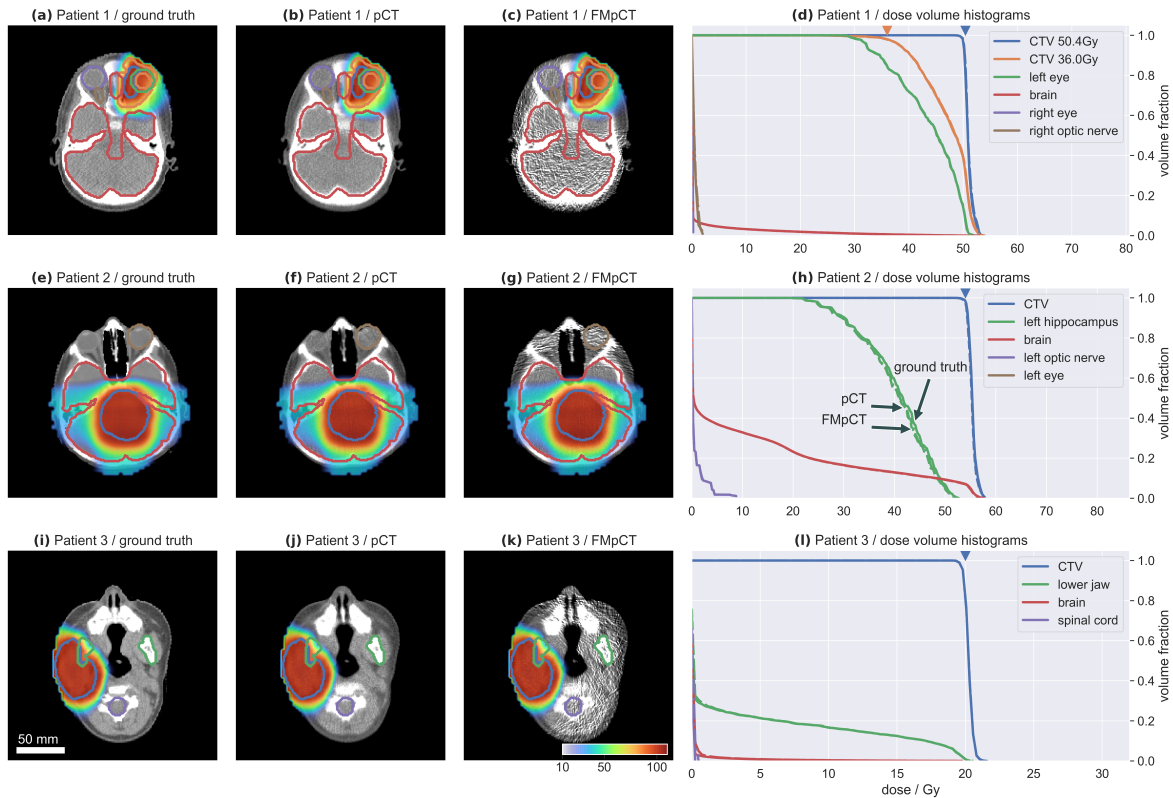
Figure 8.6: Treatment doses optimized on the (a,e,i) ground truth RSP map and re–calculated on (b,f,j) pCT and (c,g,k) FMpCT images together with (d,h,l) dose–volume histograms. Treatment doses are shown on top of the corresponding RSP maps. The color scale of treatment doses is in percent of the prescription dose. The CTV and OAR contours use the same colors as in the dose–volume histograms. In the histograms, the three dose sets are mostly indistinguishable. Triangles on top of the histogram lines indicate the corresponding prescriptions. ⓒ Reproduced from Dickmann et al. [138] under the CC BY 4.0 license.

still at 88.5 %. All other passing rates were above 90 % and in particular the passing rates comparing FMpCT to pCT were above 98 % for all patients.

Table 8.5 investigates the range accuracy of pCT and FMpCT images compared to the ground truth and lists passing rates for a 1 mm criterion. Passing rates for all comparisons were above 95 %. The best average agreement over all patients was between FMpCT and pCT. The mean absolute difference and the mean difference between the evaluated ranges, which is also shown in table 8.5, was well below 0.5 mm for all three comparisons and all patients.

Figure 8.7 shows the imaging dose for the pCT and the FMpCT scan, which were both matched to the peak noise level inside the ROI. The imaging dose of the pCT scan was homogeneous as expected, with slight reductions in bony structures due to the different elemental composition. The imaging dose for the FMpCT scans was concentrated in the ROI and drops further away from it. Imaging OARs, such as the right eye of patient 1 and both eyes of patient 2 received less imaging dose compared to their local surrounding. In particular the eye of patient 1 was at a higher dose level compared to the back of the brain, but its imaging dose was considerably less compared to other tissue with a similar distance
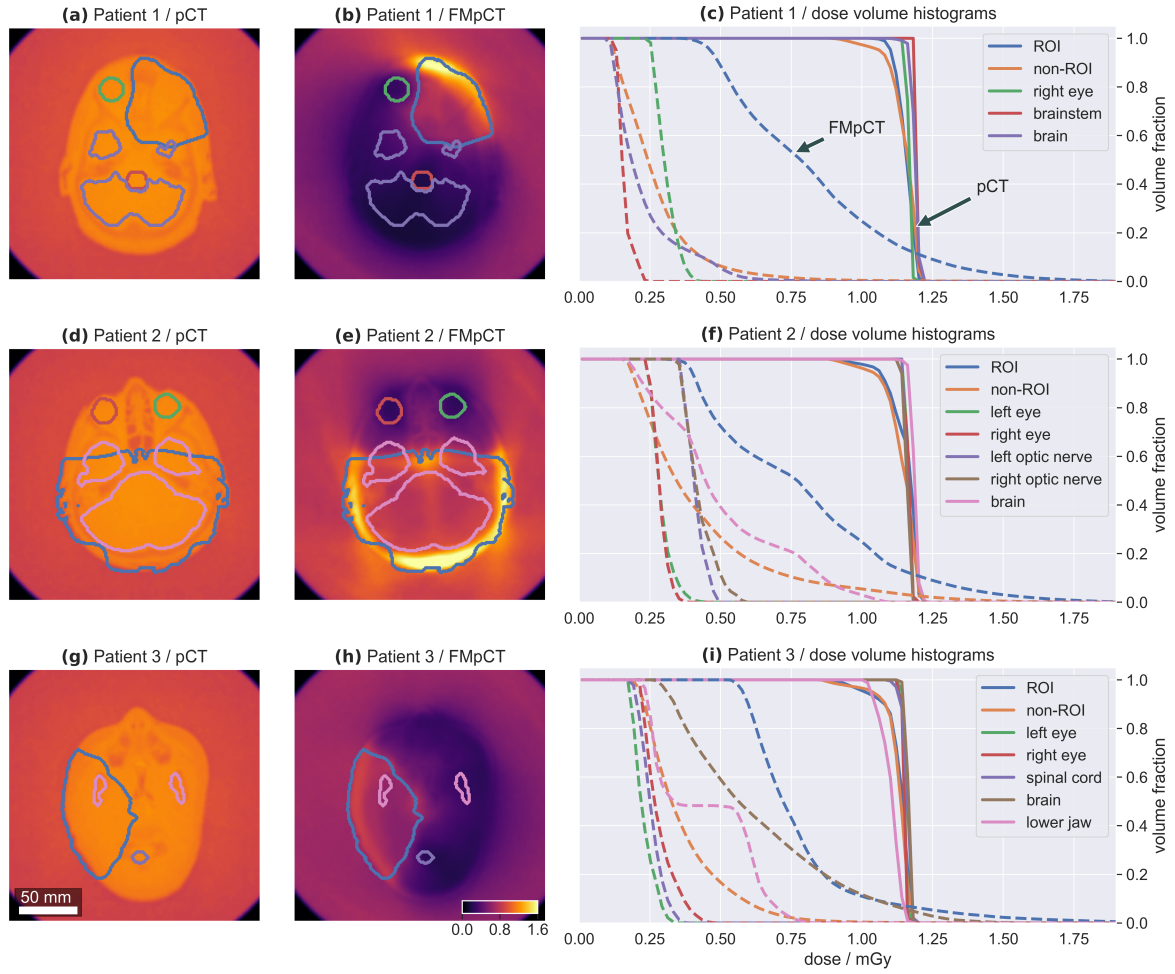
Figure 8.7: Imaging doses of (a,d,g) the pCT and (b,e,h) the FMpCT scans, and (c,f,i) corresponding imaging dose–volume histograms. ⓒ Reproduced from Dickmann et al. [138] under the CC BY 4.0 license.

to the ROI. Imaging dose inside the ROI was non–homogeneous and elevated at the edge and in regions of elevated image noise for the uniform fluence scan (compare fig. 8.5). In particular at the edge and for patients 1 and 2, imaging dose was increased compared to the unit fluence scan.

The observations are confirmed in the imaging dose–volume histograms in fig. 8.7 (c,f,i). The uniform fluence pCT dose was at around 1.2 mGy for all three patients and all structures. For limited volumes with different elemental composition (e. g. the jaw of patient 3), lower doses also occurred, but all relevant doses were above 1.0 mGy. Inside the ROI, peak dose levels were elevated for FMpCT scans compared to the pCT scans. The dose increase reached up to 1.8 mGy, but was limited to about 10 % of the total ROI volume. Most of the ROI received a lower dose in the FMpCT scan compared to the pCT scan. All OARs and in general the complete non–ROI volume received a considerably reduced imaging dose in the FMpCT scans. The only exception for this is patient 2, where 2.5 % of the non–ROI volume received a dose above 1.2 mGy.
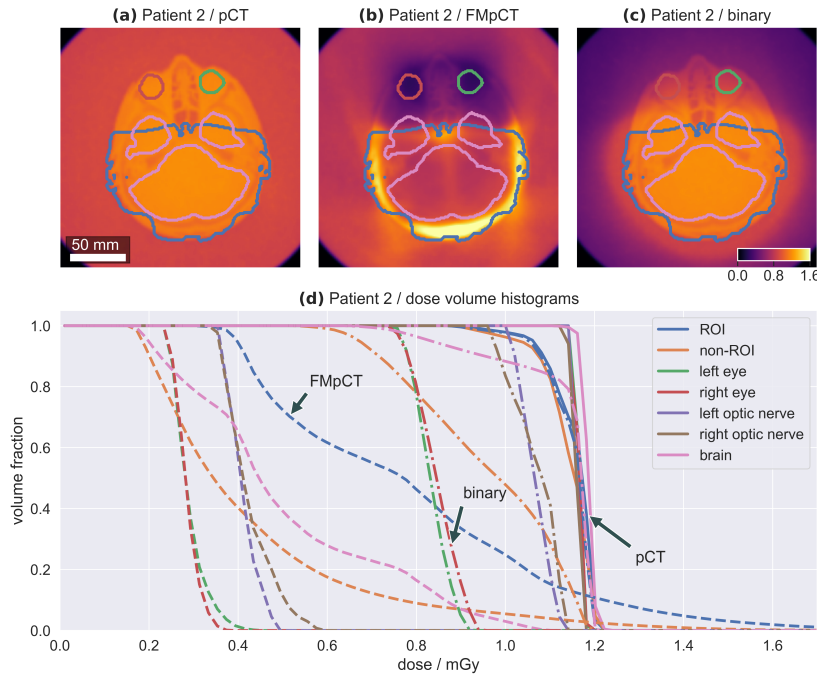
Figure 8.8: Comparison of the imaging dose in mGy for patient 2 and (a) the pCT scan, (b) the FMpCT scan, and (c) the scan with the simple binary fluence modulation, as well as (d) the corresponding imaging dose–volume histograms. ⓒ Reproduced from Dickmann et al. [138] under the CC BY 4.0 license.

In fig. 8.8 (c) the imaging dose of the binary fluence modulation in eq. (8.20) is shown for patient 2 and compared to the corresponding pCT and FMpCT imaging doses, which are shown again in (a) and (b). The dose of the binary modulation inside the ROI is, by construction, equal to the dose in the uniform fluence pCT scan. Outside of the ROI, imaging dose drops, but the decrease is much less dramatic compared to the FMpCT scan. Moreover, the eyes are not spared additionally from imaging dose. This is confirmed in the imaging dose–volume histograms in fig. 8.8 (d), where the binary fluence modulation shows a dose reduction outside of the ROI, but for example the eyes are at approximately double the imaging dose compared to the FMpCT scan.

Table 8.6 lists the median imaging doses for all three patients and for pCT, FMpCT, and binary–modulated pCT scans. The uniform fluence doses were comparable for all patients and structures and between 1.14 mGy and 1.19 mGy. For modulated scans, the percentage dose saving is given in brackets. For all three patients, the FMpCT dose savings outside the ROI and in particular in OARs was considerable and of comparable magnitude between the three patients. Dose savings for some OARs are above 80 %. The binary modulation had the same imaging dose inside the ROI as the pCT scan and dose savings outside reached 29 %, but were all considerably less compared to the FMpCT scan.

Figure 8.9 displays image variance and imaging dose maps calculated from the Monte Carlo simulation in the first column. Those maps are the same as shown in figs. 8.5 and 8.7. The second column shows the corresponding predictions of the bixel–wise forward model of eqs. (8.3) and (8.7). The third column shows line profiles along the white dashed lines

Table 8.6: Median imaging doses for all patients and for pCT, FMpCT and binary modulated scans. Doses are given for the ROI and all relevant OARs. Dose savings compared to the unit fluence scan are given in parentheses. *Abbreviations:* ROI – region–of–interest, l – left, r – right.

	region	pCT dose/mGy	FMpCT dose/mGy (saving/%)		binary dose/mGy (saving/%)	
patient 1	ROI	1.16	0.78	(−33)	—	
	non–ROI	1.16	0.24	(−80)	—	
	r. eye*	1.17	0.30	(−74)	—	
	brainstem*	1.19	0.15	(−87)	—	
	brain	1.19	0.19	(−84)	—	
patient 2	ROI	1.17	0.77	(−35)	1.17	(0)
	non–ROI	1.16	0.34	(−71)	0.98	(−16)
	l. eye*	1.17	0.28	(−76)	0.83	(−29)
	r. eye*	1.17	0.28	(−76)	0.85	(−27)
	l. optic nerve*	1.16	0.41	(−65)	1.07	(− 8)
	r. optic nerve*	1.16	0.41	(−65)	1.08	(− 7)
	brain	1.19	0.45	(−62)	1.17	(− 2)
patient 3	ROI	1.14	0.72	(−37)	—	
	non–ROI	1.14	0.33	(−71)	—	
	l. eye*	1.15	0.22	(−81)	—	
	r. eye*	1.15	0.27	(−77)	—	
	spinal cord*	1.16	0.25	(−78)	—	
	brain	1.17	0.56	(−52)	—	
	lower jaw	1.11	0.34	(−69)	—	
average	ROI	1.16	0.76	(−35)	1.17	(0)
	non–ROI	1.15	0.30	(−74)	0.98	(−16)

* imaging OAR

for both quantities as well as the variance objective for the variance profiles. The variance predictions agree well with the values scored from the Monte Carlo data. Differences are visible only outside of the patient. Variance predictions in the ROI are flatter compared to the simulated values employing pencil beams. While for patient 1 the prediction and simulation agree well, for patient 2 the model slightly overestimates and for patient 3 it slightly underestimates the simulated values. For the imaging doses, estimations and simulations agree well visually. Also in the profiles, the predictions follow the Monte Carlo values with only minor differences. Unlike the variance maps, the imaging dose profiles are comparably smooth between prediction and simulation.
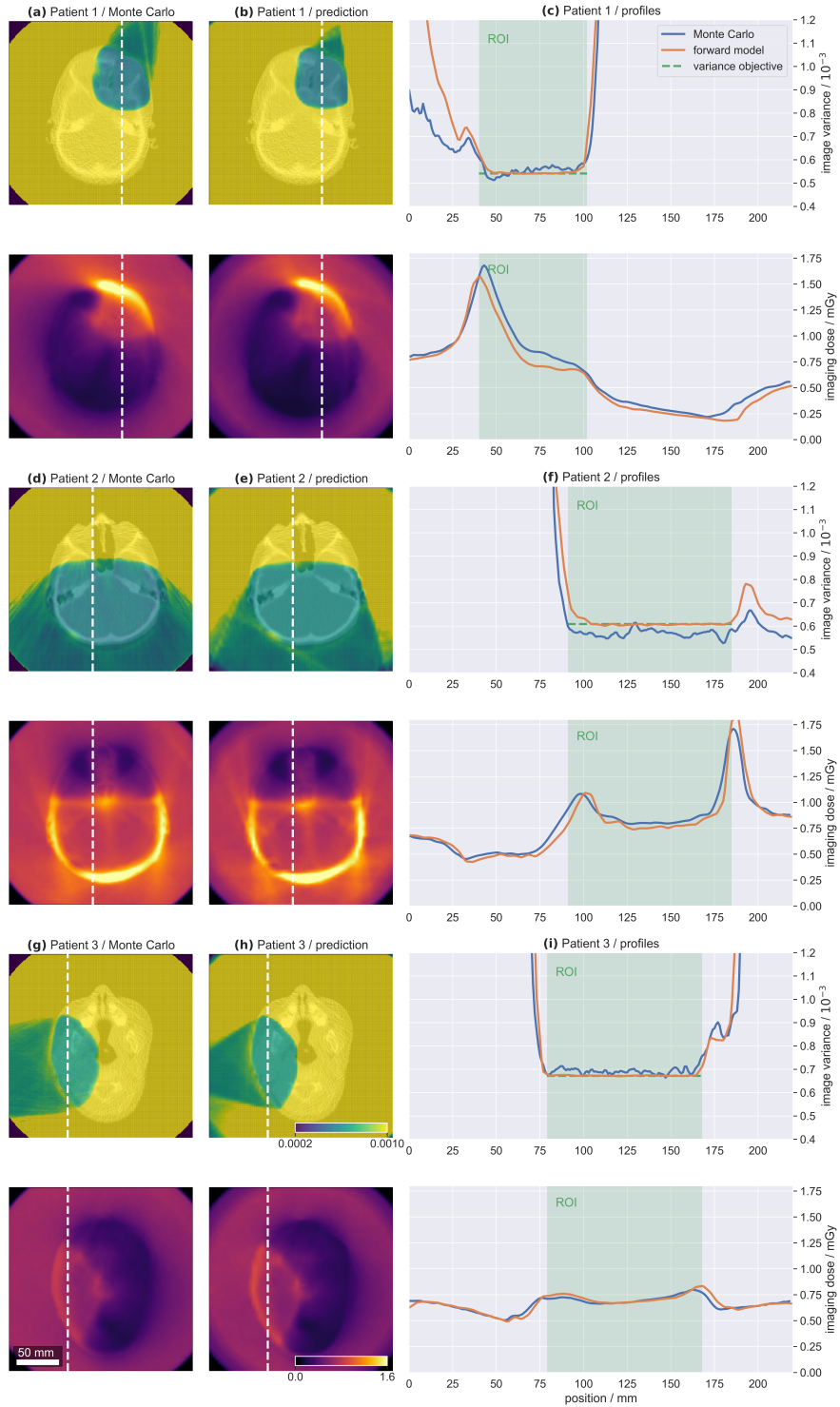
Figure 8.9: Predictions of the optimizer's forward model in the first column (a,d,g) and actual values from the Monte Carlo simulation in the second column (b,e,h) for imaging dose (in mGy) and image variance maps (dimensionless) of the converged fluence–modulations of all three patients. The last column (c,f,i) shows line profiles along the white dashed lines. The ROI is indicated as a green shaded area.

8.5 DISCUSSION

An algorithm for joint dose minimization and variance optimization

A novel optimization algorithm for FMpCT was proposed, which considers both image variance and imaging dose objectives. It consists of a bixel–wise fluence optimization followed by a pencil beam optimization to each projection fluence. Both optimizations converged fast and reliably for all three patients. The bixel–wise optimization required between 100 and 200 iterations to converge, but was run for 500 iterations to ensure optimal results. The pencil beam optimization, which considered only a single projection fluence at a time, converged more quickly within less than ten iterations. For the bixel–wise optimization, the final cost function value was different for each patient, since they had different noise targets and, more importantly, different ROI volumes. Typically, for treatment planning, the penalties would be adjusted for each patient to ensure an optimal solution. This was not done in this study to make results comparable between patients and to investigate the robustness of the method. The cost function of patient 1 was dominated by the dose term, while for the other two patients the variance term prevailed. This may be explained by the different ROI volumes of the patients, where patient 1 had the smallest ROI volume and patient 2 the largest. For a small ROI volume, less voxels contribute to the sum of the variance term in eq. (8.10) and more contribute to the dose term, which, as a result, is higher.

The proposed method makes use of an initial bixel–wise optimization and only optimizes pencil beams in a second step. The dose term could have been optimized end–to–end by calculating the dose contribution to voxel i from pencil beam k (instead of bixel i). However, this was not possible for the variance term due to the inverse relationship between fluence and variance. The variance contribution of pencil beam k would be infinity for all voxels i outside of the extent of the pencil beam. To solve this, parallel and mutually independent bixels were introduced that allowed to reproduce the variance reconstruction formula with the same bixel weights that could be used for the imaging dose forward calculation. The two–step design of the algorithm, therefore, was required to correctly account for the variance term. As a side effect, this also considerably reduced the computational cost of the dose forward model, since the dose matrix with bixels is much sparser compared to a dose matrix with pencil beams.

The evaluation of the cost function was implemented efficiently using sparse matrix multiplications, which reduced the evaluation time to few hundred milliseconds. The dose and variance matrices had in total $N \cdot M \approx 5 \times 10^{10}$ entries each, but most were zero. With a sparse implementation of the matrices, only non–zero components are stored in memory and used in multiplication operations. With the optimization itself being fast, most of the computation time was attributed to the creation of the matrices, which involved reading dose and variance projections from disk as well as querying the pencil beam model. For the creation of the matrix P_{ik}, the pencil beam model in eq. (6.1) was queried on a grid with $1\,mm^3$ voxels with each voxel requiring a computationally expensive call to the exponential function (unless the voxel was outside of a three–standard–deviation interval). However, the matrices D_{ij} in eq. (8.3), V_{ij} in eq. (8.7), F_{ij} in eq. (8.13), and P_{ik} in eq. (8.15) can be calculated as soon as the uniform fluence pCT scan and the corresponding Monte Carlo

simulation are available. This could be done automatically before any user interaction to reduce the optimization time to a few minutes if the matrices are readily available. If calculation speed is crucial, the optimization time could be further reduced by stopping the optimization already after 200 or even 100 iterations, which seems feasible with only a minor degradation of the algorithm performance given the cost function values at these iterations. This would bring the time down to few tens of seconds and could further be improved by avoiding unnecessary calculations, such as variance values for voxels where the variance weight is zero. Altogether, it appears reasonable that the algorithm is feasible time–wise to be used in a clinical context with calculation times comparable to those of treatment plan optimization.

The reduction of free parameters by quarter–shifting the pencil beam grid and increasing the spacing in u–direction was avoided in this study and a denser pencil beam grid was used instead. While the quarter shift has worked in simulations in chapter 6, it was concluded in chapter 7 that it makes the fluence delivery susceptible to small setup shifts, which can occur experimentally. This was not repeated in this study, potentially making a future delivery of such patterns more robust.

Evaluation of optimized FMpCT scans

The performance of the proposed optimization algorithm for FMpCT was evaluated by re–calculating treatment doses on simulated pCT and FMpCT scans. The agreement of the resulting treatment doses with the dose optimized on a ground truth RSP map was satisfactory — both in terms of the overall dose accuracy and in terms of the resulting range error. Passing rates were calculated for a 1 % criterion on the treatment dose and for a 1 mm criterion on the ranges. These are strict criteria compared to those used in literature, which are typically 3 %/3 mm [192–194]. All passing rates were above 88 % with respect to the ground truth dose and range distributions. Passing rates for the comparison of pCT and FMpCT images were considerably better and above 97 %, which suggests that the driving contribution of dose calculation errors is the use of pCT instead of the ground truth RSP map, and not the application of fluence modulation. The RSP errors, which can lead to dose calculation distortions, were quantified to be below 1 % compared to the ground truth and similar for pCT and FMpCT. An error of 1 % is expected for the experimental operation of the scanner [41], and also using the realistic simulation used in this study [103]. The range errors introduced by FMpCT scans were also slightly higher compared to pCT scans, but the differences were small compared to the variation in between patients. Therefore, no significant deterioration by fluence modulation was found and the dosimetric accuracy of modulated an un–modulated scans is equivalent. This is supported by the fact that accuracy differences between the ground truth and the two pCT scans was much larger than the difference in between pCT and FMpCT. Dosimetric and range errors using pCT were comparable to values determined in a previous study by Meyer et al. [43].

In agreement to chapter 6, imaging doses were constant for the uniform fluence pCT scans, but image variance was increased at the hull or close to heterogeneities. With FMpCT, it was possible to invert this and achieve constant variance inside the ROI with an inhomogeneous imaging dose distribution. Achieving a constant variance allowed to reduce the median

imaging dose even in the ROI by 35 % on average. This was possible, since pCT and FMpCT scans were matched for peak variance in the ROI and an increase of variance was possible in many regions of the ROI. This dose reduction, however, is not very relevant in reality, since it is negligible compared to the treatment dose in this region.

Variance outside of the ROI was strongly increased in the FMpCT scans, in particular further away from the ROI. This allowed to reduce the median imaging dose by on average 74 % compared to uniform fluence scans. The dose saving was consistent between all three patients and ranged from 71 % to 80 %, even though the ROI shapes and volumes differed considerably in between patients. The largest dose saving was achieved for the patient with the smallest ROI volume. Imaging doses to specific OARs were further reduced by the optimizer and OAR dose savings reached up to 87 %. The closer an OAR was to the ROI, the smaller the possible dose saving was, which is why imaging doses to OARs need to be compared that in their direct vicinity. The relative dose penalty of imaging OARs was fixed in this study, but would need to be determined based on clinical factors and the available imaging dose–volume metrics — if at all an imaging OARs is needed. A potential clinical indication requiring to limit radiation exposure by imaging could be radiation–induced cataract [195], which can occur in the eye and is particularly relevant for young patients. Other imaging OARs used here, were chosen to showcase a potential application of the algorithm and would eventually need to be defined by a physician.

In chapter 6 and in Dickmann et al. [156], dose savings of up to 40.5 % were reported outside of the ROI and compared to unit fluence scans while matching the peak variance level. The results achieved with the novel optimization algorithm for dose and variance objectives, are considerably improved by about a factor of two. This could be due to the generally smaller ROI volumes used in this study compared to the large ROI used in chapter 6. However, even for patient 2, for which the ROI in the central slice covered about half of the patient's cross section, a dose saving of over 70 % was achieved. The improvement compared to chapter 6 is, therefore, likely to stem from the optimizer's ability to achieve a given variance in the required regions while reducing imaging dose where this is beneficial instead of implicitly reducing dose via the prescription of a higher variance. Moreover, the prescription of a variance contrast of four in chapter 6 may not have been the optimal choice and the novel optimizer was not restricted to any maximum variance level outside of the ROI, apart from the minimum fluence needed in each pixel.

The relevance of imaging dose savings achieved using the proposed method must be assessed by comparing it to the local treatment dose. If the dose from imaging is irrelevant compared to the dose delivered during treatment, the effort of reducing imaging dose has no impact. This is the case for the imaging dose inside the ROI, where a single fraction of the treatment exposes the patient to doses between 1.8 Gy and 2.5 Gy (see table 8.1). Imaging doses from pCT are only at 1.2 mGy and therefore lower by three orders of magnitude. Dose savings by FMpCT in the ROI are, therefore, not relevant. This applies also to the dose increases that were observed for a small part of the ROI for FMpCT scans. The dose saving outside of the ROI, however, which was on average 0.9 mGy, may be relevant compared to the treatment dose of a single fraction, which is at 180 mGy to 250 mGy at the hull of the ROI, but quickly drops to zero due to the steep dose gradients achievable with proton therapy. The treatment doses as determined by the TPS were in fact exactly zero for most imaging

OARs. Nevertheless, it is known, that the low doses are not correctly determined in a TPS, since they often do not consider dose from secondary neutrons. Neutrons can originate from stray dose from the beam line elements or scattered from the room walls, or can be created in the patient. Neutron doses are in the order of magnitude of a few milli–Sievert per fraction according to Schneider and Hälg [196]. Assuming that pCT imaging is to be performed prior to every treatment faction, the potential dose saving due to FMpCT appears to be relevant. However, a precise study including the calculation of neutron dose is necessary to answer this question definitively.

The predictions of the optimizer for image variance and imaging dose were compared to the same quantities calculated from Monte Carlo simulations. While minor differences were identified, the agreement between the prediction and the simulated value was satisfactory. This would allow to directly employ optimized FMpCT fluence patterns without the need to wait for an additional Monte Carlo simulation to verify the optimization result.

While this study used a realistic simulation of a prototype pCT scanner and patient data to generate images, it did not investigate the potential impact of anatomical changes, which can occur during proton therapy treatment, and which may impact the required fluence modulation patterns. Translational and rotational errors are expected to be small for patients of the head and neck region, since they can be precisely positioned using thermoplastic face masks. However, internal anatomical changes due to weight loss or reduction of the tumor volume need to be investigated carefully in future studies. It may be possible to update the required fluence patterns just from the previous FMpCT scan. If this is not an option, uniform fluence scans may need to be repeated when anatomical changes are suspected, which would cause an additional imaging dose, but ensure a precise delivery of the imaging fluence.

8.6 CONCLUSIONS

In this chapter and originally in Dickmann et al. [138], a novel optimization algorithm for FMpCT was proposed, that considers both image variance and imaging dose targets. Conceptually, this is a step forward compared to the algorithm proposed in chapter 6 and in Dickmann et al. [156], since it allows to directly optimize for the quantities that are relevant in each voxel of the volume. It avoids the need to indirectly reduce imaging dose by prescribing an arbitrarily chosen higher variance target. Moreover, the algorithm allows to increase the dose penalty in pre–defined imaging OARs, where an additional dose saving can be achieved. The performance of the algorithm was demonstrated in a Monte Carlo simulation study using a realistic pCT scanner and proton beam model. For this purpose, imaging data of three pediatric patients was used to simulate realistic patient geometries and have ROI definitions based on actual proton therapy dose distributions. For the three patients, which had different locations of the tumor within the head, an imaging dose saving of 74 % outside of the ROI was achieved, which clearly outperformed the optimization algorithm in chapter 6 as well as a simple intersection–based fluence modulation. Based on the low–dose FMpCT scans as well as uniform fluence pCT, ground truth treatment doses were re–calculated on the simulated images. Passing rates for a 1 % criterion on dose and a

1 mm criterion on proton range were all well above 90 % when comparing FMpCT images to the corresponding uniform fluence images. Range uncertainties were below 0.3 mm. It is, therefore, concluded that the usage of fluence modulation does not relevantly deteriorate the dosimetric accuracy of pCT. Dose reductions achievable with FMpCT appear to be relevant compared to the out–of–field treatment dose, which is generally low for particle therapy. In conclusion, fluence–modulated pCT with dose and variance objectives allows to calculate patient–specific fluence maps, that considerably reduce the imaging dose required to acquire RSP maps with sufficient quality for treatment dose calculation and at imaging doses of only 0.3 mGy per scan outside of the ROI. This opens an interesting perspective for image–guided and adaptive particle therapy.

9 | IMAGE ACCURACY OF A PROTOTYPE SCANNER

Proton computed tomography was suggested by Cormack [36] in 1963 as an alternative to x–ray CT. Today it is envisioned to improve the accuracy of particle therapy treatment planning compared to x–ray based imaging modalities like DECT, which are already starting to become available in clinics today. For this purpose, future pCT scanners need to be on par with or outperform the accuracy achievable with DECT, for which RSP errors of 1 % have been reported [22, 68, 69, 78]. Current prototype pCT scanners appear to be performing comparably well with expected errors better than 1.6 % [87], 1.4 % [140], or 0.74 % [88]. A first direct comparison between DECT and pCT using the same phantoms was performed by Dedes et al. [41] in connection to the works presented in this thesis. They concluded that the MAPE of pCT was on par with the performance of a state–of–the–art DECT scanner.

This comparison [41] identified artifacts resulting from limited distorted WEPL intervals to degrade the performance of the pCT scanner under investigation. The WEPL intervals were linked to interfaces between the five stages of the prototype scanner and could be identified using a method developed by the author, which calculates heat–maps in image space for given distorted WEPL intervals. This has led to the development of two alternative artifact correction methods which are presented in this chapter.

The first method directly addresses the physical origin of artifacts and aims to avoid protons stopping close to stage interfaces by merging data at two incident beam energies. With this energy modulation method, the amplitude of artifacts in an experimental acquisition of the homogeneous water phantom was reduced. The inter–quartile range of RSP values, which quantifies the extent of artifacts, decreased by a factor of two to three. A negative bias in the image, which may be linked to systematic errors in the prototype scanner's calibration process, persisted and the average RSP remained underestimated by 1 %.

The second method developed to address RSP artifacts of the pCT scanner is purely empirical and makes no assumption on the origin of WEPL distortions. Instead, it uses a scan of a phantom with known RSP to isolate artifacts in the image. A decomposition of the RSP map into contributions from small–banded WEPL intervals, makes it possible to find a WEPL correction function that can corrects for artifacts in subsequent scans. Using a custom–built elliptical correction phantom, the method was applied to experimental scans of the water phantom, the CTP phantom and the head phantom. In all phantoms, the negative offset seen before and most RSP artifacts were removed. The MAPE calculated on the data of the CTP phantom and the water phantom improved from 0.86 % without the correction

down to below 0.48 % using the proposed method. This reduction by on average 47 % is a substantial improvement of pCT image accuracy.

Please note, that this chapter is organized differently compared to the previous chapters: section 9.1 gives a short motivation, and sections 9.2 to 9.4 present independent, but connected, investigations. Each section contains a consecutive presentation of materials & methods, results and a discussion. A common conclusion of all investigations is presented in section 9.5.

Some results presented in this chapter were published in *Physica Medica* [197] and other results were submitted for publication [141]. The author also developed a method, which is presented here and which was used in a publication in *Physics in Medicine & Biology* [41]. This specific method was presented at the MCMA conference in Montréal, Canada [198]. The author also contributed to a review article on the role of Monte Carlo simulation for pCT, which was published in *Zeitschrift für Medizinische Physik* [102]. The corresponding publications are stated explicitly at the beginning of sections 9.2 to 9.4.

PUBLICATIONS IN THIS CHAPTER

[197] » J. Dickmann et al.: "Proof of concept image artifact reduction by energy-modulated proton computed tomography (EMpCT)". *Physica Medica* **81** (2021), 237–244. ISSN: 11201797. DOI: 10.1016/j.ejmp.2020.12.012

[141] » J. Dickmann et al.: "An empirical artifact correction for proton computed tomography". *submitted for publication* (2021)

CONFERENCE CONTRIBUTIONS IN THIS CHAPTER

[198] » J. Dickmann et al.: "Understanding image artifacts for a prototype proton computed tomography scanner via Monte Carlo simulations". *International Conference on Monte Carlo Techniques for Medical Applications (MCMA), Montréal, Canada* (2019)

PUBLICATION CONTRIBUTIONS IN THIS CHAPTER

[41] » G. Dedes, J. Dickmann et al.: "Experimental comparison of proton CT and dual energy x-ray CT for relative stopping power estimation in proton therapy". *Physics in Medicine & Biology* **64**(16) (2019), 165002. ISSN: 1361-6560. DOI: 10.1088/1361-6560/ab2b72

[102] » G. Dedes, J. Dickmann et al.: "The role of Monte Carlo simulation in understanding the performance of proton computed tomography". *Zeitschrift für Medizinische Physik* (available online, 2020). ISSN: 0939-3889. DOI: 10.1016/j.zemedi.2020.06.006

9.1 EXPERIMENTAL COMPARISON OF PROTON CT AND X–RAY CT

The state–of–the–art for calculation of RSP maps for proton therapy range calculation is the use of DECT scanners, that probe the x–ray attenuation coefficient at two different

spectra and thereby allow for a better conversion to RSP. As discussed in section 2.3.2, they allow for an RSP accuracy of 1 %, which seems to be comparable to the accuracy achievable with current prototype pCT scanners. While still only few clinics use DECT routinely for proton range calculation [70], it is likely that they will supersede single energy CT in the near future. Consequently, a direct comparison between the performance of clinical DECT scanners and prototype pCT scanners is of high interest.

Dedes et al. [41] used the pCT prototype scanner, which was also used in this work, to scan two phantoms with in total 13 tissue–equivalent inserts. For each of the inserts, a ground truth RSP value was determined using water–column measurements. The same phantoms were then also scanned at a clinical DECT scanner (SOMATOM Definition FORCE, *Siemens Healthinieers*, Forchheim, Germany) and RSP maps were calculated using the state–of–the–art DECT calibration of Saito and Sagara [199]. This allowed for a direct comparison of RSP errors with respect to the ground truth values between the pCT and DECT scans. The accuracy of each imaging modality was quantified by the MAPE over all inserts of both phantoms. For DECT, this resulted in a MAPE of 0.67 %. For pCT the MAPE was slightly better at 0.55 %. The pCT, performance was, therefore, concluded to be on par with the DECT performance.

The study [41] also investigated intensively the sources of inaccuracies of the prototype pCT scanner. In particular, artifacts at fixed WEPL intervals due to protons stopping close to stage interfaces as well as due to the shape of the calibration phantom were suspected to degrade imaging performance. The artifacts were studied by comparing the experimental scans to realistic and idealized simulations of the imaging setup. A method for calculating the impact of distorted WEPL intervals on the reconstructed volume was developed by the author and is described in section 9.2. This method helped to identify sources of image artifacts and Dedes et al. [41] conclude that their mitigation is critical for the further improvement of pCT imaging performance.

9.2   IDENTIFICATION OF PROTON CT IMAGE ARTIFACTS

◫ The method presented in this section was developed by the author and used in Dedes et al. [41].

MATERIALS & METHODS

WEPL intervals at the stage interfaces were suspected to distort the pCT signal and cause circular rings in the reconstruction of homogeneous cylindrical phantoms. This is illustrated in fig. 9.1, where protons stopping at stage interfaces are indicated in red. For a homogeneous cylindrical phantom, these protons will always fall in the same annulus over the full rotation of the tomographic reconstruction. A distortion of protons stopping at stage interfaces would, consequently, result in a ring–shaped artifact in the reconstructed image. In addition, the WEPL calibration curve had irregularities at multiples of the calibration phantom's brick thickness (illustrated in fig. 3.2), which would result in similar rings, but which are not considered in the following since they were of a minor magnitude.
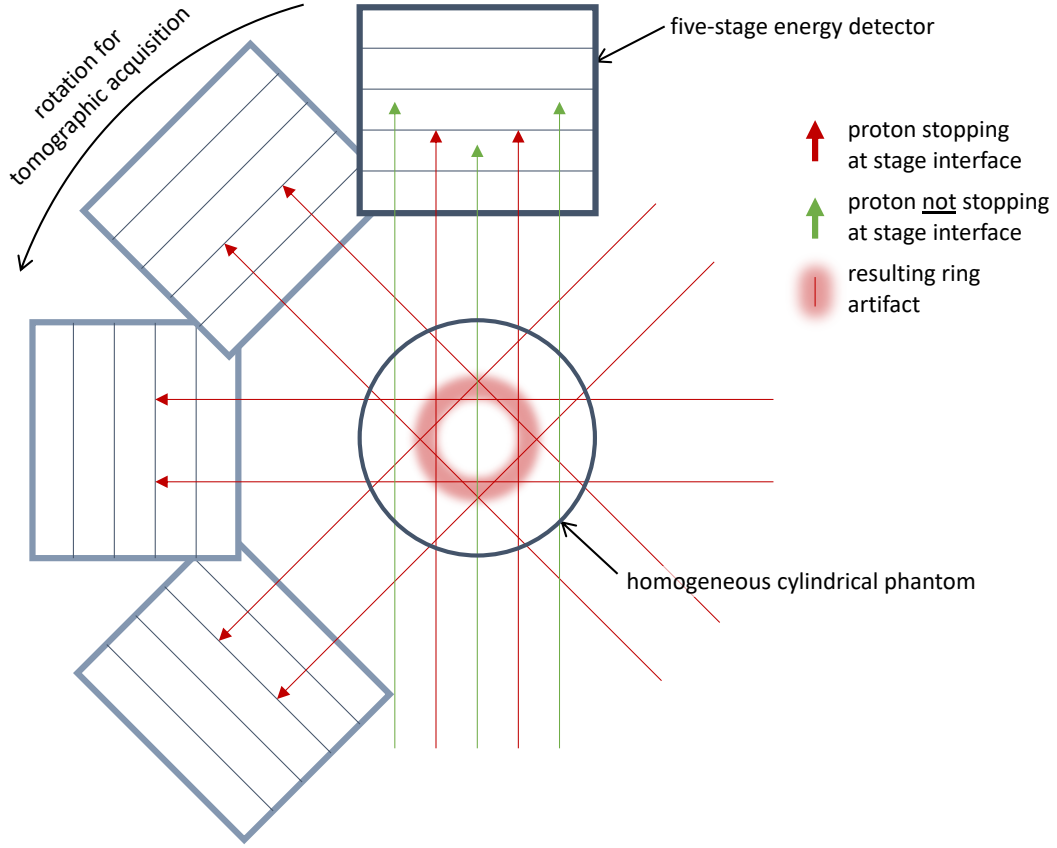
Figure 9.1: Illustration of ring–shaped image artifacts resulting from stage transitions of the five–stage energy detector for a homogeneous phantom.

Each voxel of an image is reconstructed from projection data potentially covering a broad distribution of WEPLs from different rotation angles. To estimate the impact of a given distorted WEPL interval, WEPL heat–maps were calculated, that, for each voxel of the image, indicate the corresponding fraction of WEPLs from a suspected interval. For this purpose, a ground truth RSP map was calculated from experimental data by thresholding a phantom's reconstruction to the nearest true (known) RSP value. This ground truth RSP map was then forward projected in parallel beam geometry to calculate the sinogram $s(\xi, \theta)$, which contains WEPL values. Subsequently, the sinogram was thresholded and a new sinogram s_{thr} was calculated as

$$s_{thr}(\xi, \theta) = \begin{cases} 1 & \text{if } s(\xi, \theta) \in \{w\} \\ 0 & \text{else} \end{cases}, \tag{9.1}$$

where the set $\{w\}$ contained all WEPLs from the suspected intervals. This sinogram was then backprojected according to the same geometry to form the WEPL heatmap h as

$$h(x, y) = \frac{1}{N} \sum_{n=1}^{N} s_{thr}(x \cos \theta_n + y \sin \theta_n, \theta_n), \tag{9.2}$$

where the sum was over $N = 360$ projection angles covering a full rotation. The interpolation that was needed to query the ξ–coordinate was implemented as a nearest neighbor interpolation.

The WEPL heat–maps result in a value of $h(x, y) = 1$ if the WEPL of every ray that intersected the voxel at (x, y) came from the set $\{w\}$ of suspected WEPLs and the corresponding RSP value is expected to be distorted. When $h(x, y) = 0$, none of the corresponding WEPLs was in the interval. For values in between, a relative fraction is calculated, which gives a qualitative idea of the amount of distortion to be expected, but it is not to be taken quantitatively, since eq. (9.2) uses a parallel beam geometry and also disregards the reconstruction filter. Nevertheless, the WEPL heat–maps can give a good first estimate of where in the image distortions are to be expected for certain suspected WEPL intervals.

To demonstrate the use of WEPL heat–maps for the identification of distorted measurements, experimental data of the water phantom acquired for the study [41] presented in section 9.1, was analyzed. Artifacts were compared to a WEPL heat–map for the suspected WEPL intervals at $(36 \pm 3)\,\text{mm}$, $(88 \pm 3)\,\text{mm}$, and $(139 \pm 3)\,\text{mm}$, which correspond to protons stopping in between two stages of the energy detector (see fig. 9.1 and section 3.2.2) and which were suspected to cause distortions. To demonstrate that the reconstruction algorithm would correctly process un–distorted data, an idealized simulation was performed, which did not model any detector effects and thus did not have distorted WEPLs intervals. Idealized simulations do not model the scanner's geometry, but simply score the energy of protons prior and after the object. The two energies can be converted to WEPL according to eq. (3.3).

RESULTS

Figure 9.2 (a) and (b) shows the experimental reconstruction as well as an ideal simulation of the water phantom, for which an RSP of one is expected by definition. While the water

(a) measurement **(b)** simulation (ideal) **(c)** distorted WEPL fraction in percent

C/W = 1.0/0.15   C/W = 1.0/0.15   0   15   30
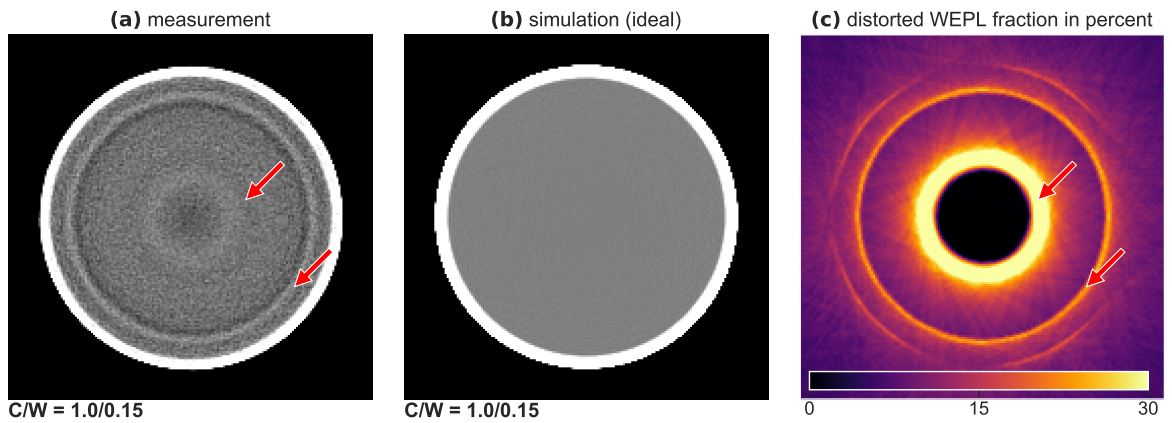
Figure 9.2: RSP reconstructions of the water phantom for (a) experimental data, (b) idealized data simulated without detector effects, and (c) the voxel–wise fraction in percent of WEPLs originating from distorted WEPL intervals close to stage transitions of the pCT scanner. Arrows in (a) and (c) are at the same locations with respect to the phantom.

body is flat as intended for the idealized simulation, the experimental reconstruction is distorted by two ring artifacts, which occur at different radii and which are marked with arrows. For both rings, towards the center of the phantom RSP is decreased and it is increased towards the hull. The location of the two rings is confirmed by the WEPL heat–map in fig. 9.2 (c), where the rings occur at the same locations. An additional third ring is visible, but is located at the PMMA hull of the water and not visible for this grayscale setting. The central ring appears to be stronger in the WEPL heat–map despite the fact that the RSP artifact itself is weaker. While the outer ring agrees precisely with the prediction of the heat–map, the inner ring is slightly closer to the center in the reconstruction.

DISCUSSION

WEPL heat–maps were used to predict the location of distorted RSP images from stage transitions of the prototype scanner's energy detector. They agreed well with ring artifacts in the experimental acquisition, in particular for the outer ring. As expected, heat–maps cannot be taken quantitatively and the weaker ring artifact was represented by a higher fraction of distorted WEPLs in the heat–map. Slight differences in the location of the predicted rings may also be caused from the parallel beam geometry used for the calculation of the heat–maps, which only roughly agrees with actual reconstruction geometry. Nevertheless, the calculation of WEPL heat–maps provides a fast and simple method to investigate the origin of ring artifacts in a pCT image. Dedes et al. [41] used this tool and also included artifacts caused by the geometry of the wedge–shaped calibration phantom in the heat–maps, which was required to reach a better agreement for some phantoms, but are not included here.

9.3 ACCURACY IMPROVEMENT WITH ENERGY−MODULATION

⊟ The results presented in this section were published in *Physica Medica* [197].

MATERIALS & METHODS

9.3.1 Data acquisition

Having concluded that the performance of the prototype pCT scanner was mainly limited by image artifacts that originate from well–defined intervals of WEPLs, a potential improvement in accuracy may be achieved by avoiding such intervals during the acquisition. The predominant origin of distortions was protons that stopped at or close to stage interfaces of the five–stage energy detector [41]. Those are small intervals of WEPL that occur in multiples of 52.9 mm, which is the WEPL of each stage. In this chapter, data was acquired at two incident energies, at 200 MeV and at 187.5 MeV. The range in water of protons at these two energies is 260 mm and 232 mm, respectively [53]. The difference of 28 mm corresponds to approximately half the thickness of a stage plus the wrapping foil. Thereby, a WEPL that makes protons stop at a stage interface for one dataset makes them stop in the

center of the stage for the other dataset. Consequently, the two datasets can be combined to completely cover the phantom while avoiding protons stopping close to stage interfaces. This combined dataset will then produce an image with reduced artifacts. In the future, such energy–modulated proton computed tomography (EMpCT) scans may be acquired using small pencil beams such as for FMpCT. This would allow to scan the phantom while alternating the beam energy such that protons never stop at stage interfaces. However, this initial feasibility study employed twice the imaging dose and performed one complete scan for each energy and merged the data in postprocessing.

Experimental data of the calibration phantom and the water phantom were acquired with a broad beam using the prototype pCT scanner for the two incident energies. The two datasets of the water phantom are referred to as *low–E* (187.5 MeV) and *high–E* (200 MeV) in the following. After data filtration, 98.2 million protons were intersecting the reconstruction volume for the high–E dataset and 99.5 million for the low–E dataset and 360 projections were used for reconstruction. This corresponded to a single projection fluence of $15.7\,\text{mm}^{-2}$ and $15.8\,\text{mm}^{-2}$ respectively and to imaging doses of 0.72 mGy and 0.75 mGy for the high–E and low–E datasets as estimated from corresponding simulations.

9.3.2 Dataset selection

In fig. 9.3 (a) and (b), calibration histograms are shown for the two incident energies. The data for the high–E dataset in (b) agrees with fig. 3.3 and with what is expected for the normal operation of the scanner at 200 MeV. In contrast to that, the low–E data are shifted towards lower WEPLs, since the same energy deposit to the stopping stage requires a lower WEPL (with less energy loss) due to the initially lower energy.

Dashed lines in fig. 9.3 (a) and (b) indicate stage transitions, where, for an increasing WEPL, the energy deposit to a given stage goes to zero and suddenly increases to the maximum energy deposit, but to the previous stopping stage. The WEPLs corresponding to these stage interfaces are listed in table 9.1. The difference between the stage transitions of the two datasets, which is noted in the last column of the table, confirms that they are indeed shifted by approximately half a stage thickness, which is 26 mm.

To avoid protons stopping at stage interfaces to be used in the reconstruction, the two datasets of the water phantom were merged in postprocessing. The selection, whether or not a proton was to be included in the merged dataset, was made based on the WEPL. Figure 9.3 (c) shows calibration curves for both datasets that resulted from the calibration histograms in (a) and (b). The WEPL intervals for each dataset are indicated as solid lines for the calibration curves and as bars on both sides of the plot. To determine the WEPL interval used for the low–E dataset, the WEPLs of the stage transitions in table 9.1 of the high–E dataset were reduced by 5 mm to determine the start of the interval. The end of the interval was then defined as the start of the interval of the high–E dataset, and thus as the stage transition of the low–E dataset, minus 5 mm. For example: the low–E dataset was selected for WEPLs between $36\,\text{mm} - 5\,\text{mm} = 31\,\text{mm}$ and $61\,\text{mm} - 5\,\text{mm} = 56\,\text{mm}$. The high–E dataset was then selected between 56 mm and $88\,\text{mm} - 5\,\text{mm} = 83\,\text{mm}$ and so on. There was no overlap between datasets and the choice of 5 mm was such that the used
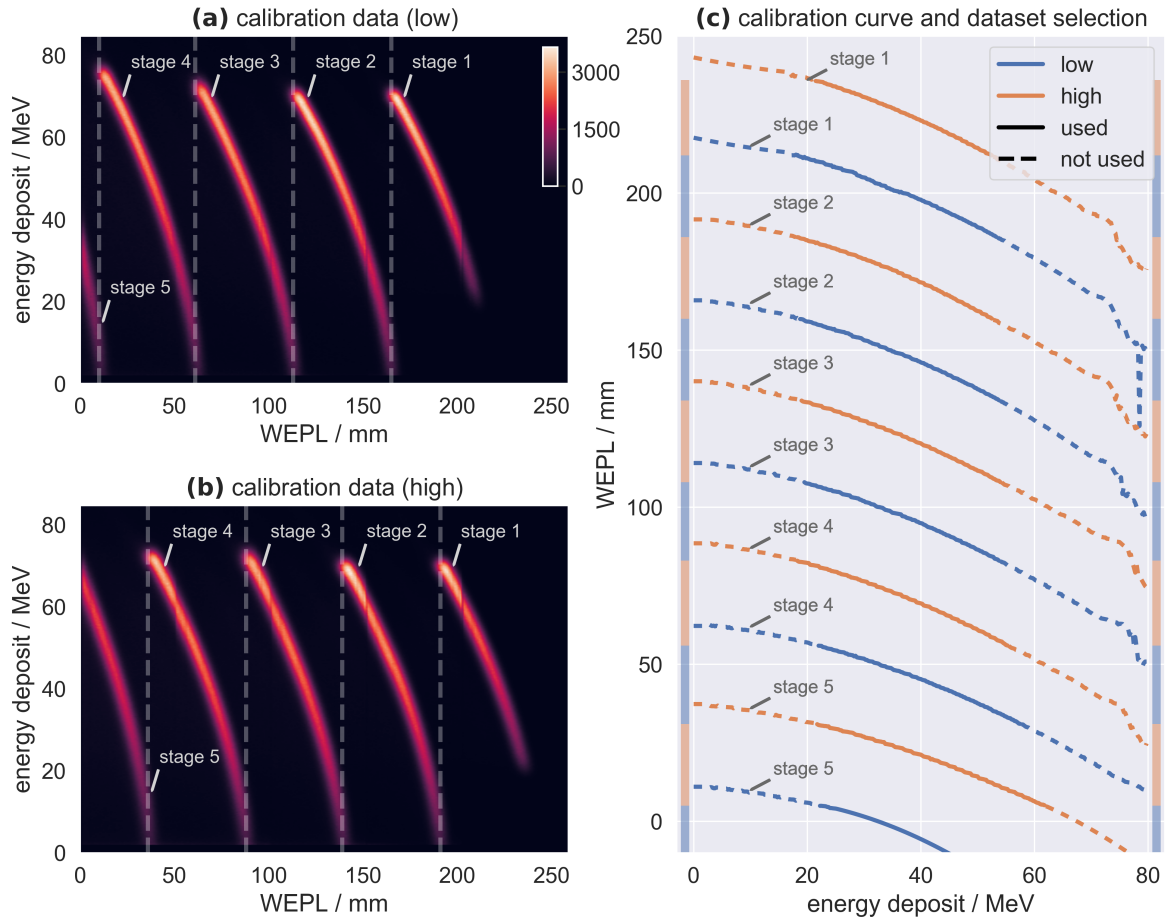
Figure 9.3: Calibration histograms for the (a) low energy and the (b) high energy dataset, and (c) the resulting calibration curves for both datasets with an indication of intervals where the two datasets were used. Dashed lines in (a) and (b) indicate transitions between two stages of the energy detector. The color scale is shared between (a) and (b). In (c), solid lines of the calibration curve and bars on both sides of the plot indicate if a given WEPL was used in the merged dataset and the dashed calibration curve was still used for single–energy datasets. ⎘ Reproduced with permission from Dickmann et al. [197].

intervals were approximately in the middle of each stage in terms of energy deposits, as can be appreciated in fig. 9.3 (c).

Two different strategies were followed when deciding if a given proton was to be included in the merged dataset:

- *dataset selection A:* for each proton, use its calibrated WEPL and check if it falls into the accepted WEPL interval of the corresponding dataset, or

- *dataset selection B:* for each proton, use its tracking coordinates at the front tracker to query the mean WEPL of protons within the same $1\,\mathrm{mm} \times 1\,\mathrm{mm}$ pixel, and decide based on this mean WEPL whether to include the proton or not.

Strategy A may be faster since it only requires knowledge of the individual proton's WEPL, strategy B may be more robust to noise, since it considers the average WEPL of

Table 9.1: Stage interfaces of the high–E and the low–E datasets and corresponding WEPL values. The furthest stage is marked as stage 5 and corresponds to the lowest WEPL. Differences in WEPL between stage transitions are noted in the last column.

stages	dataset	WEPL/mm	difference/mm
$5 \to 4$	low–E	10	—
	high–E	36	26
$4 \to 3$	low–E	61	25
	high–E	88	27
$3 \to 2$	low–E	113	25
	high–E	139	26
$2 \to 1$	low–E	165	26
	high–E	191	26

all surrounding protons. Since WEPL intervals were chosen without overlaps, the merged datasets had the same number of protons compared to the original datasets.

To visualize the dataset distributions within a projection of the water phantom, protons from a single projection were assigned a value of 1 if they came from the low–E dataset and a value of 2 if they came from the high–E dataset. From this data, the average dataset number was calculated in bins of $1\,\text{mm} \times 1\,\text{mm}$ resulting in projection values of 1, if the given pixel was exclusively reconstructed from the low dataset, and 2 if exclusively from the high–dataset. For an exemplary value of 1.2, the fraction of protons from the low–E dataset would be 20 %. This evaluation was also done for two simulated projections of a pediatric head for the data of patient 1 in chapter 8.

9.3.3 Evaluation of energy–modulated scans

Images of the water phantom were reconstructed for the two initial datasets as well as for two merged datasets with dataset selection A and B. To evaluate the accuracy of these two scans, for each voxel in the image, the distance to the rotation axis was calculated. Then, the mean RSP value was calculated in radial bins, since distortions for the cylindrical water phantom were expected to be ring–shaped. The radial bins were chosen such that the spacing was unequal in terms of the radius, but each annulus covered a constant area of $A = 500\,\text{mm}^2$. This was achieved by rounding the voxels' radii to the nearest radius of the set $\{r_n\}$ with

$$r_0 = 0\,\text{mm}, \tag{9.3}$$

$$r_1 = \sqrt{A/\pi}, \tag{9.4}$$

$$r_{n+1} = \sqrt{r_1^2 + r_n^2}. \tag{9.5}$$

With this approach, the mean in each annulus was calculated over the same number of voxels and, therefore, with a comparable uncertainty. In each radial bin, the mean RSP value was calculated as a function of the radius. In addition, the median RSP of all bins as well as

the inter–quartile range (IQR) from the 25th to the 75th quartile of all mean RSP values was calculated. The median RSP represents the average accuracy and the IQR is representative of the magnitude of rings present in the image.

RESULTS

Figure 9.4 (a) shows a projection of the water phantom for the high–E dataset, where the top and bottom lids as well as the phantom holder can be seen in contrast to the central water region. At the top edge of the water region, a small air bubble reduces the WEPL. In fig. 9.4 (b), the corresponding average dataset number for dataset selection A is shown, where the dataset was selected based on each proton's WEPL. While certain regions with predominant datasets can be identified, most pixels have average dataset numbers between 1 and 2, which indicates a relevant contribution from both datasets. In general, the dataset map is noisy. Figure 9.4 (d) shows the same evaluation for dataset selection B, where the average WEPL was used to choose the dataset. This map is much less noisy and most pixels are at the values 1 or 2, except for regions close to the edges of the phantom. Those merged regions occur because the dataset was selected at the front tracker at $d = -167\,\mathrm{mm}$, while the data shown here is at the isocenter at $d = 0\,\mathrm{mm}$. The regions which are scanned with a homogeneous energy dataset are approximately $10\,\mathrm{mm}$ wide, except for the very edge of the phantom, where both datasets are used even for dataset selection B. Figure 9.4 (c) shows a histogram of WEPLs in the projection of the water phantom, where two peaks can be identified: one peak around $150\,\mathrm{mm}$ corresponding to the largest WEPL observed in the water region and a second peak around $170\,\mathrm{mm}$ corresponding to WEPLs at the lids and the holder, which are made from a higher RSP material. Colors in the stacked histogram
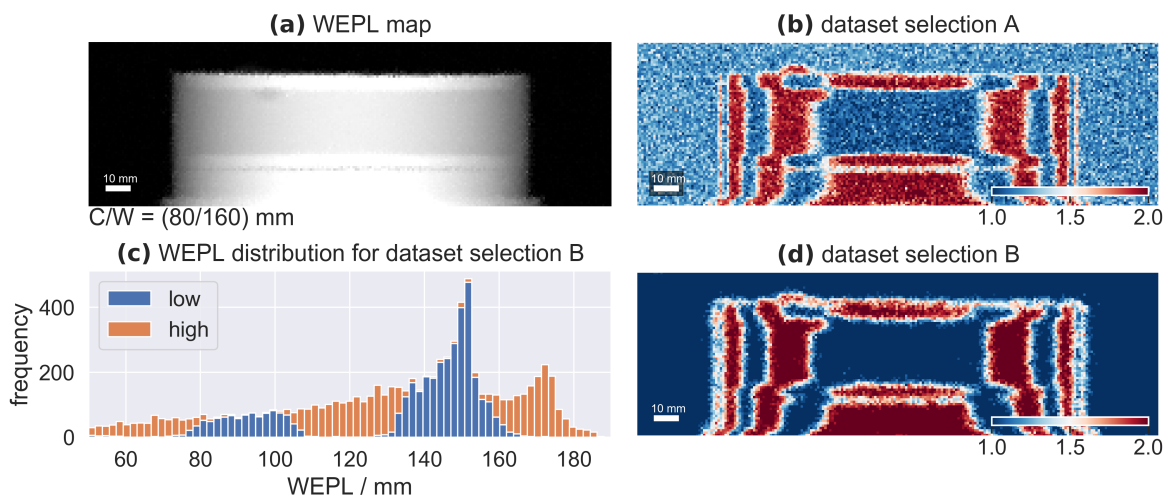
Figure 9.4: Spatial distribution of the two energies in the energy–modulated scan in one projection of the water phantom: (a) WEPL map of the high–E dataset, (b,d) average dataset number for selection strategy A and B, and (c) a histogram of WEPLs in the projection with the corresponding dataset selection. All data are shown at the isocenter binning depth $d = 0\,\mathrm{mm}$. ⏃ Reproduced with permission from Dickmann et al. [197].
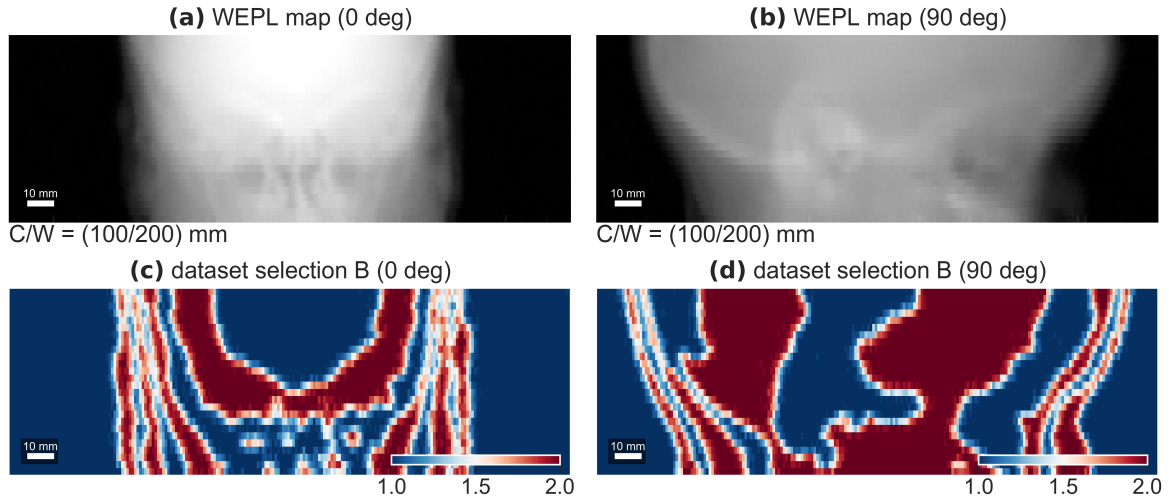
Figure 9.5: Spatial distribution of the two energies in the energy–modulated scan in one projection of a pediatric head: (a,b) WEPL map at a rotation angle of 0 and 90 degrees, (c,d) the corresponding dataset selection for selection strategy B. Reproduced with permission from Dickmann et al. [197].

indicate the distribution of WEPLs to the two datasets. The transition between two datasets spans approximately a WEPL range of 5 mm.

Figure 9.5 shows simulated WEPL maps as well as dataset maps for two projections of a pediatric head scan and dataset selection B. While in particular in fig. 9.5 (c) around the nasal cavity in the center, smaller features occur in the dataset map, most regions of uniform dataset number are 10 mm wide or larger.

Figure 9.6 (a) and (b) shows reconstructions of the experimental data of the water phantom for the low–E and high–E dataset, and in (c) and (d) for the merged dataset using selection strategies A and B. The original datasets at the two energies show ring artifacts at different radii. Artifacts for the high–E dataset are stronger than those for the low–E dataset. This is confirmed in the corresponding radial profiles in fig. 9.6 (e), where the modulation seen for the high–E dataset is more prominent. Both datasets suffer from an offset of approximately 1 % compared to the expected RSP for water of one. In particular the central radial bin is affected by a negative offset.

The merged datasets in fig. 9.6 (c) and (d) as well as the corresponding line profiles in (f) show considerably less rings, in particular when compared to the high–E dataset, which is the current standard acquisition. In the profiles, data selection strategy B is slightly flatter compared to strategy A, which can also be seen visually. Both merged datasets are also offset by 1 % and their central values are further reduced.

In table 9.2 the median RSP and the IQR from the 25th to the 75th percentile are listed for all four datasets. The median values reflect the 1 % bias, that was already seen in the profile plots. The low–E dataset as well as the merged dataset with energy modulation B are closest to the RSP of water and therefore produce the best results in terms of accuracy. The IQR is highest in the high–E dataset, but also elevated in the low–E dataset. The energy–modulated scan with selection B has the smallest IQR, which is three– to four–fold better than the high–E dataset and about two–fold better than the low–E dataset.
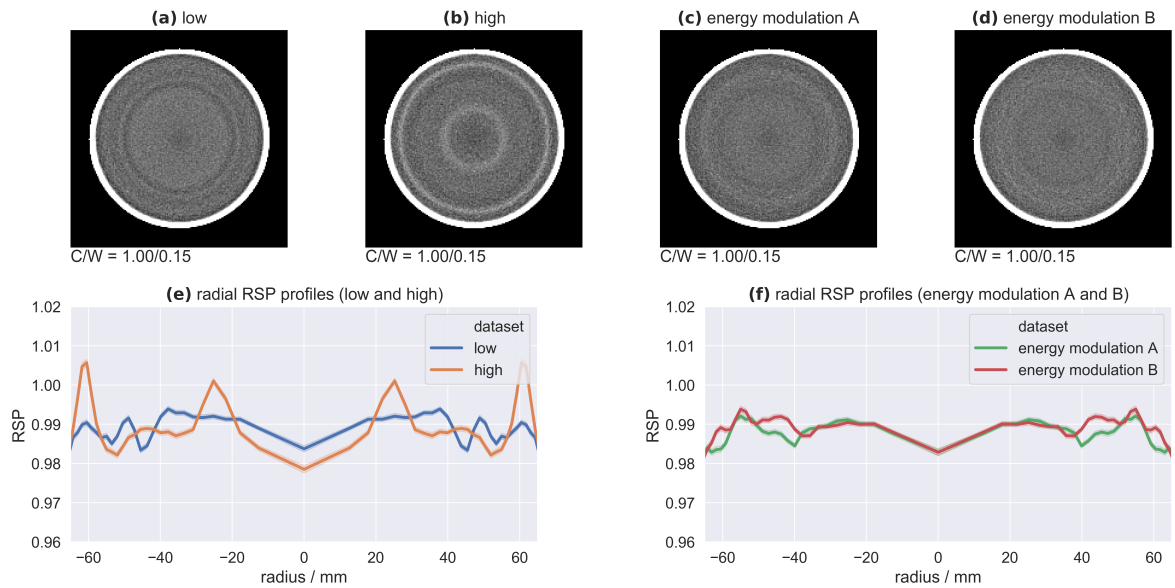
Figure 9.6: Reconstructions of the water phantom for the (a) low and the (b) high energy dataset as well as (c,d) for the two merged datasets. Radial profiles through (a,b) are shown in (e) and through (c,d) are shown in (f). ⎘ Reproduced with permission from Dickmann et al. [197].

Table 9.2: Evaluation of the median and the inter–quartile range (from the 25th to the 75th quartile) of all radial bins of the water phantom for the high–E, the low–E and both merged datasets. For both quantities the best value among the four datasets is indicated.

dataset	median	IQR (25/75)
low–E	0.990*	0.5 %
high–E	0.988	0.7 %
energy modulation A	0.988	0.4 %
energy modulation B	0.990*	0.2 %*

* best value among the four datasets

DISCUSSION

Ring artifacts stemming from detection inaccuracies of protons stopping close to stage interfaces were considerably reduced in experimental acquisitions by merging the proton–by–proton data of two scans at different energies in an EMpCT scan. The energies were chosen such that their corresponding ranges differed by half the thickness of a stage, thereby allowing a selection of protons that completely avoided stage interfaces.

While the design of the experiment as used here employed double the imaging dose compared to a normal scan, it may be feasible in the future to acquire EMpCT scans with small pencil beams with alternating energy avoiding to scan the patient twice at both energies. Were an object to be scanned with energy–modulated pencil beams, its RSP map would need to be known before the initial scan to plan the EMpCT acquisition using optimization, similar to FMpCT. This RSP map may come from a prior x–ray CT scan and EMpCT scans would be possible at little additional imaging dose. If such scan is not available, a full pCT

scan could be performed at one energy, which would allow to guide the acquisition of the second energy only in regions where this is necessary to avoid artifacts. In that case, the imaging dose would only be elevated by about 50 %.

Calculation of dataset maps for a projection of the water phantom and two projections of a pediatric head scan have shown that modulation features are approximately of the size of 10 mm or above for most of the phantom. This would allow them to be scanned with energy–modulated pencil beams of the extent determined in chapter 7, which was between 6 mm and 9 mm (FWHM). At the transition regions as well as for very heterogeneous regions such as the nasal cavity, more than one energy may be needed allowing to reject some protons, which in turn would increase imaging dose. In principle, the use of energy–modulation with more than two energies would allow to avoid the five–stage design of the energy detector and directly build a monolithic detector with only one stage as done by DeJongh et al. [89] and Ordoñez et al. [200]. This would additionally reduce image noise due to energy straggling in the detector (see chapter 5), but may require longer scan times.

The comparison of the two dataset selection strategies showed that selection A, where the selection of protons was made based on their individual WEPL was noisier in the dataset maps and also resulted in worse results in terms of image accuracy. Acquisition of such selection maps with small energy–modulated pencil beams would not be achievable due to the mixing of datasets throughout the projection. On the other hand, dataset selection B achieved better results in terms of RSP accuracy and also had clearly separated dataset maps, which would be achievable with small pencil beams in most regions.

The acquisition of the low–E scan showed less severe image artifacts compared to the high–E scan, which is the typical acquisition energy. Since the calibration phantom was not moved in between the calibrations at the two energies, this difference cannot be explained by a changing quality of the two calibrations. It may, however, be explained by the dimension of the water phantom, for which the most frequent WEPLs are between 140 mm and 155 mm, which is close to a stage interface for the high–E dataset and further away from an interface for the low–E dataset. In addition, the low–E dataset only experienced two stage transitions over the complete WEPL range of the water phantom, while for the high–E dataset this was three. As shown by Dedes et al. [41], a different phantom size can indeed have an impact on the magnitude of artifacts observed.

The merged datasets, and in particular the datasets with selection strategy B had a clear visual reduction of ring artifacts compared to both single–energy datasets. All datasets suffered from a bias of 1 %, which may be caused by wrong assumptions on the RSP of the wedge–shaped calibration phantom or other inaccuracies in the calibration. It is unlikely to be caused by the reconstruction algorithm, which gives unbiased reconstructions for ideal datasets [113]. The IQR, and therefore the magnitude of ring artifacts, was reduced considerably using energy–modulation B. The median error with EMpCT was also the lowest among all datasets using selection strategy B, but the accuracy is still to be improved given the offset of 1 %, which could by design not be addressed by EMpCT, since it merges two datasets that both suffer from the bias.

9.4 ACCURACY IMPROVEMENT WITH AN EMPIRICAL CORRECTION

▣ The results presented in this section were submitted for publication [141].

9.4.1 An empirical correction method for proton CT

The EMpCT acquisitions in the previous section allowed to reduce ring artifacts in pCT images by avoiding the physical reason for their occurrence, which were protons stopping close to stage interfaces. However, the method did not completely remove artifacts and failed to correct for a systematic bias, that was present in the original data. In this section, an empirical correction method is developed that makes use of a tomographic scan of a phantom with known RSP to isolate and quantify WEPL distortions. With the distortion known, a correction function can be found, that removes artifacts from subsequent experimental scans. The proposed method is based on the empirical cupping correction of Kachelrieß, Sourbelle, and Kalender [201], which was developed to reduce cupping artifacts in x–ray CT images, and which was adapted here to the requirements of pCT. Unlike EMpCT, it makes no assumption on the origin of artifacts and is purely empirical and based on the scan of a custom–made elliptical phantom of known RSP, which was constructed specifically for this work.

A schematic workflow of the proposed method is shown in fig. 9.7 and will be detailed in the following. The notation in this section is adapted to the one of the original publication [201]. Let $q(u, v)$ be the measured WEPL projection values at the isocenter binning depth $d = 0\,\mathrm{mm}$. In the following, the dependence on (u, v) will be dropped to simplify the notation. Instead of using the FDK–DDB algorithm, a simple filtered backprojection
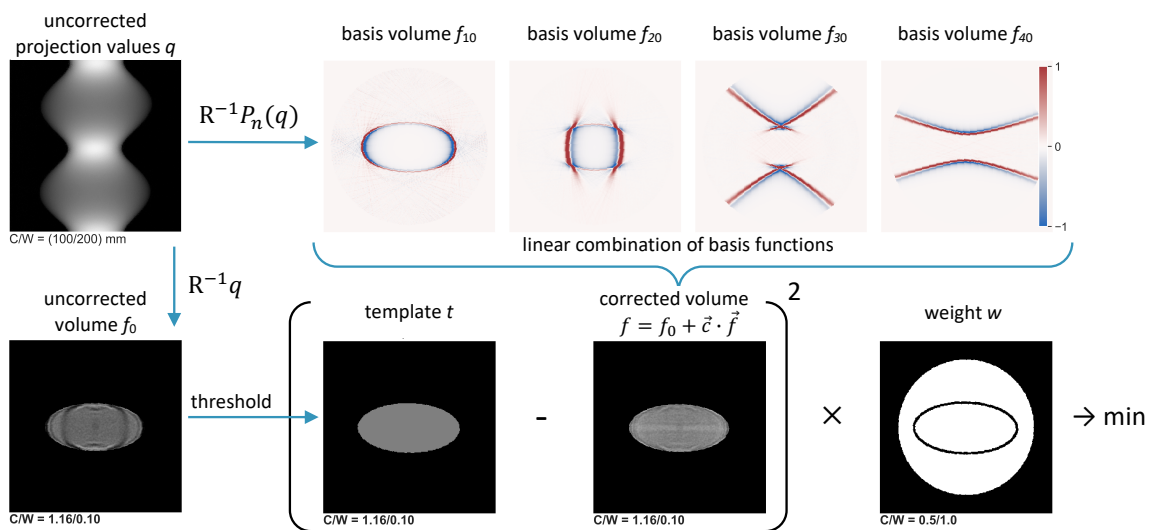
Figure 9.7: Schematic for the optimization of the correction function with examples of all quantities of the algorithm. The basis volumes f_n are normalized to their maximum.

is applied to reconstruct an image f from the projection values q, which allows to reduce the computational cost of the method. A correction function P is defined, which calculates corrected projection values $p = P(q)$ from the measured ones. It is assumed, that P can be expressed as a linear combination of $N = 47$ basis functions $P_n(q)$ as

$$P(q) = q + \sum_{n=1}^{N} c_n P_n(q) = q + \vec{c} \cdot \vec{P}(q), \tag{9.6}$$

where the coefficient vector $\vec{c}$ needs to be determined. An additional term q is added before the sum compared to the original method, such that the weighted sum of basis functions $\vec{c} \cdot \vec{P}(q)$ only determines the WEPL error $P(q) - q$. In the original method, this was not necessary, since q was one of the basis functions. The choice of N is motivated later.

Exploiting the linearity of the inverse Radon transform R^{-1}, which denotes the reconstruction operation, the corrected image $f(\vec{r})$ with $\vec{r} = (x, y, z)$ can be expressed as a linear combination of basis images $f_n(\vec{r})$, which can be calculated as

$$f_n(\vec{r}) = R^{-1} P_n(q), \tag{9.7}$$

with $1 \leqslant n \leqslant N$. The corrected image then is

$$f(\vec{r}) = R^{-1} p = R^{-1} q + \sum_{n=1}^{N} c_n f_n(\vec{r}) = f_0(\vec{r}) + \vec{c} \cdot \vec{f}(\vec{r}), \tag{9.8}$$

were $f_0 = R^{-1} q$ is the uncorrected image. All reconstructions were performed using the FDK algorithm for distance–driven binned data at the central binning depth and using the *Reconstruction Toolkit* of Rit et al. [122].

The basis functions were not chosen to be monomials as in the original publication [201], but instead radial basis functions were used, that have been employed in the context of machine learning to approximate arbitrary functional dependencies [202]. The basis functions were Gaussian functions with a small standard deviation $\sigma = 2\,\text{mm}$ and equally spaced peaks, which were defined as

$$P_n(q) = A \cdot \exp\left[-\left(\frac{q - (n-1) \cdot s - s_0}{\sqrt{2}\sigma}\right)^2\right], \tag{9.9}$$

with an amplitude $A = 1\,\text{mm}$, a step size of $s = 2\sigma = 4\,\text{mm}$ and a minimum WEPL of $s_0 = 8\,\text{mm}$. Since each basis function had an amplitude of $1\,\text{mm}$, setting the weight c_n of basis function n to, for example, $c_n = 2$, would result in the corresponding WEPL interval around $(n-1) \cdot s + s_0$ being shifted up by $2\,\text{mm}$.

The coefficients $\vec{c}$ were found by assuming that a ground truth RSP value of a given homogeneous correction phantom is known and a template $t(\vec{r})$ can be obtained by thresholding $f_0(\vec{r})$ to the true RSP value inside the phantom or zero outside the phantom. This makes no assumptions on the actual shape or location of the correction phantom, and only its RSP must be known. With this, the cost function

$$E^2 = \int d^3 r \, w(\vec{r}) \, (f(\vec{r}) - t(\vec{r}))^2 \tag{9.10}$$

is to be minimized, where $w(\vec{r})$ is a spatial weight, which is set to zero at the edges of the correction phantom. This avoids distortions of the method due to the limited spatial resolution of the scanner. As suggested in the original publication [201], the weights can be found by solving the system of linear equations $\vec{a} = \mathbf{B} \cdot \vec{c}$ for

$$a_i = \int d^3 r\, w(\vec{r}) f_i(\vec{r})(t(\vec{r}) - f_0(\vec{r})), \qquad (9.11)$$

$$B_{ij} = \int d^3 r\, w(\vec{r}) f_i(\vec{r}) f_j(\vec{r}), \qquad (9.12)$$

where $1 \leqslant i, j \leqslant N$. Equation (9.11) has an additional term $f_0(\vec{r})$ here, due to the explicit inclusion of q in eq. (9.6). The system was solved using the trust region reflective algorithm with an implementation of the *SciPy* library [203].

Subsequent scans were then corrected by passing each of the WEPL values of the three–dimensional distance–driven binned projections through the correction function $P(q)$. The projections were then reconstructed using the FDK–DDB algorithm.

The previous assumption that the correction phantom consists of a single material makes the method simpler. However, this is not strictly required, since the template $t(\vec{r})$ can represent any number and distribution of known RSP values as long as they can be clearly identified by thresholding.

9.4.2 Evaluation of corrected proton CT scans

Experimental scans were performed using the prototype pCT scanner and with the water phantom, the CTP phantom and the pediatric head phantom. Scans were performed at a proton energy of 200 MeV and at 187.5 MeV, which effectively doubled the amount of test cases since artifacts are expected to be different at different incident energies, as shown in section 9.3. Scans at two energies, however, are not required for the proposed method to work and most likely the method would be applied to a single energy only in the future. To apply the correction method, the custom–made elliptical phantom shown in fig. 3.6 (d) was also scanned at both energies. The correction phantom consisted of PMMA and had a height of 80 mm and diameters of 165 mm and 80 mm. The flat surface of the phantom was milled to a flatness of better than 0.02 mm allowing for a determination of the phantom's RSP as

$$RSP_{\text{ref.}} = 1.160 \pm 0.001 \qquad (9.13)$$

using a multi–layer ionization chamber and protons with an incident energy of 150 MeV. The RSP was determined and averaged at five distinct locations across the flat surface.

The elliptical shape of the phantom was chosen to have a smooth outer hull while covering the widest possible WEPL range down to even small WEPL values. The phantom's maximum WEPL was 191 mm, and thus larger than the maximum WEPL of all other phantoms, which were 152 mm for the water phantom, 173 mm for the CTP phantom, and approximately 176 mm for the head phantom. This also justifies the choice of $N = 47$ basis functions, which cover WEPLs from $s_{\text{min}} = s_0 = 8$ mm up to $s_{\text{max}} = (N-1) \cdot s + s_0 = 192$ mm, which is 1 mm more than the maximum WEPL of the correction function. The minimum WEPL was chosen

according to the smallest non–zero WEPLs observed in the ellipse data, which is about $8\,mm$ due to the pixel size of the projections.

The performance of the proposed correction method was evaluated by comparing corrected and uncorrected reconstructions of the water and the ellipse phantom while averaging all homogeneous slices of the two phantoms to reduce noise. For the CTP phantom, the MAPE was calculated over all inserts and the body similar to chapter 7. The mean RSP for each insert was determined within a radius of $3\,mm$ around its center, which was $50\,\%$ of the insert radius. The RSP of the body was calculated for a central circular ROI with a radius of $10\,mm$. The mean RSP values were then compared to the reference values $RSP_{ref.}$, resulting in the relative error $(RSP_{mean} - RSP_{ref.})/RSP_{ref.}$. This allowed to calculate the MAPE from all M inserts as

$$\text{MAPE} = \frac{100\,\%}{M} \cdot \sum_{m=1}^{M} \frac{|RSP_{mean,m} - RSP_{ref.,m}|}{RSP_{ref.,m}}, \tag{9.14}$$

where the body of the CTP phantom and the water phantom were also considered as one of the M values. Results of the accuracy analysis were put into perspective to results determined by Dedes et al. [41], where the center of the CTP phantom as well as the accuracy of the water phantom were re–evaluated for this study, since they were not reported there. The MAPE reported here also differs from the original publication, since it contains values only from one phantom, whereas the original publications considered more than one phantom.

Since the proposed method requires an additional scan for each beamtime, it was tested if the application of the correction makes the original calibration omissible. For this, data of the CTP phantom were calibrated using a calibration run acquired 15 months before the data acquisition for this study. Subsequently the correction method was applied using data of the ellipse phantom with the same, old calibration. Image accuracy was evaluated as described before and compared to a run using the calibration of the day and the proposed method.

For the scan of the head phantom no ground truth RSP values were known, but the phantom was expected to consist of piece–wise constant materials. Images were, therefore, inspected visually, and, to isolate the effect of the correction function, difference images between corrected and uncorrected scans were calculated.

RESULTS

Figure 9.8 displays the two correction functions that were found using the proposed method for the two datasets at an initial energy of $200\,MeV$ and $187.5\,MeV$. The functions represent the WEPL error $P(q) - q = \vec{c} \cdot \vec{P}(q)$ as a function of the WEPL q. Both functions were strictly positive, which was not enforced by the optimization, and appear to follow the same linear trend with a positive slope of approximately $1.3\,\%$. They have a common feature at around $130\,mm$, where both corrections agree precisely. At around $40\,mm$ and $70\,mm$ for the $200\,MeV$ function and at $80\,mm$ for the $187.5\,MeV$ function distinct features were observed, where different corrections were applied to the two datasets. Since there was no basis function support below $8\,mm$ and above $192\,mm$, both correction functions drop to zero outside of this interval.

In fig. 9.9 (a) and (b) uncorrected and corrected scans of the ellipse phantom are shown for both incident energies. The uncorrected scans, based on which the correction function was
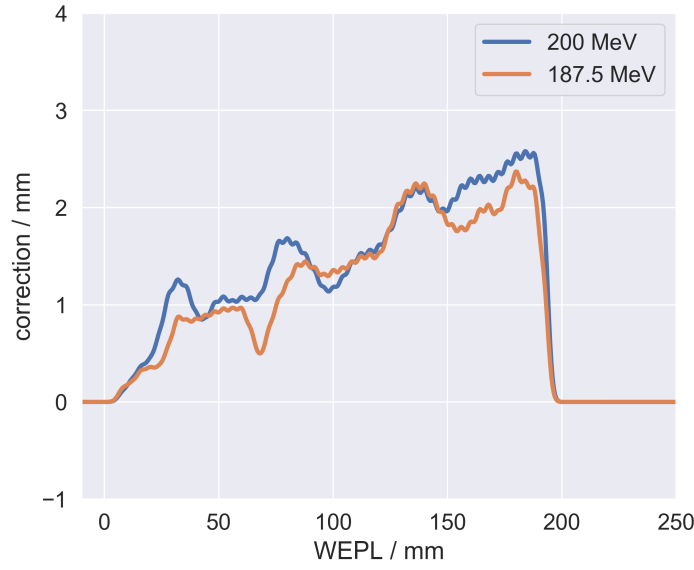
Figure 9.8: Optimized correction functions $P(q) - q = \vec{c} \cdot \vec{P}(q)$ for the high and low energy dataset, where q is the WEPL.

found exhibited clear artifacts, which had hyperbolic shapes. These artifacts were greatly reduced, when the correction function was applied to the ellipse data as a consistency check. A central dip remained in the center as well as minor artifacts in particular in the dataset at $200\,MeV$. These observations are confirmed in the last row of the figure, where line profiles are shown which were taken horizontally through the center of each phantom. Uncorrected scans of the ellipse phantom at both energies had a bias which was outside of the $\pm 1\,\%$ band shown as a shaded green area. Using the correction, this offset was removed and the scans were on average at the reference value. The magnitude of ring artifacts as seen in the profiles, was also considerably reduced.

Figure 9.9 (c) and (d) shows the same evaluation for the water phantom and using the correction function that was found based on the ellipse data. Ring artifacts in the uncorrected scans were strong, which can be seen both visually and in the line profiles. A similar negative bias as for the uncorrected ellipse scans was observed. The magnitude of maximum to minimum RSP value within the profiles was slightly above $\pm 2\,\%$. Application of the correction function reduced the ring artifacts substantially, with only subtle modulations remaining. In the line profiles, all RSP values were within the $\pm 1\,\%$ band around the reference RSP and their magnitude was reduced considerably.

The mean RSP in the ellipse phantom and the water phantom for both energies are reported in table 9.3. The mean values across the entire phantom body confirmed the negative bias of uncorrected scans that was already seen in the line profiles. Application of the correction substantially improved the accuracy and reduces the bias. Results for the ellipse phantom were slightly better than those for the water phantom, but all relative errors of corrected scans were below $0.3\,\%$.

Figure 9.10 displays the scan of the CTP phantom at $200\,MeV$ without any correction in (a), with the proposed correction method in (b), and in (c) the scan of the study of Dedes et al. [41]
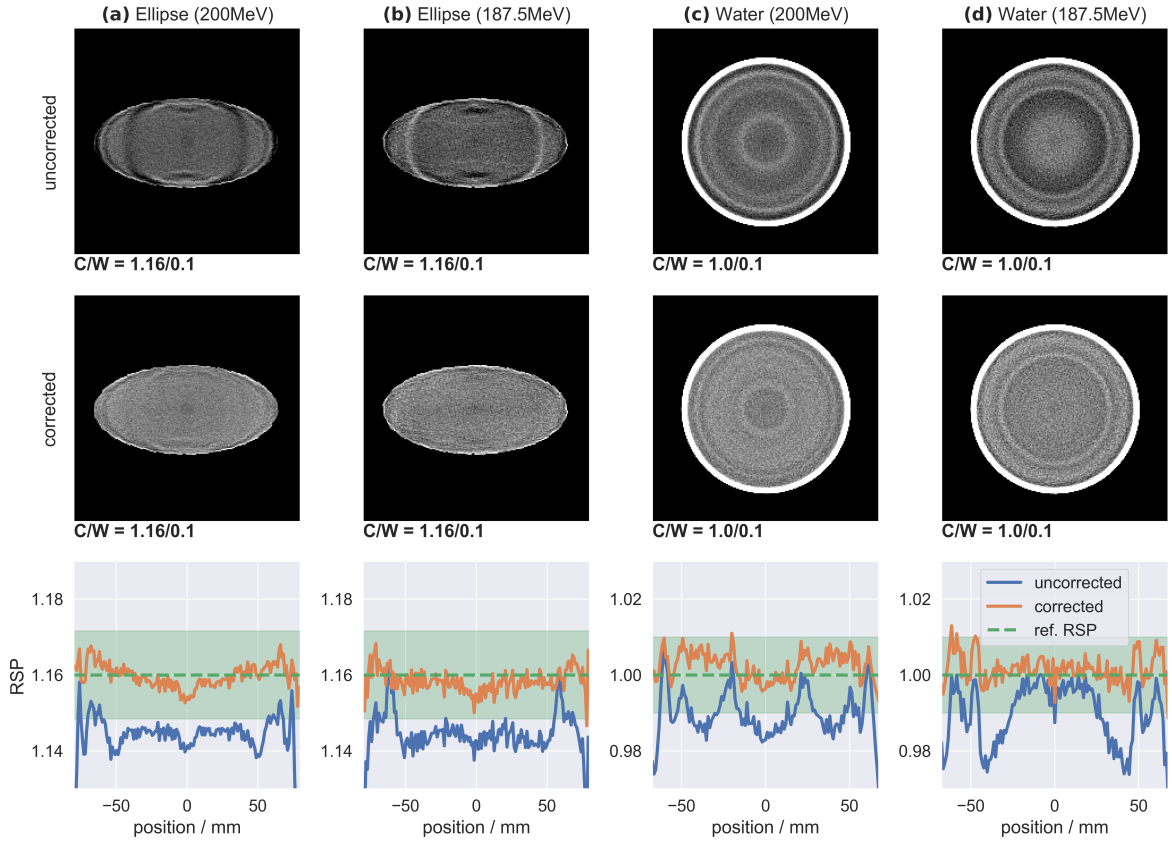
Figure 9.9: Corrected and uncorrected scans of (a,b) the ellipse and (c,d) the water phantom at both energies. The last row shows horizontal line profiles along the center of the phantom together with the reference RSP and a shaded area indicating a ±1 % interval around that value.

is shown again as a comparison. Artifacts visible in (a) and (c) were at similar locations and of similar magnitude. They were not ring–shaped due to the heterogeneity of the phantom as discussed in Dedes et al. [41]. The corrected data in (b) was visibly more homogeneous and artifacts were reduced considerably. The RSP was also at a higher level compared to (a) and (c).

In fig. 9.10 (d), the relative RSP errors of all inserts and the body are shown for the three scans in (a) – (c). For the two uncorrected scans, errors were very similar and negative, except for the PMP insert. The inserts are ordered by RSP and there was a tendency of a stronger underestimation for higher RSP values. The relative errors of corrected scans were in general closer to zero and all were within the ±1 % interval, except for the PMP insert, for which the lower uncertainty bound overlapped with the interval. While in general, errors were reduced, the trend of overestimating low–RSP inserts and underestimating high–RSP insets remained.

Table 9.4 lists relative errors for all inserts of the CTP phantom and the body of the water phantom both for uncorrected and corrected scans at 200 MeV and 187.5 MeV. In addition, the results of Dedes et al. [41] are reported in the last column, where the body of the CTP

Table 9.3: Mean RSP values with and without the empirical correction for the water and the ellipse phantom.

phantom	RSP$_{ref.}$	energy / MeV	uncorrected		corrected	
			RSP	rel. error / %	RSP	rel. error / %
ellipse	1.160	200	1.145	−1.3	1.160	0.0
		187.5	1.145	−1.3	1.159	0.1
water	1.000	200	0.990	−1.0	1.003	0.3
		187.5	0.988	−1.2	1.002	0.2

Table 9.4: Relative RSP errors of all inserts and the body of the CTP phantom at the two incident energies. In the last column relative errors from a previous study [41] are reported, which were re–evaluated for this study. In the last row, the mean absolute percentage error (MAPE) is calculated for all errors reported in the table. *Abbreviations:* PMP – polymethylpentene, LDPE – low–density polyethylene.

rel. error in %		uncorrected error		corrected error		Dedes [41]
material	RSP$_{ref.}$	200 MeV	187.5 MeV	200 MeV	187.5 MeV	200 MeV
CTP phantom						
PMP	0.883	0.19 ± 0.10	-0.19 ± 0.12	1.16 ± 0.10	1.23 ± 0.12	1.08 ± 0.11
LDPE	0.979	-0.25 ± 0.08	-0.22 ± 0.10	0.87 ± 0.08	0.99 ± 0.11	-0.49 ± 0.11
polyst.	1.024	0.55 ± 0.08	-0.89 ± 0.10	0.32 ± 0.08	0.29 ± 0.10	0.04 ± 0.10
epoxy	1.144	-1.33 ± 0.00	-0.55 ± 0.00	-0.06 ± 0.00	0.37 ± 0.00	$-1.20 \pm 0.00*$
acrylic	1.160	-0.81 ± 0.07	-0.96 ± 0.10	0.29 ± 0.07	0.24 ± 0.09	-0.30 ± 0.10
Delrin	1.359	-1.36 ± 0.06	-1.22 ± 0.08	0.06 ± 0.06	0.07 ± 0.08	-1.16 ± 0.09
Teflon	1.790	-1.56 ± 0.05	-1.66 ± 0.06	-0.38 ± 0.05	-0.49 ± 0.07	-1.31 ± 0.05
water phantom						
water	1.000	-0.99 ± 0.00	-1.15 ± 0.00	0.35 ± 0.00	0.18 ± 0.00	$-0.22 \pm 0.00*$
MAPE		0.87 ± 0.02	0.86 ± 0.03	0.44 ± 0.02	0.48 ± 0.03	0.72 ± 0.03

*data were re–evaluated for this study

phantom (epoxy) and the water phantom were re–evaluated for this study. The reference RSP is given for each insert and inserts are ordered by RSP.

The MAPE was calculated over all values listed in the table and was at 0.87 ± 0.02 and 0.86 ± 0.03 for the two uncorrected scans. The re–evaluated MAPE of the previous study was calculated to be 0.72 ± 0.03, which is different from the value reported in the original publication, since here only the CTP and water phantom were considered. Using the proposed correction method, the MAPE was reduced considerably to 0.44 ± 0.02 for the 200 MeV dataset and 0.48 ± 0.03 for the 187.5 MeV dataset. This is a reduction of 49 % and 44 % compared to the corresponding uncorrected scans.

Figure 9.11 shows the same evaluation as in fig. 9.10, but for the scan of the CTP phantom, that was calibrated using a 15 months old calibration. The corresponding uncorrected image in (a) was consequently distorted and the under–representations in (d) were larger than 2 %. Application of the proposed method to this distorted dataset in (b) restored the accuracy and
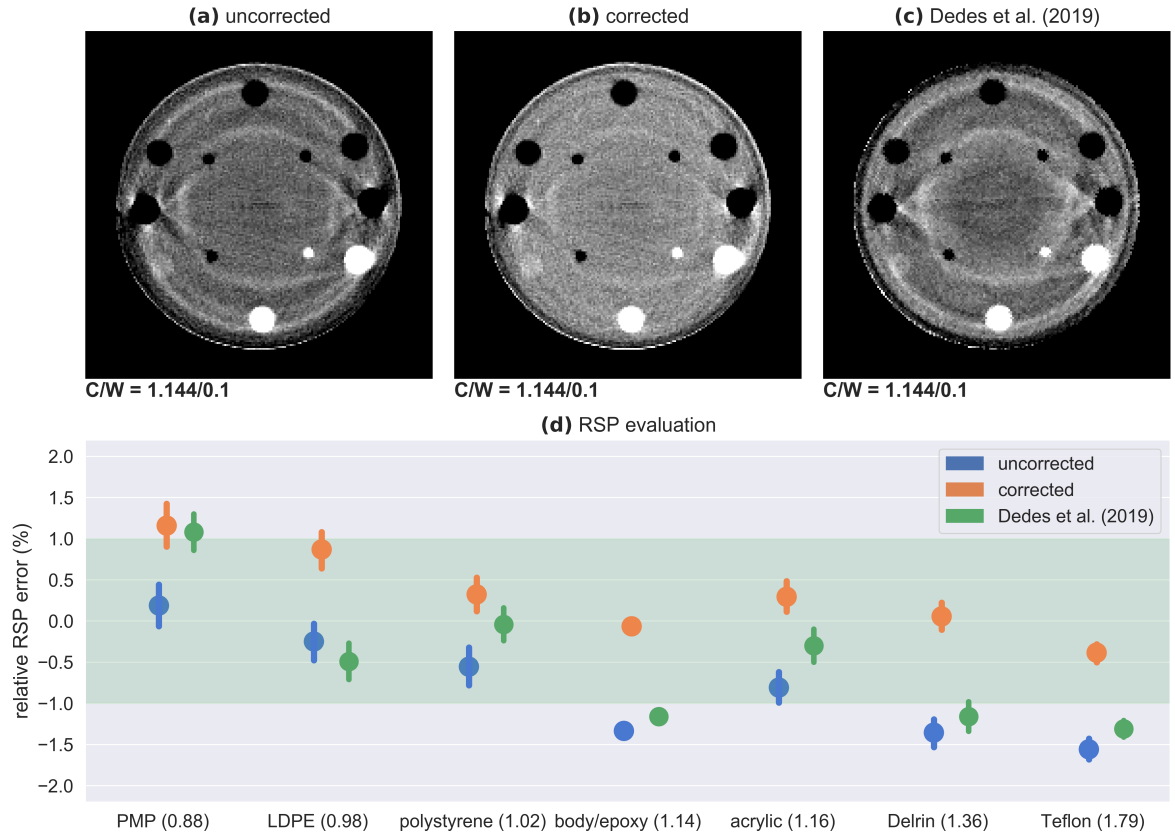
Figure 9.10: Evaluation of the RSP accuracy of the (a) uncorrected and (b) corrected CTP phantom acquired at $200\,\mathrm{MeV}$ and (c) comparison to the data of Dedes et al. [41]; (d) the relative RSP error for all inserts and datasets with the green area indicating a $\pm1\,\%$ interval. The RSP of inserts is given in brackets behind the name and inserts are sorted by RSP.

reduced ring artifacts. For comparison, in (c) the scan with the correct calibration and the proposed method is shown and in (d) little differences were observed between the accuracy of (b) and (c). The MAPE of the CTP phantom for the uncorrected scan was $(1.94 \pm 0.03)\,\%$. Using the proposed method this was reduced to $(0.32 \pm 0.03)\,\%$, which is even slightly better than the accuracy of the scan using the correct calibration and the proposed method. However, accuracy values between the two corrected scans overlapped for all inserts within the uncertainty interval.

In fig. 9.12 reconstructed data for the head phantom at both incident energies is shown with and without the correction. In addition, difference maps were calculated between the corrected and the uncorrected scans. Arrows indicate locations of visible artifacts and are repeated at the same locations for the corrected scan and the difference map. For the soft–tissue region of the phantom, the material was constant and no modulation is expected. Modulations in the uncorrected scan could, therefore, be attributed to artifacts. For all marked locations, the artifacts reduced considerably or were completely removed using the proposed method. In the difference maps, a general underestimation could be observed for the uncorrected phantom. Artifacts that were removed using the proposed method were in
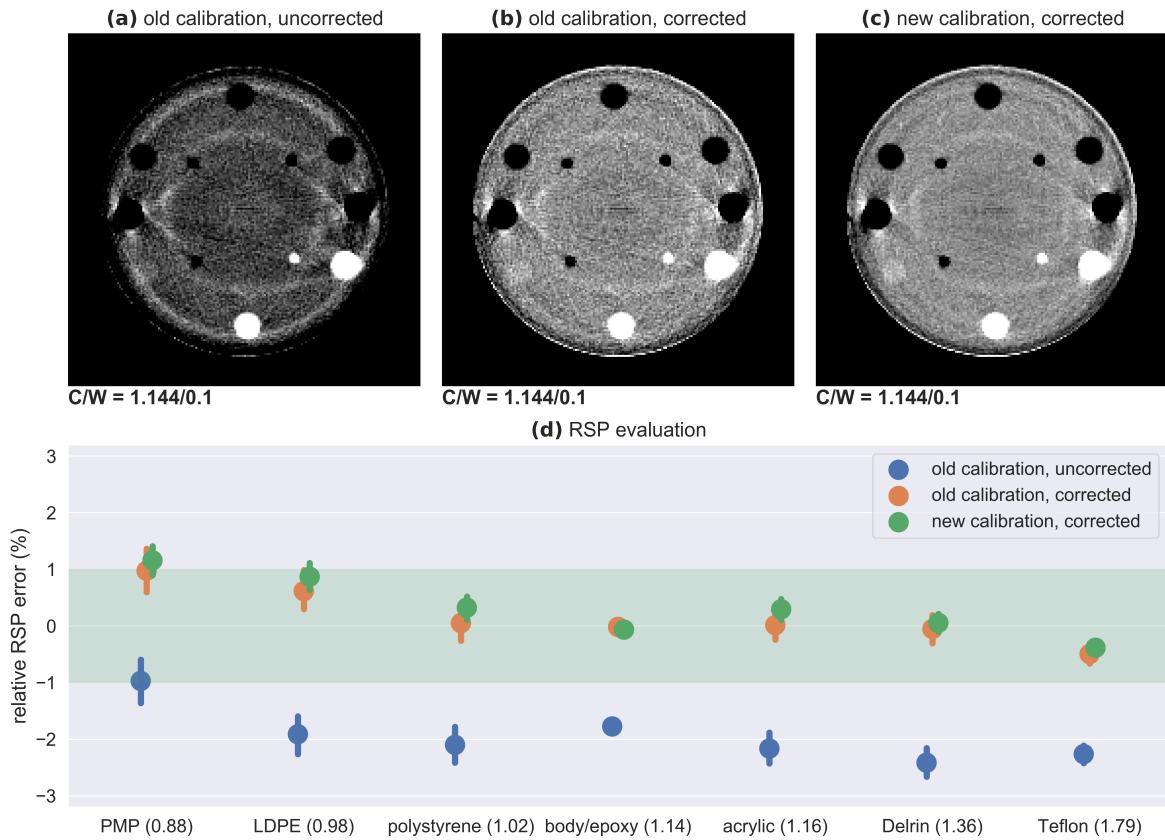
Figure 9.11: Evaluation of the RSP accuracy of the CTP phantom at $200\,\mathrm{MeV}$ and (a) using the calibration of another beamtime, (b) using the calibration of another beamtime and the correction method, (c) using the the correct calibration and the correction method; (d) the relative RSP error for all inserts and datasets with the green area indicating a $\pm 1\,\%$ interval.

the order of magnitude of ± 0.02 in terms of RSP. The scan at $187.5\,\mathrm{MeV}$ expectedly had a higher noise level, since less data were acquired for it.

DISCUSSION

An empirical artifact correction was adapted to the requirements of pCT and applied to improve the accuracy of experimental scans using a prototype scanner. Based on data of an elliptical phantom with known RSP, correction functions were found for two datasets at different incident energies. Both functions had an approximate slope of 1.3 %, which corrected for a proportional under–representation of RSP values in uncorrected scans. This bias was already observed in earlier experiments described in section 9.3 and could not be corrected by EMpCT. It may stem from an incorrect measurement of the calibration phantom's RSP or geometry. In both datasets, the offset is captured and corrected by the proposed empirical method in subsequent scans. In addition to that, the calibration functions exhibited both common and distinct features, that corrected for small band WEPL

(a) Head (200MeV), uncorrected **(b)** Head (200MeV), corrected **(c)** uncorrected - corrected

(d) Head (187.5MeV), uncorrected **(e)** Head (187.5MeV), corrected **(f)** uncorrected - corrected
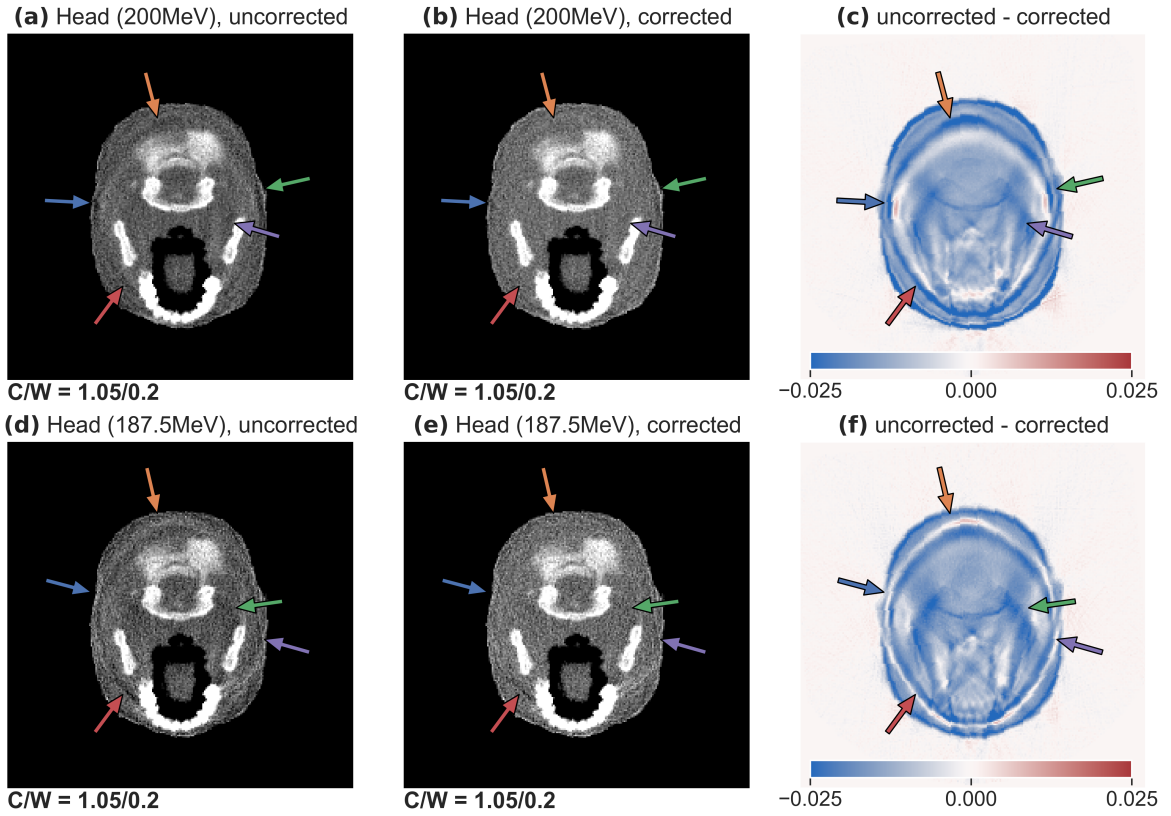
Figure 9.12: Reconstructions of the head phantom at (a,b) the high and the (d,e) the low energy with and without the correction method. In (c,f) absolute difference maps between the corrected and the uncorrected scan are shown. Arrows mark locations of the most visible artifacts in the uncorrected images.

inaccuracies. Distinct features are likely to be caused by stage interfaces and were already improved by EMpCT. Common features, however, could not be improved by EMpCT and may, again, result from an inaccurate manufacturing or knowledge of the geometry of the wedge–shaped calibration phantom. The correction functions had, by design, no effect below $s_{min} = 8\,\text{mm}$ or above $s_{max} = 192\,\text{mm}$. These two values corresponded to the minimum and maximum WEPL seen in the data of the ellipse phantom and s_{max} was larger than the maximum WEPL seen in any other phantom in this investigation.

The proposed method was applied to data of the ellipse phantom as a consistency check as well as to data of the water phantom to test the performance on a homogeneous object. In both cases, the initial bias as well as RSP artifacts were considerably reduced. Minor artifacts remained also in the data of the ellipse phantom, but line profiles through the center of the two homogeneous phantoms were within the intended $\pm 1\,\%$ interval around the ground truth value. The magnitude of ring artifacts in the uncorrected images was more than $2\,\%$ (minimum to maximum) and merely a shift of the mean value would not have sufficed to bring RSP values into the $\pm 1\,\%$ interval.

Application of the proposed method to data of the CTP phantom allowed to quantify the RSP accuracy before and after the correction and compare values to data of a previous

study. Ring artifacts as well as the negative bias were distorting the uncorrected scan as well as the data of Dedes et al. [41]. Using the correction, both artifacts and the bias were reduced and the MAPE of all inserts reduced by 49 % and 44 % in the two datasets, which is a considerable improvement. On average, the MAPE of the two incident energies was at 0.46 %, which clearly outperforms the results of the previous study on pCT RSP accuracy [41]. Nevertheless, a conclusion on the performance of artifact–corrected pCT scans with respect to the state–of–the–art DECT requires a larger study with more phantoms and inserts. A remaining trend of overestimated low–RSP inserts and underestimated high–RSP inserts remains to be studied and is not understood. It cannot be caused by the spatial resolution of the scanner, which is at 0.5 lp/mm for this prototype [204], and thus small compared to the 3 mm margin at the edge of each insert that was disregarded for the evaluation. It may be beneficial to apply the proposed method to data of a phantom composed of several RSP values. A homogeneous correction phantom was chosen in this study to avoid that the heterogeneity in connection with the limited spatial resolution of the scanner distorts correction results. However, a phantom consisting of two or three different RSP values may help to reduce the remaining RSP trend and further improve accuracy.

An interesting implication of the correction method was studied by processing the data of the CTP phantom with a 15 months old calibration and subsequently applying the proposed method. The resulting scan had a MAPE which agreed with the MAPE of a scan that was processed with the correct calibration and the proposed method. Consequently, the proposed method makes daily calibration runs using the wedge–shaped calibration phantom omissible and reduces the regular calibration effort to just one scan of the ellipse phantom. This may be beneficial to reduce the time needed for calibrations in upcoming beamtimes, since a calibration run typically takes 30 min to 40 min, while a single scan of the ellipse phantom would be just 6 min.

Application of the proposed method to data of the anthropomorphic head phantom allowed to reduced visible ring artifacts considerably and to improve the homogeneity of regions that are expected to be homogeneous from the design of the phantom. RSP errors of up to ±0.02 were reduced, which is a considerable inaccuracy if not corrected appropriately. While distinct WEPL artifacts (with positive and negative amplitude) may average out to a certain extent for the proton range calculation, also the head phantom showed a negative bias, which would proportionally result in biased range calculations. This bias, as for the other phantoms, was corrected using the proposed method.

9.5 CONCLUSIONS

In section 9.2, a method was presented to isolate ring artifacts in scans of the prototype pCT scanner. With the method, rings could be linked to distinct WEPL regions of the calibration curve. Thereby, protons stopping close to stage interfaces could be identified by Dedes et al. [41] to cause spurious signals in the detector and distort the image. This has led to two different attempts to tackle the problem of distortions in pCT scans. While EMpCT in section 9.3 and originally in Dickmann et al. [197] directly addressed the physical reason of artifacts and tried to avoid using protons stopping close to stage interfaces, it failed to

correct for an inherent offset in the data. Therefore, in a second study in section 9.4 an empirical method was applied, that makes no assumptions on the origin of artifacts, but just tries to correct them in postprocessing using data of a phantom with known RSP.

The use of energy modulation on the homogeneous water phantom allowed to reduce the magnitude of ring artifacts considerably and make the image visually more homogeneous. Datasets were merged in postprocessing, but EMpCT was designed keeping in mind the acquisition of data with small pencil beams at modulated energies. The use of two energies was tailored for the specific prototype scanner under investigation, where it allowed to avoid protons stopping close to stage interfaces. Using more than two energies would eventually allow to change the design of the detector to contain just a single instead of five stages. Using EMpCT, the IQR, which is a measure of the magnitude of ring artifacts, was reduced by a factor of two to three. An offset of RSP values of about 1 % could, by design, not be corrected and remained in the merged data.

Using the empirical artifact correction, which adapted a method for x–ray CT to the requirements of pCT, allowed to address the biased RSP values of pCT scans. A correction function was found based on a scan of an elliptical homogeneous correction phantom and at two incident energies. The two energies were utilized to increase the number of test cases and a single energy is sufficient to apply the proposed method. The optimized correction functions captured an offset of 1.3 %, but also corrected for small–banded WEPL distortions, with some of them common in both datasets and some distinct. Resulting corrected RSP maps of the water phantom, the CTP phantom, and the head phantom were analyzed for distortions and their accuracy. Images of the water phantom were more homogeneous and the magnitude of remaining ring artifacts was small compared to those in the uncorrected data. The MAPE of the CTP phantom improved by 47 % on average and was below 0.48 % for both energies. Image artifacts in an anthropomorphic head phantom were clearly reduced. Using this empirical method, therefore, allows to considerably improve accuracy of pCT imaging.

10 | CONCLUSIONS & PERSPECTIVES

I MAGING WITH PROTON COMPUTED TOMOGRAPHY, and in particular the use of fluence modulation to reduce patient dose exposure, was the topic of investigation in this thesis. In five chapters, scientific questions have been answered and novel algorithms were developed. In chapter 5, contributions to image noise of a pCT scanner were investigated using Monte Carlo simulations and compared to measurements. Chapter 6 proposed an algorithm for optimizing FMpCT fluence patterns with a variance objective. These patterns were realized experimentally using a prototype pCT scanner and a PBS beamline in chapter 7. In chapter 8, a second optimization algorithm for FMpCT was developed, that applied concepts of treatment planning and that could optimize fluences based on imaging dose and image variance objectives. Lastly, the experimental operation of the prototype pCT scanner resulted in the development of two artifact reduction methods presented in chapter 9, that allowed to improve the accuracy of the prototype scanner. In the following, the main conclusions of these five chapters are summarized and future perspectives are identified.

IMAGE NOISE   Fluence modulation for pCT requires a precise model of the image noise expected for a given fluence setting. For this, a Monte Carlo simulation with a detailed model of the prototype pCT scanner was improved to model the effect of quenching in the scintillator and the initial energy spread of the proton beam. Both effects were modeled based on experimental data of the scanner. With this, a good agreement for noise prediction could be obtained, with discrepancies between predicted and experimental noise of less than 7 % in terms of the standard deviation. The use of variance reconstruction allowed to estimate noise not only at the detector level, but also within every voxel of the reconstructed volume. Since the proposed approach does not require to evaluate noise in a small ROI with constant RSP or many repetitions of the experiment or simulation, this is an important contribution for the future investigation of pCT image noise in experiments with heterogeneous phantoms, or patients.

Using variance reconstruction, image noise of the prototype pCT scanner was compared between homogeneous phantoms like the water phantom and an anthropomorphic head phantom. From considerations based on homogeneous media of Bashkirov et al. [56], the noise of both phantoms, which had a similar diameter, should be the same and roughly constant, since a WEPL uncertainty of 3 mm independent of the WEPL is to be expected for the prototype scanner. However, noise for the water phantom was increased considerably at the edges and the standard deviation in the center of the head phantom was elevated by

a factor of two compared to the center of the water phantom. Such differences were even observed between two non–adjacent slices of the same scan of the head phantom. This is a relevant observation, since image noise should ideally be constant if the entire image is to be used for diagnostic purposes and, compensating for a factor of two difference in standard deviation requires four times the patient dose exposure. To study this, the noise level of the prototype scanner was disentangled into contributions from the energy detection process, energy straggling in in the object, tracking, MCS and the initial beam energy spread. This revealed that the noise increase seen in heterogeneous regions and at the outer hull of a phantom was driven by MCS. The non–straight path of protons can make particles with different WEPL histories scatter to the same location in the phantom, thereby increasing the variance inside a pixel of the distance–driven binned projection. Since this is a physical effect, it cannot be avoided and is an inherent noise contribution of pCT. Bashkirov et al. [56] only considered data from the homogeneous calibration phantom, for which two scattered proton trajectories always experience very similar WEPLs, since there are no material changes and protons scattering out of the phantom were not considered. They were, therefore, right to conclude a WEPL uncertainty of 3 mm independent of the WEPL. However, this situation is not representative of heterogeneous phantoms or at the edge of phantoms, where variance due to MCS comes on top of the 3 mm uncertainty from the detection process.

PCT image noise at a uniform fluence, therefore, is inherently non–uniform. However, fluence modulation can be used to achieve constant image noise. At the same peak noise level, this saves imaging dose in homogeneous regions where noise is reduced in a uniform fluence scan. This first feature of FMpCT was demonstrated by calculating a fluence profile that would yield constant noise in the image for the water phantom, which is the equivalent of a bow–tie filter in x–ray CT. Calculating such profiles requires a precise model of pCT image noise, which was established in this thesis for a specific prototype pCT scanner.

In future works, the dose efficiency of pCT may need to be revisited. While Schulte et al. [44] concluded an advantage of pCT over x–ray–based imaging for requiring less dose at the same noise level, this comparison was done at the center of a homogeneous water phantom, where the noise contribution of MCS was negligible. However, pCT noise would increase in heterogeneous regions and in particular in a patient scan of the head. Consequently, pCT may require more dose at the same noise level in such regions and the dose efficiency benefit may diminish. At the same time, it is interesting to study the impact of other particle types such as helium–ions, which have been used for particle imaging [83]. On one hand, they exhibit less noise due to reduced energy straggling and MCS because of their larger mass, but fragmentation and increased nuclear interactions require a strict filtration of detected particles [85], which reduces statistics and increases noise. The framework developed here may also be used to optimize the design of future pCT detectors, in particular with respect to image noise and comparable to the study of Krah et al. [108], who studied different scanner designs in terms of spatial resolution.

FMPCT WITH VARIANCE OBJECTIVES Using the precise image variance model via Monte Carlo simulation, an optimization algorithm was developed, that calculates fluence patterns based on image variance objectives. The algorithm makes use of an iterative variance forward projection, to transform the variance prescription from the image to the projections,

where fluences can be independently optimized based on a pencil beam model. The forward projection method was based on a ray–tracing operation on the prescription variance followed by a ramp filtration. Since this caused negative and therefore unphysical variance values, the variance projections needed to be thresholded to positive values, which introduced a discrepancy to the prescription. This discrepancy was reduced iteratively, for which the convergence speed depended on the shape of the prescription.

In a Monte Carlo simulation study based on three phantoms, for which physical counterparts existed, three different fluence modulation targets were investigated. The bow–tie filter calculation could be repeated using the optimization both for the water phantom, where it resulted in a dose reduction of 8.9 % and for the heterogeneous head phantom, where the dose reduction at equal peak noise level was 16.0 %. In these two scans, the image noise was aimed to be homogeneous, allowing to decrease imaging fluence and dose in regions were noise was initially lower. Such fluence modulations are, in the form of bow–tie filters, a standard component of almost every x–ray CT scanner. For pCT, a corresponding implementation is more complicated since it requires a dynamic fluence modulation and an initial knowledge of the patient's geometry. A simple static fluence modulation pattern would not be sufficient. This highlights the fact, that dynamic fluence modulation is a required and essential component for pCT imaging to keep the patient dose exposure low and to allow for a frequent repetition of imaging during the particle therapy treatment. In fact, dose savings in the context of particle therapy can further be increased by prescribing more sophisticated noise targets. In subsequent simulations, two FMpCT scans were investigated, where the ROIs followed a typical shape of a proton therapy treatment plan. With this, image quality can be maintained, precisely where it is needed to calculate the therapeutic dose, and imaging dose can be reduced elsewhere. For the two FMpCT scans, imaging dose savings between 25.7 % and 40.5 % were achieved outside of the ROI. This is a substantial reduction and opens an interesting perspective for daily image guidance and adaptive particle therapy.

While promising results were achieved using the proposed method, limitations, that were addressed in following works of this thesis, were identified. The optimization algorithm operated solely based on image variance. Imaging dose was only implicitly minimized by prescribing higher variance outside of the ROI. This variance value was not optimized and higher values and in particular inhomogeneous variance distributions outside of the ROI may have resulted in better dose savings. In addition, the algorithm considered variances from single projections independently and could not compensate for an increased variance in one projection by a decreased variance in another projection. For this purpose, a second algorithm based on variance and dose objectives was developed later. Moreover, this simulation study was designed for the experimental realization of FMpCT. It, therefore, did not include any patient data, but only phantoms, for which physical counterparts existed, that could be used in a subsequent experimental acquisition.

EXPERIMENTAL REALIZATION OF FMPCT   The simulation study for the FMpCT optimization with variance objectives was specifically designed to be applied in an experimental scan. For this reason, it modeled the pencil beams of the PBS beamline in Chicago as well as the entire scanner and phantom geometries. This allowed to directly use the optimized

pencil beam weights in an experimental acquisition. The PBS delivery system was interfaced to deliver modulated fluence patterns using a grid of regularly spaced pencil beams. In the resulting scans, an unexpected spatial dependency of the initial energy within a single pencil beam was identified, that distorted the RSP images. This dependency was isolated and could be corrected in all scans. In addition, a small misalignment of the quarter–shifted fluence pattern caused a distortion of the fluence delivery that resulted in inhomogeneous image variance where homogeneous variance was expected. These modulations could not be corrected, but they were only of a small extent compared to the overall variance modulation that was achieved with FMpCT. In general, the agreement of the resulting image variance with the planned image variance from simulations was satisfactory. No distortion of the RSP accuracy was observed to be caused by the FMpCT acquisition. A good agreement of fluence maps confirmed that the dose savings, although they were not measured, were comparable to those determined in the corresponding simulations. This is the first experimental realization of optimized FMpCT scans using a prototype pCT scanner and completes the proof–of–concept works of Dedes et al. [46].

Future experimental work should focus on the automation of the fluence delivery in synchrony with the acquisition of the scanner. Due to a missing link between the PBS system and the pCT scanner, acquisition times were increased because of the manual initiation of each projection of the scan. The beam–on time just amounted to 26 % of the total acquisition time. Establishing a link may be possible with little modifications of the hardware and would allow for a near–continuous delivery of the modulated imaging fluence, which in turn would speed up acquisition times dramatically. In addition, the quarter–shifted pattern is to be avoided in the future, since already minimal shifts of 0.5 mm as for some scans in this study can result in distortions of the fluence pattern. Moreover, automated alignment procedures could be found that allow for a precise determination of the scanners coordinate system with respect to the PBS nozzle.

FMPCT WITH DOSE AND VARIANCE OBJECTIVES Despite its success in reducing the imaging dose, the first optimization algorithm presented in this thesis had clear limitations in particular due to the lack of a dose term in the cost function. For this purpose, a novel optimizer for FMpCT with dose and variance objectives was developed. This directly optimizes the two quantities, each in the region where it is relevant: variance can be maintained in the ROI and outside of it, dose is minimized. The optimization algorithm employed concepts of treatment planning for radiation therapy by using matrices that describe the dose contribution of a given bixel to a given voxel. For variance, it was possible to replicate the variance reconstruction formula using a similar matrix multiplication. This required to first optimize fluences based on small independent bixels, and subsequently fit the pencil beam model to calculate pencil beam weights. Due to its fast implementation on a multi–core system and using dedicated libraries, the optimization was performed efficiently and fast. Spatial weights in the dose term of the cost function could be used to further increase the dose saving in pre–defined imaging OARs. In addition, the trade–off between achieving a given image variance in the ROI and saving dose to critical structures can be decided by the user based on medical considerations. This is similar to treatment planning, where a trade–off between the conformity of the prescription dose with the PTV and the dose saving

to OARs has to be made. The proposed algorithm is very flexible to adapt to arbitrary ROI shapes and sizes, and task–specific dose and variance prescriptions as well as independently weighted sparing of imaging OARs.

The performance of the algorithm was tested in a simulation study, which was based on imaging data of three pediatric patients undergoing radiotherapy treatment of the head. Proton therapy treatment plans were optimized for each patient based on the original delineations and prescriptions. From these therapeutic dose distributions, the imaging ROI was derived and FMpCT plans were optimized. Subsequently, the accuracy of pCT and FMpCT scans was tested by re–calculating the optimized treatment doses. Passing rates for strict criteria were satisfactory and distortions of the proton range calculations were below 0.3 mm for all scans. No relevant deterioration of the dosimetric accuracy was observed. At the same time, imaging doses for the three patients could be reduced by 74 % outside of the ROI, which is a considerable improvement compared to uniform fluence pCT scans, but also compared to the previous optimization algorithm. The dose savings were similar for all patients and thus comparable reductions are expected for similar cases of the head and neck region. Such dose savings may pave the way towards daily adaptive particle therapy with pCT image guidance.

In future works, the clinical applicability of optimized FMpCT scans may be investigated by including larger cohorts of patients and studying limitations of using pCT in general and FMpCT in particular. Such limitations could be the size of patients with respect to the size of the scanner and its field–of–view. Limitations could also be with respect to the treatment workflow and studies should investigate the time budget that daily pCT imaging and FMpCT optimization would require and if or how it would fit into clinical routine.

IMAGE ACCURACY Dose reductions achievable with FMpCT have little value if inaccuracies in the scans prevent their usage for the calculation of particle therapy treatment plans. For this reason, the performance of pCT was compared to the state–of–the–art x–ray based imaging modality DECT by Dedes et al. [41]. Within this work, to which the author contributed, pCT was found to be on par with DECT in terms of RSP accuracy. Ring artifacts due to distorted WEPL intervals have been identified to degrade the performance of the prototype pCT scanner used in the context of this thesis. This initiated the development of two correction algorithms that aimed to reduce image artifacts and improve the performance of the prototype scanner in producing accurate RSP maps. The first algorithm used energy modulation to avoid protons stopping close to interfaces of the scanner's five–stage energy detector. Such protons were suspected to cause distorted WEPL measurements and cause ring–shaped artifacts in cylindrical phantoms. With EMpCT it was possible to avoid these critical structures and artifacts were reduced considerably. However, an offset and minor artifacts remained, which could not be corrected by the method. For this purpose, a second empirical correction method was developed, that makes no assumption on the origin of artifacts. It calculates a WEPL correction function based on the scan of a custom–built phantom with known RSP. This function was applied to subsequent scans, which removed a 1.3 % offset and considerably reduced image artifacts present in the uncorrected scans. The performance of the method was tested on several phantoms, including the sensitometric CTP phantom, for which the MAPE was reduced by 47 % and thus almost to half of its value

for uncorrected scans. The RSP accuracy was better than 0.48 % for two scans at different initial energies. In addition, artifacts in the anthropomorphic head phantom were clearly reduced. The performance of the scanner was, therefore, improved considerably, without the need for additional hardware. In addition, the proposed correction method made a daily calibration of the scanner omissible, and, thereby, reduced the pre–scan preparation time of future beamtimes with the prototype scanner.

While the results of the EMpCT study may help in the design of monolithic detectors with only a single stage, where energy modulation is required, the empirical correction method appears to be most promising to be applied in future scans of the particular prototype pCT scanner, and other scanners. Future studies should investigate its impact on a direct comparison with DECT with more phantoms than those used in this thesis. Such a future comparative study could also include the use of DECT and corrected pCT for the calculation of proton ranges. It would also be interesting to see such results for fresh meat samples as it was already done for DECT by Niepel et al. [78].

––––––––––

Proton computed tomography, as first envisioned by Cormack [36] in 1963, has seen considerable developments in recent years due to the availability of fast hardware that allows for the acquisition of single–particle tracking pCT data. This has resulted in the development of numerous prototype scanners, of which some still need to be realized. Proton CT seems to be the natural choice as imaging modality used for particle therapy treatment planning, since it measures the interaction of protons with matter directly: via the energy loss of protons in matter. Thereby, it employs only little imaging dose and prototype scanners report promising accuracies that can compete with the state–of–the–art in the clinics. This thesis engaged with both: reducing the patient dose of pCT and improving its accuracy. Using fluence–modulated proton CT an elegant method has been established — in simulations and experiments — that saves imaging dose by avoiding to acquire data that is not needed for treatment planning. The full workflow of fluence–modulated pCT was developed including clinically relevant optimization targets and important aspects of its experimental realization. At the same time, artifacts of the prototype pCT scanner were isolated and successfully corrected, which improved the accuracy of the scanner. With these two contributions, proton computed tomography has, hopefully, come a small step closer to the clinics.

ACKNOWLEDGMENTS

While in the last year our office was at home, I was lucky to share for two years an office in Garching with Dr. Liheng Tian, Katharina Niepel, Michael Stanislawski, Ze Huang, Lourival Beltrão Martins, and Philipp Wesp and would like to thank you for your support and friendship. I would also like to thank my colleagues and friends at the department, in particular Katrin Schnürle who was always there when coffee was needed, Henning Schmitz and Sascha Huck for always interesting discussions, Ronaldo Kalunga who is the best chef, Giulio Lovatti who has drawn the beautiful illustration on page IX and also was my flatmate, Juliana Martins who always joined to eat Krapfen, and Franz Englbrecht who tried to explain Bavaria to me.

I would like to thank my parents, my brother Marcel and Oma Karin for their support. And last but definitely not least, I want to thank Giulia for having shared with me a wonderful time with countless adventures to come.

PUBLICATIONS

PUBLICATIONS IN SCIENTIFIC JOURNALS

[100] » J. Dickmann et al.: "Prediction of image noise contributions in proton computed tomography and comparison to measurements". *Physics in Medicine & Biology* **64**(14) (2019), 145016. ISSN: 1361-6560. DOI: 10.1088/1361-6560/ab2474

[156] » J. Dickmann et al.: "An optimization algorithm for dose reduction with fluence-modulated proton CT". *Medical Physics* **47**(4) (2020), 1895–1906. ISSN: 0094-2405. DOI: 10.1002/mp.14084

[99] » J. Dickmann et al.: "Experimental realization of dynamic fluence field optimization for proton computed tomography". *Physics in Medicine & Biology* **65**(19) (2020), 195001. DOI: 10.1088/1361-6560/ab9f5f

[138] » J. Dickmann et al.: "Fluence-modulated proton CT optimized with patient-specific dose and variance objectives for proton dose calculation". *Physics in Medicine & Biology* **66**(6) (2021), 064001. DOI: 10.1088/1361-6560/abe3d2

[197] » J. Dickmann et al.: "Proof of concept image artifact reduction by energy-modulated proton computed tomography (EMpCT)". *Physica Medica* **81** (2021), 237–244. ISSN: 11201797. DOI: 10.1016/j.ejmp.2020.12.012

[141] » J. Dickmann et al.: "An empirical artifact correction for proton computed tomography". *submitted for publication* (2021)

CONTRIBUTIONS TO PUBLICATIONS

[41] » G. Dedes, J. Dickmann et al.: "Experimental comparison of proton CT and dual energy x-ray CT for relative stopping power estimation in proton therapy". *Physics in Medicine & Biology* **64**(16) (2019), 165002. ISSN: 1361-6560. DOI: 10.1088/1361-6560/ab2b72

[102] » G. Dedes, J. Dickmann et al.: "The role of Monte Carlo simulation in understanding the performance of proton computed tomography". *Zeitschrift für Medizinische Physik* (available online, 2020). ISSN: 0939-3889. DOI: 10.1016/j.zemedi.2020.06.006

CONFERENCE CONTRIBUTIONS

💬 [147] » J. Dickmann et al.: "Modelling of contributions to image variance in proton CT for application in low-dose fluence-modulated imaging". *Conference of the Particle Therapy Cooperative Group (PTCOG), Manchester, United Kingdom* (2019)

💬 [157] » J. Dickmann et al.: "Method for Fluence Field Optimization to Achieve Nonconvex Image Noise Prescriptions with Fluence-Modulated Proton CT". *Annual Meeting of the American Association of Physicists in Medicine (AAPM), San Antonio, USA, proceedings published in Medical Physics* **46**(6) (2019), E477–E477. ISSN: 24734209. DOI: 10.1002/mp.13589

💬 [198] » J. Dickmann et al.: "Understanding image artifacts for a prototype proton computed tomography scanner via Monte Carlo simulations". *International Conference on Monte Carlo Techniques for Medical Applications (MCMA), Montréal, Canada* (2019)

💬 [158] » J. Dickmann et al.: "Dosiseinsparung durch fluenzmodulierte Protonen-Computertomographie mit Anwendung in der Protonentherapie". *Jahrestagung der Deutschen Gesellschaft für Medizinische Physik (DGMP), Stuttgart, Germany* (2019)

💬 [148] » J. Dickmann et al.: "Monte Carlo study of image noise contributions of a prototype proton computed tomography scanner". *Jagiellonian Symposium on Fundamental and Applied Subatomic Physics, Krakow, Poland* (2019)

💬 [159] » J. Dickmann et al.: "PO-1706: Low dose fluence-modulated proton CT: simulation study and first experimental results". *Conference of the European Society for Radiotherapy and Oncology (ESTRO), Vienna, Austria (online), proceedings published in Radiotherapy and Oncology* **152**(S1) (2020), S941. DOI: 10.1016/S0167-8140(21)01724-2

💬 [182] » J. Dickmann et al.: "Prescribing Image Noise Using Dynamic Fluence Field Optimization: Experimental Results Using a Pre-Clinical Proton CT Scanner". *Joint Annual Meeting of the American Association of Physicists in Medicine and the Canadian Organization of Medical Physics (AAPM/COMP), Vancouver, Canada (online), proceedings published in Medical Physics* **47**(6) (2020), 2580–2674. ISSN: 24734209. DOI: 10.1002/mp.14315

💬 [183] » J. Dickmann et al.: "Dynamic Fluence Modulation using Proton CT for Low-dose Imaging in Particle Therapy". *International Conference on Image Formation in X-Ray Computed Tomography (CT Meeting), Regensburg, Germany (online)* (2020)

💬 [184] » J. Dickmann et al.: "Joint Dose Minimization and Variance Optimization for Fluence-Modulated Proton CT". *IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), Boston, USA (online)* (2020)

ETHICS STATEMENT

T HE USE OF PATIENT IMAGING DATA in this retrospective study was exempt from requiring ethics approval. Bavarian state law (Bayerisches Krankenhausgesetz/Bavarian Hospital Law §27 Absatz 4 Datenschutz (data protection)) allows the use of patient data for research, provided that any person's related data are kept anonymous. German radiation protection laws request a regular analysis of outcomes in the sense of quality control and assurance, thus in the case of purely retrospective studies no additional ethical approval is needed under German law.

BIBLIOGRAPHY

[1] D. Hanahan and R. A. Weinberg: "Hallmarks of cancer: The next generation". *Cell* **144**(5) (2011), 646–674. ISSN: 00928674. DOI: 10.1016/j.cell.2011.02.013.

[2] T. Reya, S. J. Morrison, M. F. Clarke, and I. L. Weissmann: "Stem cells, cancer, and cancer stem stells". *Nature* **414** (2001), 105–111. DOI: 10.1038/35102167.

[3] B. Steward and C. Wild: *World Cancer Report 2014*. Ed. by B. Steward and C. Wild. International Association for Research and Cancer, 2014. ISBN: 9789283204299.

[4] R. Goldbrunner, G. Minniti, M. Preusser, M. D. Jenkinson, K. Sallabanda, E. Houdart, A. V. Deimling, P. Stavrinou, F. Lefranc, M. Lund-Johansen, E. Cohen, J. Moyal, D. Brandsma, R. Henriksson, R. Soffi, and M. Weller: "EANO guidelines for the diagnosis and treatment of meningiomas". *Lancet Oncology* **17**(9) (2016), E383–E391. DOI: 10.1016/S1470-2045(16)30321-7.

[5] K. D. Held, H. Kawamura, T. Kaminuma, A. E. S. Paz, Y. Yoshida, Q. Liu, H. Willers, and A. Takahashi: "Effects of Charged Particles on Human Tumor Cells". *Frontiers in Oncology* **6** (2016). ISSN: 2234-943X. DOI: 10.3389/fonc.2016.00023.

[6] D. A. Jaffray: "Image-guided radiotherapy: from current concept to future perspectives". *Nature Reviews Clinical Oncology* **9**(12) (2012). ISSN: 1759-4774. DOI: 10.1038/nrclinonc.2012.194.

[7] J. M. Galvin, A. R. Smith, and B. Lally: "Characterization of a multileaf collimator system". *International Journal of Radiation Oncology*Biology*Physics* **25**(2) (1993). ISSN: 03603016. DOI: 10.1016/0360-3016(93)90339-W.

[8] C. Scholz, S. Nill, and U. Oelfke: "Comparison of IMRT optimization based on a pencil beam and a superposition algorithm". *Medical Physics* **30**(7) (2003), 1909–1913. ISSN: 00942405. DOI: 10.1118/1.1586452.

[9] F. Tommasino, A. Nahum, and L. Cella: "Increasing the power of tumour control and normal tissue complication probability modelling in radiotherapy: Recent trends and current issues". *Translational Cancer Research* **6** (2017), S807–S821. ISSN: 22196803. DOI: 10.21037/tcr.2017.06.03.

[10] W. P. Levin, H. Kooy, J. S. Loeffler, and T. F. DeLaney: "Proton beam therapy". *British Journal of Cancer* **93**(8) (2005), 849–854. ISSN: 00070920. DOI: 10.1038/sj.bjc.6602754.

[11] M. Durante, R. Orecchia, and J. S. Loeffler: "Charged-particle therapy in cancer: clinical uses and future perspectives". *Nature Reviews Clinical Oncology* **14**(8) (2017). ISSN: 1759-4774. DOI: 10.1038/nrclinonc.2017.30.

[12] M. Durante and J. Flanz: "Charged particle beams to cure cancer: Strengths and challenges". *Seminars in Oncology* **46**(3) (2019). ISSN: 00937754. DOI: 10.1053/j.seminoncol.2019.07.007.

[13] D. C. Weber, R. Schneider, G. Goitein, T. Koch, C. Ares, J. H. Geismar, A. Schertler, A. Bolsi, and E. B. Hug: "Spot Scanning-Based Proton Therapy for Intracranial Meningioma: Long-Term Results From the Paul Scherrer Institute". *International Journal of Radiation Oncology*Biology*Physics* **83**(3) (2012), 865–871. ISSN: 03603016. DOI: 10.1016/j.ijrobp.2011.08.027.

[14] J. Park, Y. Park, S. U. Lee, T. Kim, Y.-K. Choi, and J.-Y. Kim: "Differential dosimetric benefit of proton beam therapy over intensity modulated radiotherapy for a variety of targets in patients with intracranial germ cell tumors". *Radiation Oncology* **10**(1) (2015), 135. ISSN: 1748-717X. DOI: 10.1186/s13014-015-0441-5.

[15] K. Nakajima, H. Iwata, H. Ogino, Y. Hattori, S. Hashimoto, K. Hayashi, T. Toshito, K. Akita, M. Iwana, F. Baba, K. Nakamae, Y. Shibamoto, and J. Mizoe: "Clinical Outcomes of Image-Guided Proton Therapy for Stage I Non–small Cell Lung Cancer". *International Journal of Radiation Oncology*Biology*Physics* **99**(2) (2017), E483–E484. ISSN: 03603016. DOI: 10.1016/j.ijrobp.2017.06.1759.

[16] D. Schardt, T. Elsässer, and D. Schulz-Ertner: "Heavy-ion tumor therapy: Physical and radiobiological benefits". *Reviews of Modern Physics* **82**(1) (2010), 383–425. ISSN: 00346861. DOI: 10.1103/RevModPhys.82.383.

[17] R. Mohan, C. R. Peeler, F. Guan, L. Bronk, W. Cao, and D. R. Grosshans: "Radiobiological issues in proton therapy". *Acta Oncologica* **56**(11) (2017), 1367–1373. ISSN: 1651226X. DOI: 10.1080/0284186X.2017.1348621.

[18] C. Grau, M. Durante, D. Georg, J. A. Langendijk, and D. C. Weber: "Particle therapy in Europe". *Molecular Oncology* **14**(7) (2020), 1492–1499. ISSN: 18780261. DOI: 10.1002/1878-0261.12677.

[19] J. B. Farr, J. B. Flanz, A. Gerbershagen, and M. F. Moyers: "New horizons in particle therapy systems". *Medical Physics* **45**(11) (2018), e953–e983. ISSN: 00942405. DOI: 10.1002/mp.13193.

[20] D. A. Jaffray, S. Das, P. M. Jacobs, R. Jeraj, and P. Lambin: "How Advances in Imaging Will Affect Precision Radiation Oncology". *International Journal of Radiation Oncology*Biology*Physics* **101**(2) (2018). ISSN: 03603016. DOI: 10.1016/j.ijrobp.2018.01.047.

[21] A. Bolsi, M. Peroni, D. Amelio, A. Dasu, M. Stock, I. Toma-Dasu, P. W. Nyström, and A. Hoffmann: "Practice patterns of image guided particle therapy in Europe: A 2016 survey of the European Particle Therapy Network (EPTN)". *Radiotherapy and Oncology* **128**(1) (2018), 4–8. ISSN: 18790887. DOI: 10.1016/j.radonc.2018.03.017.

[22] N. Hudobivnik, F. Schwarz, T. Johnson, L. Agolli, G. Dedes, T. Tessonnier, F. Verhaegen, C. Thieke, C. Belka, W. H. Sommer, K. Parodi, and G. Landry: "Comparison of proton therapy treatment planning for head tumors with a pencil beam algorithm on dual and single energy CT images". *Medical Physics* **43**(1) (2016), 495–504. ISSN: 00942405. DOI: 10.1118/1.4939106.

[23] M. Wedenberg, C. Beltran, A. Mairani, and M. Alber: "Advanced Treatment Planning". *Medical Physics* **45**(11) (2018). ISSN: 0094-2405. DOI: 10.1002/mp.12943.

[24] G. Landry and C.-H. Hua: "Current state and future applications of radiological image guidance for particle therapy". *Medical Physics* **45**(11) (2018), e1086–e1095. ISSN: 00942405. DOI: 10.1002/mp.12744.

[25] L. Nenoff, M. Matter, J. H. Lindmar, D. C. Weber, A. J. Lomax, and F. Albertini: "Daily adaptive proton therapy – the key to innovative planning approaches for paranasal cancer treatments". *Acta Oncologica* **58**(10) (2019), 1423–1428. DOI: 10.1080/0284186X.2019.1641217.

[26] A.-C. Knopf, D. Boye, A. Lomax, and S. Mori: "Adequate margin definition for scanned particle therapy in the incidence of intrafractional motion". *Physics in Medicine & Biology* **58**(17) (2013). ISSN: 0031-9155. DOI: 10.1088/0031-9155/58/17/6079.

[27] S. K. Vinod, M. G. Jameson, M. Min, and L. C. Holloway: "Uncertainties in volume delineation in radiation oncology: A systematic review and recommendations for future studies". *Radiotherapy and Oncology* **121**(2) (2016), 169–179. ISSN: 18790887. DOI: 10.1016/j.radonc.2016.09.009.

[28] G. Meschini, A. Vai, C. Paganelli, S. Molinelli, G. Fontana, A. Pella, L. Preda, V. Vitolo, F. Valvo, M. Ciocca, M. Riboldi, and G. Baroni: "Virtual 4DCT from 4DMRI for the management of respiratory motion in carbon ion therapy of abdominal tumors". *Medical Physics* **47**(3) (2020), 909–916. ISSN: 00942405. DOI: 10.1002/mp.13992.

[29] W. Cao, G. Lim, L. Liao, Y. Li, S. Jiang, X. Li, H. Li, K. Suzuki, X. R. Zhu, D. Gomez, and X. Zhang: "Proton energy optimization and reduction for intensity-modulated proton therapy". *Physics in Medicine & Biology* **59**(21) (2014), 6341–6354. ISSN: 13616560. DOI: 10.1088/0031-9155/59/21/6341.

[30] P. Wohlfahrt, C. Möhler, K. Stützer, S. Greilich, and C. Richter: "Dual-energy CT based proton range prediction in head and pelvic tumor patients". *Radiotherapy and Oncology* **125**(3) (2017), 526–533. ISSN: 01678140. DOI: 10.1016/j.radonc.2017.09.042.

[31] E. G. Troost, D. Thorwarth, and W. J. Oyen: "Imaging-based treatment adaptation in radiation oncology". *Journal of Nuclear Medicine* **56**(12) (2015), 1922–1929. ISSN: 2159662X. DOI: 10.2967/jnumed.115.162529.

[32] C. Kontaxis, G. H. Bol, J. J. Lagendijk, and B. W. Raaymakers: "A new methodology for inter- and intrafraction plan adaptation for the MR-linac". *Physics in Medicine & Biology* **60**(19) (2015), 7485–7497. ISSN: 13616560. DOI: 10.1088/0031-9155/60/19/7485.

[33] C. Kurz, G. Buizza, G. Landry, F. Kamp, M. Rabe, C. Paganelli, G. Baroni, M. Reiner, P. J. Keall, C. A. V. D. Berg, and M. Riboldi: "Medical physics challenges in clinical MR-guided radiotherapy". *Radiation Oncology* **15**(1) (2020). ISSN: 1748717X. DOI: 10.1186/s13014-020-01524-4.

[34] J. Thariat, J. Hérault, A. Beddok, L. Feuvret, D. Dauvergne, M. Gérard, J. Balosso, G. Noël, and S. Valable: "Imaging issues specific to hadrontherapy (proton, carbon, helium therapy and other charged particles) for radiotherapy planning, setup, dose monitoring and tissue response assessment". *Cancer/Radiothérapie* **24**(5) (2020), 429–436. ISSN: 17696658. DOI: 10.1016/j.canrad.2020.01.010.

[35] R. Ricotti, A. Pella, B. Tagaste, G. Elisei, G. Fontana, M. Bonora, M. Ciocca, F. Valvo, R. Orecchia, and G. Baroni: "Long-time clinical experience in patient setup for several particle therapy clinical indications: management of patient positioning and evaluation of setup reproducibility and stability". *The British Journal of Radiology* **93**(1107) (2020), 20190595. DOI: 10.1259/bjr.20190595.

[36] A. M. Cormack: "Representation of a Function by Its Line Integrals, with Some Radiological Applications". *Journal of Applied Physics* **34**(9) (1963), 2722–2727. ISSN: 0021-8979. DOI: 10.1063/1.1729798.

[37] K. M. Hanson, J. N. Bradbury, T. M. Cannon, R. L. Hutson, D. B. Laubacher, R Macek, M. A. Paciotti, and C. A. Taylor: "Application of protons to computer tomography". *Nuclear Science Symposium* (1977).

[38] W. D. Newhauser and R. Zhang: "The physics of proton therapy". *Physics in Medicine & Biology* **60**(8) (2015), R155–R209. ISSN: 0031-9155. DOI: 10.1088/0031-9155/60/8/R155.

[39] A. J. Lomax: "Charged Particle Therapy: The Physics of Interaction". *The Cancer Journal* **15**(4) (2009), 285–292. DOI: 10.1097/PPO.0b013e3181af5cc7.

[40] D. C. Hansen, J. Seco, T. S. Sørensen, J. B. B. Petersen, J. E. Wildberger, F. Verhaegen, and G. Landry: "A simulation study on proton computed tomography (CT) stopping power accuracy using dual energy CT scans as benchmark". *Acta Oncologica* **54**(9) (2015), 1638–1642. ISSN: 0284-186X. DOI: 10.3109/0284186X.2015.1061212.

[41] G. Dedes, J. Dickmann, K. Niepel, P. Wesp, R. P. Johnson, M. Pankuch, V. Bashkirov, S. Rit, L. Volz, R. W. Schulte, G. Landry, and K. Parodi: "Experimental comparison of proton CT and dual energy x-ray CT for relative stopping power estimation in proton therapy". *Physics in Medicine & Biology* **64**(16) (2019), 165002. ISSN: 1361-6560. DOI: 10.1088/1361-6560/ab2b72.

[42] N. Arbor, D. Dauvergne, G. Dedes, J. M. Létang, K. Parodi, C. T. Quiñones, E. Testa, and S. Rit: "Monte Carlo comparison of x-ray and proton CT for range calculations of proton therapy beams". *Physics in Medicine & Biology* **60**(19) (2015), 7585–7599. ISSN: 0031-9155. DOI: 10.1088/0031-9155/60/19/7585.

[43] S. Meyer, F. Kamp, T. Tessonnier, A. Mairani, C. Belka, D. J. Carlson, C. Gianoli, and K. Parodi: "Dosimetric accuracy and radiobiological implications of ion computed tomography for proton therapy treatment planning". *Physics in Medicine & Biology* **64**(12) (2019), 125008. ISSN: 1361-6560. DOI: 10.1088/1361-6560/ab0fdf.

[44] R. W. Schulte, V. Bashkirov, M. C. Klock, T. Li, A. J. Wroe, I. Evseev, D. C. Williams, and T. Satogata: "Density resolution of proton computed tomography". *Medical Physics* **32**(4) (2005), 1035–1046. ISSN: 00942405. DOI: 10.1118/1.1884906.

[45] G. Dedes, L. De Angelis, S. Rit, D. Hansen, C. Belka, V Bashkirov, R. P. Johnson, G. Coutrakon, K. E. Schubert, R. W. Schulte, K. Parodi, and G. Landry: "Application of fluence field modulation to proton computed tomography for proton therapy imaging". *Physics in Medicine & Biology* **62**(15) (2017), 6026–6043. ISSN: 1361-6560. DOI: 10.1088/1361-6560/aa7734.

[46] G. Dedes, R. P. Johnson, M. Pankuch, N. Detrich, W. M. A. Pols, S. Rit, R. W. Schulte, K. Parodi, and G. Landry: "Experimental fluence-modulated proton computed tomography by pencil beam scanning". *Medical Physics* **45**(7) (2018), 3287–3296. ISSN: 00942405. DOI: 10.1002/mp.12989.

[47] M. Durante and H. Paganetti: "Nuclear physics in particle therapy: A review". *Reports on Progress in Physics* **79**(9) (2016). ISSN: 00344885. DOI: 10.1088/0034-4885/79/9/096702.

[48] R. P. Johnson: "Review of medical radiography and tomography with proton beams". *Reports on Progress in Physics* **81**(1) (2018), 016701. ISSN: 0034-4885. DOI: 10.1088/1361-6633/aa8b1d.

[49] H. Bethe: "Zur Theorie des Durchgangs schneller Korpuskularstrahlen durch Materie". *Annalen der Physik* **397**(3) (1930), 325–400. DOI: https://doi.org/10.1002/andp.19303970303.

[50] F. Bloch: "Zur Bremsung rasch bewegter Teilchen beim Durchgang durch Materie". *Annalen der Physik* **408**(3) (1933), 285–320. DOI: https://doi.org/10.1002/andp.19334080303.

[51] U. Fano: "Penetration of Protons, Alpha Particles, and Mesons". *Annual Review of Nuclear Science* **13**(1) (1963), 1–66. DOI: 10.1146/annurev.ns.13.120163.000245.

[52] W. Barkas, J. Dyer, and H. Heckman: "Resolution of the Σ^--Mass Anomaly". *Physical Review Letters* **11**(1) (1963), 26–undefined. DOI: 10.1103/PhysRevLett.11.26.

[53] M. Berger, J. Coursey, M. Zucker, and J. Chang: *ESTAR, PSTAR, and ASTAR: Computer Programs for Calculating Stopping-Power and Range Tables for Electrons, Protons, and Helium Ions (version 1.2.3).* 2005. URL: https://physics.nist.gov/PhysRefData/Star/Text/PSTAR.html.

[54] R. W. Schulte, S. N. Penfold, J. T. Tafas, and K. E. Schubert: "A maximum likelihood proton path formalism for application in proton computed tomography". *Medical Physics* **35**(11) (2008), 4849–4856. ISSN: 00942405. DOI: 10.1118/1.2986139.

[55] C. Tschalär: "Straggling distributions of large energy losses". *Nuclear Instruments and Methods* **61**(2) (1968), 141–156. ISSN: 0029-554X. DOI: 10.1016/0029-554X(68)90535-1.

[56] V. A. Bashkirov, R. W. Schulte, R. F. Hurley, R. P. Johnson, H. F.-W. Sadrozinski, A Zatserklyaniy, T. Plautz, and V. Giacometti: "Novel scintillation detector design and performance for proton radiography and computed tomography". *Medical Physics* **43**(2) (2016), 664–674. ISSN: 00942405. DOI: 10.1118/1.4939255.

[57] J. Janni: "Proton range-energy tables, 1 keV-10GeV". *Atomic Data and Nuclear Data Tables* **27** (1982), 341–529.

[58] L. Eyges: "Multiple Scattering with Energy Loss". *Physical Review* **74**(10) (1948), 1534–1535. DOI: 10.1103/PhysRev.74.1534.

[59] V. L. Highland: "Some practical remarks on multiple scattering". *Nuclear Instruments and Methods* **129**(2) (1975), 497–499. ISSN: 0029-554X. DOI: 10.1016/0029-554X(75)90743-0.

[60] G. R. Lynch and O. I. Dahl: "Approximations to multiple Coulomb scattering". *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* **58**(1) (1991), 6–10. ISSN: 0168-583X. DOI: 10.1016/0168-583X(91)95671-Y.

[61] S. Agostinelli et al.: "Geant4 - a simulation toolkit". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506**(3) (2003), 250–303. ISSN: 01689002. DOI: 10.1016/S0168-9002(03)01368-8. arXiv: 1005.0727v1.

[62] J. Hubbell and S. Seltzer: *NIST: X-Ray Mass Attenuation Coefficients*. 1996. URL: https://www.nist.gov/pml/x-ray-mass-attenuation-coefficients.

[63] A. Einstein: "Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt". *Annalen der Physik* **322**(6) (1905), 132–148. ISSN: 00033804. DOI: 10.1002/andp.19053220607.

[64] A. H. Compton: "A Quantum Theory of the Scattering of X-rays by Light Elements". *Physical Review* **21**(5) (1923), 483–502. ISSN: 0031-899X. DOI: 10.1103/PhysRev.21.483.

[65] M. Krämer, O. Jäkel, T. Haberer, G. Kraft, D. Schardt, and U. Weber: "Treatment planning for heavy-ion radiotherapy: physical beam model and dose optimization". *Physics in Medicine & Biology* **45**(11) (2000), 3299–3317. ISSN: 0031-9155. DOI: 10.1088/0031-9155/45/11/313.

[66] S. Warren, M. Partridge, A. Bolsi, A. J. Lomax, C. Hurt, T. Crosby, and M. A. Hawkins: "An Analysis of Plan Robustness for Esophageal Tumors: Comparing Volumetric Modulated Arc Therapy Plans and Spot Scanning Proton Planning". *International Journal of Radiation Oncology*Biology*Physics* **95**(1) (2016). ISSN: 03603016. DOI: 10.1016/j.ijrobp.2016.01.044.

[67] C. Goma, I. P. Almeida, and F. Verhaegen: "Revisiting the single-energy CT calibration for proton therapy treatment planning: A critical look at the stoichiometric method". *Physics in Medicine & Biology* **63**(23) (2018). ISSN: 13616560. DOI: 10.1088/1361-6560/aaede5.

[68] M. Yang, X. R. Zhu, P. C. Park, U. Titt, R. Mohan, G. Virshup, J. E. Clayton, and L. Dong: "Comprehensive analysis of proton range uncertainties related to patient stopping-power-ratio estimation using the stoichiometric calibration". *Physics in Medicine & Biology* **57**(13) (2012), 4095–4115. ISSN: 0031-9155. DOI: 10.1088/0031-9155/57/13/4095. arXiv: NIHMS150003.

[69] H. Paganetti: "Range uncertainties in proton therapy and the role of Monte Carlo simulations". *Physics in Medicine & Biology* **57**(11) (2012), R99–R117. ISSN: 0031-9155. DOI: 10.1088/0031-9155/57/11/R99. arXiv: 1510.03300.

[70] P. Wohlfahrt, C. Möhler, V. Hietschold, S. Menkel, S. Greilich, M. Krause, M. Baumann, W. Enghardt, and C. Richter: "Clinical Implementation of Dual-energy CT for Proton Treatment Planning on Pseudo-monoenergetic CT scans". *International Journal of Radiation Oncology Biology Physics* **97**(2) (2017), 427–434. ISSN: 1879355X. DOI: 10.1016/j.ijrobp.2016.10.022.

[71] G. Landry, J. Seco, M. Gaudreault, and F. Verhaegen: "Deriving effective atomic numbers from DECT based on a parameterization of the ratio of high and low linear attenuation coefficients". *Physics in Medicine & Biology* **58**(19) (2013), 6851–6866. ISSN: 00319155. DOI: 10.1088/0031-9155/58/19/6851.

[72] A. E. Bourque, J.-F. Carrier, and H. Bouchard: "A stoichiometric calibration method for dual energy computed tomography". *Physics in Medicine & Biology* **59**(8) (2014). ISSN: 0031-9155. DOI: 10.1088/0031-9155/59/8/2059.

[73] A. E. Bourque, J. F. Carrier, and H. Bouchard: "Erratum: A stoichiometric calibration method for dual energy computed tomography". *Physics in Medicine & Biology* **59**(18) (2014), 5611–5612. ISSN: 13616560. DOI: 10.1088/0031-9155/59/18/5611.

[74] M. Yang, G. Virshup, J. Clayton, X. R. Zhu, R. Mohan, and L. Dong: "Theoretical variance analysis of single- and dual-energy computed tomography methods for calculating proton stopping power ratios of biological tissues". *Physics in Medicine & Biology* **55**(5) (2010), 1343–1362. ISSN: 00319155. DOI: 10.1088/0031-9155/55/5/006.

[75] N. Hünemohr, B. Krauss, C. Tremmel, B. Ackermann, O. Jäkel, and S. Greilich: "Experimental verification of ion stopping power prediction from dual energy CT data in tissue surrogates". *Physics in Medicine & Biology* **59**(1) (2014), 83–96. ISSN: 00319155. DOI: 10.1088/0031-9155/59/1/83.

[76] V. T. Taasti, G. J. Michalak, D. C. Hansen, A. J. Deisher, J. J. Kruse, B. Krauss, L. P. Muren, J. B. Petersen, and C. H. McCollough: "Validation of proton stopping power ratio estimation based on dual energy CT using fresh tissue samples". *Physics in Medicine & Biology* **63**(1) (2018). ISSN: 13616560. DOI: 10.1088/1361-6560/aa952f.

[77] E. Bär, A. Lalonde, R. Zhang, K. W. Jee, K. Yang, G. Sharp, B. Liu, G. Royle, H. Bouchard, and H. M. Lu: "Experimental validation of two dual-energy CT methods for proton therapy using heterogeneous tissue samples". *Medical Physics* **45**(1) (2018), 48–59. ISSN: 00942405. DOI: 10.1002/mp.12666.

[78] K. B. Niepel, M. Stanislawski, M. Würl, F. Dörringer, M. Pinto, O. Dietrich, B. Ertl-Wagner, A. Lalonde, H. Bouchard, E. Pappas, I. Yohannes, M. Hillbrand, G. Landry, and K. Parodi: "Animal tissue-based quantitative comparison of dual-energy CT to SPR conversion methods using high-resolution gel dosimetry". *Physics in Medicine & Biology* (2020). ISSN: 0031-9155. DOI: 10.1088/1361-6560/abbd14.

[79] K. Taguchi and J. S. Iwanczyk: "Vision 20/20: Single photon counting x-ray detectors in medical imaging". *Medical Physics* **40**(10) (2013), 100901. ISSN: 00942405. DOI: 10.1118/1.4820371.

[80] V. T. Taasti, D. C. Hansen, G. J. Michalak, A. J. Deisher, J. J. Kruse, L. P. Muren, J. B. Petersen, and C. H. McCollough: "Theoretical and experimental analysis of photon counting detector CT for proton stopping power prediction". *Medical Physics* **45**(11) (2018), 5186–5196. ISSN: 00942405. DOI: 10.1002/mp.13173.

[81] H.-W. Sadrozinski, T. Geoghegan, E. Harvey, R. Johnson, T. Plautz, A. Zatserklyaniy, V. Bashkirov, R. Hurley, P. Piersimoni, R. Schulte, P. Karbasi, K. Schubert, B. Schultze, and V. Giacometti: "Operation of the preclinical head scanner for proton CT". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **831** (2016), 394–399. ISSN: 01689002. DOI: 10.1016/j.nima.2016.02.001.

[82] R. P. Johnson, V. Bashkirov, L. DeWitt, V. Giacometti, R. F. Hurley, P. Piersimoni, T. E. Plautz, H. F. Sadrozinski, K. Schubert, R. Schulte, B. Schultze, and A. Zatserklyaniy: "A Fast Experimental Scanner for Proton CT: Technical Performance and First Experience With Phantom Scans". *IEEE Transactions on Nuclear Science* **63**(1) (2016), 52–60. ISSN: 0018-9499. DOI: 10.1109/TNS.2015.2491918.

[83] P. Piersimoni, B. A. Faddegon, J. R. Méndez, R. W. Schulte, L. Volz, and J. Seco: "Helium CT: Monte Carlo simulation results for an ideal source and detector with comparison to proton CT". *Medical Physics* **45**(7) (2018), 3264–3274. ISSN: 0094-2405. DOI: 10.1002/mp.12942.

[84] L. Volz, C. A. Collins-Fekete, P. Piersimoni, R. P. Johnson, V. Bashkirov, R. Schulte, and J. Seco: "Stopping power accuracy and achievable spatial resolution of helium ion imaging using a prototype particle CT detector system". *Current Directions in Biomedical Engineering* **3**(2) (2017), 401–404. ISSN: 23645504. DOI: 10.1515/cdbme-2017-0084.

[85] L. Volz, P. Piersimoni, V. A. Bashkirov, S. Brons, C.-A. Collins-Fekete, R. P. Johnson, R. W. Schulte, and J. Seco: "The impact of secondary fragments on the image quality of helium ion imaging". *Physics in Medicine & Biology* **63**(19) (2018), 195016. ISSN: 1361-6560. DOI: 10.1088/1361-6560/aadf25.

[86] I. Rinaldi, S. Brons, J. Gordon, R. Panse, B. Voss, O. Jäkel, and K. Parodi: "Experimental characterization of a prototype detector system for carbon ion radiography and tomography". *Physics in Medicine & Biology* **58**(3) (2013), 413–427. ISSN: 0031-9155. DOI: 10.1088/0031-9155/58/3/413.

[87] M. Esposito, C. Waltham, J. T. Taylor, S. Manger, B. Phoenix, T. Price, G. Poludniowski, S. Green, P. M. Evans, P. P. Allport, S. Manolopoulos, J. Nieto-Camero, J. Symons, and N. M. Allinson: "PRaVDA: The first solid-state system for proton computed tomography". *Physica Medica* **55** (2018), 149–154. ISSN: 11201797. DOI: 10.1016/j.ejmp.2018.10.020.

[88] C. Civinini, M. Scaringella, M. Brianzi, M. Intravaia, N. Randazzo, V. Sipala, M. Rovituso, F. Tommasino, M. Schwarz, and M. Bruzzi: "Relative stopping power measurements and prosthesis artifacts reduction in proton CT". *Physics in Medicine & Biology* **65** (2020), 225012. ISSN: 0031-9155. DOI: 10.1088/1361-6560/abb0c8.

[89] E. A. DeJongh, D. F. DeJongh, I. Polnyi, V. Rykalin, C. Sarosiek, G. Coutrakon, K. L. Duffin, N. T. Karonis, C. E. Ordoñez, M. Pankuch, J. R. Winans, and J. S. Welsh: "Technical Note: A fast and monolithic prototype clinical proton radiography system optimized for pencil beam scanning". *Medical Physics* (2020). ISSN: 0094-2405. DOI: 10.1002/mp.14700.

[90] R. Zhang, K. W. Jee, E. Cascio, G. C. Sharp, J. B. Flanz, and H. M. Lu: "Improvement of single detector proton radiography by incorporating intensity of time-resolved dose rate functions". *Physics in Medicine & Biology* **63**(1) (2018). ISSN: 13616560. DOI: 10.1088/1361-6560/aa9913.

[91] N. Krah, L. D. Marzi, A. Patriarca, G. Pittá, and I. Rinaldi: "Proton radiography with a commercial range telescope detector using dedicated post processing methods". *Physics in Medicine & Biology* **63**(20) (2018). ISSN: 13616560. DOI: 10.1088/1361-6560/aae043.

[92] S. Meyer: "On the Clinical Potential of Ion Computed Tomography with Different Detector Systems and Ion Species". Ludwig-Maximilians-Universität München, 2019.

[93] S. Meyer, C. Gianoli, L. Magallanes, B. Kopp, T. Tessonnier, G. Landry, G. Dedes, B. Voss, and K. Parodi: "Comparative Monte Carlo study on the performance of integration- and list-mode detector configurations for carbon ion computed tomography". *Physics in Medicine & Biology* **62**(3) (2017), 1096–1112. ISSN: 0031-9155. DOI: 10.1088/1361-6560/aa5602.

[94] C. Gianoli, S. Meyer, L. Magallanes, C. Paganelli, G. Baroni, and K. Parodi: "Analytical simulator of proton radiography and tomography for different detector configurations". *Physica Medica* **59** (2019), 92–99. ISSN: 1724191X. DOI: 10.1016/j.ejmp.2019.03.002.

[95] K. Parodi, W. Assmann, C. Belka, J. Bortfeldt, D. A. Clevert, G. Dedes, R. Kalunga, S. Kundel, N. Kurichiyanil, P. Lämmer, J. Lascaud, K. Lauber, G. Lovatti, S. Meyer, M. Nitta, M. Pinto, M. J. Safari, K. Schnürle, J. Schreiber, P. G. Thirolf, H. P. Wieser, and M. Würl: "Towards a novel small animal proton irradiation platform: the SIRMIO project". *Acta Oncologica* **58**(10) (2019), 1470–1475. ISSN: 1651226X. DOI: 10.1080/0284186X.2019.1630752.

[96] S. Meyer, J. Bortfeldt, P. L. mmer, F. S. Englbrecht, M. Pinto, K. Schnurle, M. Wurl, and K. Parodi: "Optimization and performance study of a proton CT system for pre-clinical small animal imaging". *Physics in Medicine & Biology* **65**(15) (2020). ISSN: 13616560. DOI: 10.1088/1361-6560/ab8afc.

[97] H. F. Sadrozinski, R. P. Johnson, S. MacAfee, A. Plumb, D. Steinberg, A. Zatserklyaniy, V. A. Bashkirov, R. F. Hurley, and R. W. Schulte: "Development of a head scanner for proton CT". **699**. 2013, 205–210. DOI: 10.1016/j.nima.2012.04.029.

[98] R. Johnson, V. Bashkirov, G. Coutrakon, V. Giacometti, P. Karbasi, N. Karonis, C. Ordoñez, M. Pankuch, H.-W. Sadrozinski, K. Schubert, and R. Schulte: "Results from a Prototype Proton-CT Head Scanner". *Physics Procedia* **90**(8) (2017), 209–214. ISSN: 18753892. DOI: 10.1016/j.phpro.2017.09.060.

[99] J. Dickmann, C. Sarosiek, V. Rykalin, M. Pankuch, S. Rit, N. Detrich, G. B. Coutrakon, R. P. Johnson, R. W. Schulte, K. Parodi, G. Landry, and G. Dedes: "Experimental realization of dynamic fluence field optimization for proton computed tomography". *Physics in Medicine & Biology* **65**(19) (2020), 195001. DOI: 10.1088/1361-6560/ab9f5f.

[100] J. Dickmann, P. Wesp, M. Rädler, S. Rit, M. Pankuch, R. P. Johnson, V Bashkirov, R. W. Schulte, K. Parodi, G. Landry, and G. Dedes: "Prediction of image noise contributions in proton computed tomography and comparison to measurements". *Physics in Medicine & Biology* **64**(14) (2019), 145016. ISSN: 1361-6560. DOI: 10.1088/1361-6560/ab2474.

[101] A. Artikov, J. Budagov, I. Chirikov-Zorin, D. Chokheli, M. Lyablin, G. Bellettini, A. Menzione, S. Tokar, N. Giokaris, and A. Manousakis-Katsikakis: "Properties of the Ukraine polystyrene-based plastic scintillator UPS 923A". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **555**(1-2) (2005), 125–131. ISSN: 01689002. DOI: 10.1016/j.nima.2005.09.021.

[102] G. Dedes, J. Dickmann, V. Giacometti, S. Rit, N. Krah, S. Meyer, V. Bashkirov, R. Schulte, R. P. Johnson, K. Parodi, and G. Landry: "The role of Monte Carlo simulation in understanding the performance of proton computed tomography". *Zeitschrift für Medizinische Physik* (available online, 2020). ISSN: 0939-3889. DOI: 10.1016/j.zemedi.2020.06.006.

[103] V. Giacometti, V. A. Bashkirov, P. Piersimoni, S. Guatelli, T. E. Plautz, H. F. Sadrozinski, R. P. Johnson, A. Zatserklyaniy, T. Tessonnier, K. Parodi, A. B. Rosenfeld, and R. W. Schulte: "Software platform for simulation of a prototype proton CT scanner". *Medical Physics* **44**(3) (2017), 1002–1016. ISSN: 00942405. DOI: 10.1002/mp.12107.

[104] D. F. DeJongh, E. A. DeJongh, V. Rykalin, G. DeFillippo, M. Pankuch, A. W. Best, G. Coutrakon, K. L. Duffin, N. T. Karonis, C. E. Ordoñez, C. Sarosiek, R. W. Schulte, J. R. Winans, A. M. Block, C. L. Hentz, and J. S. Welsh: "A Comparison of Proton Stopping Power Measured with Proton CT and X-Ray CT in Fresh Post-Mortem Porcine Structures" (2020). arXiv: 2012.06629 [physics.med-ph].

[105] J. T. Taylor, G Poludniowski, T Price, C Waltham, P. P. Allport, G. L. Casse, M Esposito, P. M. Evans, S. Green, S Manger, S Manolopoulos, J. Nieto-Camero, D. J. Parker, J Symons, and N. M. Allinson: "An experimental demonstration of a new type of proton computed tomography using a novel silicon tracking detector". *Medical Physics* **43**(11) (2016), 6129–6136. ISSN: 00942405. DOI: 10.1118/1.4965809.

[106] H. Pettersen, J. Alme, A. Biegun, A. van den Brink, M. Chaar, D. Fehlker, I. Meric, O. Odland, T. Peitzmann, E. Rocco, K. Ullaland, H. Wang, S. Yang, C. Zhang, and D. Röhrich: "Proton tracking in a high-granularity Digital Tracking Calorimeter for proton CT purposes". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **860** (2017), 51–61. ISSN: 01689002. DOI: 10.1016/j.nima.2017.02.007. arXiv: 1611.02031.

[107] H. E. S. Pettersen, J. Alme, G. G. Barnaföldi, R. Barthel, A. van den Brink, M. Chaar, V. Eikeland, A. García-Santos, G. Genov, S. Grimstad, O. Grøttvik, H. Helstrup, K. F. Hetland, S. Mehendale, I. Meric, O. H. Odland, G. Papp, T. Peitzmann, P. Piersimoni, A. Ur Rehman, M. Richter, A. T. Samnøy, J. Seco, H. Shafiee, E. V. Skjæveland, J. R. Sølie, G. Tambave, K. Ullaland, M. Varga-Kofarago, L. Volz, B. Wagner, S. Yang, and D. Röhrich: "Design optimization of a pixel-based range telescope for proton computed tomography". *Physica Medica* **63** (2019), 87–97. ISSN: 11201797. DOI: 10.1016/j.ejmp.2019.05.026.

[108] N. Krah, F. Khellaf, J. M. Létang, S. Rit, and I. Rinaldi: "A comprehensive theoretical comparison of proton imaging set-ups in terms of spatial resolution". *Physics in Medicine & Biology* **63**(13) (2018), 135013. ISSN: 1361-6560. DOI: 10.1088/1361-6560/aaca1f.

[109] F. Ulrich-Pur, T. Bergauer, A. Burker, S. Hatamikia, A. Hirtl, C. Irmler, S. Kaser, P. Paulitsch, F. Pitters, and V. Teufelhart: "Imaging with protons at MedAustron". *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **978** (2020). ISSN: 01689002. DOI: 10.1016/j.nima.2020.164407.

[110] R. F. Hurley, R. W. Schulte, V. A. Bashkirov, A. J. Wroe, A. Ghebremedhin, H. F. Sadrozinski, V. Rykalin, G. Coutrakon, P. Koss, and B. Patyal: "Water-equivalent path length calibration of a prototype proton CT scanner". *Medical Physics* **39**(5) (2012), 2438–2446. ISSN: 00942405. DOI: 10.1118/1.3700173.

[111] G. Coutrakon, V. Bashkirov, F. Hurley, R. Johnson, V. Rykalin, H. Sadrozinski, and R. Schulte: "Design and construction of the 1st proton CT scanner". *AIP Conference Proceedings* **1525** (2013), 327–331. ISSN: 0094243X. DOI: 10.1063/1.4802343.

[112] N. Krah, C. T. Quiñones, J.-M. Letang, and S. Rit: "Scattering proton CT". *Physics in Medicine & Biology* **65**(22) (2020), 225015. ISSN: 0031-9155. DOI: 10.1088/1361-6560/abbd18.

[113] S. Rit, G. Dedes, N. Freud, D. Sarrut, and J. M. Létang: "Filtered backprojection proton CT reconstruction along most likely paths". *Medical Physics* **40**(3) (2013), 031103. ISSN: 00942405. DOI: 10.1118/1.4789589.

[114] F. Khellaf, N. Krah, J.-M. Letang, C.-A. Collins-Fekete, and S. Rit: "A comparison of direct reconstruction algorithms in proton computed tomography". *Physics in Medicine & Biology* **65**(10) (2020), 105010. ISSN: 0031-9155. DOI: 10.1088/1361-6560/ab7d53.

[115] C.-A. Collins-Fekete, S. Brousmiche, S. K. N. Portillo, L. Beaulieu, and J. Seco: "A maximum likelihood method for high resolution proton radiography/proton CT". *Physics in Medicine & Biology* **61**(23) (2016), 8232–8248. ISSN: 0031-9155. DOI: 10.1088/0031-9155/61/23/8232.

[116] C.-A. Collins-Fekete, L. Volz, S. K. N. Portillo, L. Beaulieu, and J. Seco: "A theoretical framework to predict the most likely ion path in particle imaging". *Physics in Medicine & Biology* **62**(5) (2017), 1777–1790. ISSN: 0031-9155. DOI: 10.1088/1361-6560/aa58ce. arXiv: 1610.05774.

[117] N. Krah, J.-M. Letang, and S. Rit: "Polynomial modelling of proton trajectories in homogeneous media for fast most likely path estimation and trajectory simulation". *Physics in Medicine & Biology* **64**(19) (2019). ISSN: 0031-9155. DOI: 10.1088/1361-6560/ab3d0b.

[118] D. Lazos, C.-A. Collins-Fekete, P. M. Evans, and N. Dikaios: "Molière maximum likelihood proton path estimation approximated by cubic Bézier curve for scatter corrected proton CT reconstruction". *Physics in Medicine & Biology* **65** (2020), 175003. ISSN: 0031-9155. DOI: 10.1088/1361-6560/ab9413.

[119] F. Khellaf, N. Krah, I. Rinaldi, J.-M. Létang, and S. Rit: "Effects of transverse heterogeneities on the most likely path of protons". *Physics in Medicine & Biology* **65** (2019), 105010. ISSN: 0031-9155. DOI: 10.1088/1361-6560/ab02a8.

[120] C.-A. Collins-Fekete, P. Doolan, M. F. Dias, L. Beaulieu, and J. Seco: "Developing a phenomenological model of the proton trajectory within a heterogeneous medium required for proton imaging". *Physics in Medicine & Biology* **60**(13) (2015), 5071–5082. ISSN: 0031-9155. DOI: 10.1088/0031-9155/60/13/5071.

[121] M. D. Brooke and S. N. Penfold: "An inhomogeneous most likely path formalism for proton computed tomography". *Physica Medica* **70** (2020), 184–195. ISSN: 11201797. DOI: 10.1016/j.ejmp.2020.01.025. arXiv: 1808.00122.

[122] S. Rit, M. V. Oliva, S. Brousmiche, R. Labarbe, D. Sarrut, and G. C. Sharp: "The Reconstruction Toolkit (RTK), an open-source cone-beam CT reconstruction toolkit based on the Insight Toolkit (ITK)". *Journal of Physics: Conference Series* **489** (2014). ISSN: 1742-6588. DOI: 10.1088/1742-6596/489/1/012079.

[123] M. McCormick, X. Liu, J. Jomier, C. Marion, and L. Ibanez: "ITK: enabling reproducible research and open science". *Frontiers in Neuroinformatics* **8** (2014). ISSN: 1662-5196. DOI: 10.3389/fninf.2014.00013.

[124] M. Rädler, G. Landry, S. Rit, R. W. Schulte, K. Parodi, and G. Dedes: "Two-dimensional noise reconstruction in proton computed tomography using distance-driven filtered back-projection of simulated projections". *Physics in Medicine & Biology* **63**(21) (2018), 215009. ISSN: 1361-6560. DOI: 10.1088/1361-6560/aae5c9.

[125] G. N. Ramachandran and A. V. Lakshminarayanan: "Three-dimensional Reconstruction from Radiographs and Electron Micrographs: Application of Convolutions instead of Fourier Transforms". *Proceedings of the National Academy of Sciences* **68**(9) (1971), 2236–2240. ISSN: 0027-8424. DOI: 10.1073/pnas.68.9.2236.

[126] A. C. Kak and M. Slaney: "Principles of Tomographic Imaging" (1987).

[127] L. A. Feldkamp, L. C. Davis, and J. W. Kress: "Practical cone-beam algorithm". *Journal of the Optical Society of America A* **1**(6) (1984), 612–619. DOI: 10.1364/JOSAA.1.000612.

[128] D. C. Hansen, T. Sangild Sørensen, and S. Rit: "Fast reconstruction of low dose proton CT by sinogram interpolation". *Physics in Medicine & Biology* **61**(15) (2016), 5868–5882. ISSN: 0031-9155. DOI: 10.1088/0031-9155/61/15/5868.

[129] G. Poludniowski, N. M. Allinson, and P. M. Evans: "Proton computed tomography reconstruction using a backprojection-then-filtering approach". *Physics in Medicine & Biology* **59**(24) (2014), 7905–7918. ISSN: 0031-9155. DOI: 10.1088/0031-9155/59/24/7905.

[130] S. Rit, R. Clackdoyle, J. Hoskovec, and J. M. Letang: "List-mode proton CT reconstruction using their most likely paths via the finite Hilbert transform of the derivative of the backprojection". *Proceedings of Fully3D* (2015), 324–327.

[131] F. Khellaf, N. Krah, J. M. Létang, and S. Rit: "2D directional ramp filter". *Physics in Medicine & Biology* **65**(8) (2020). ISSN: 13616560. DOI: 10.1088/1361-6560/ab7875.

[132] S. N. Penfold, A. B. Rosenfeld, R. W. Schulte, and K. E. Schubert: "A more accurate reconstruction system matrix for quantitative proton computed tomography". *Medical Physics* **36**(10) (2009), 4511–4518. ISSN: 00942405. DOI: 10.1118/1.3218759.

[133] S. N. Penfold, R. W. Schulte, Y. Censor, and A. B. Rosenfeld: "Total variation superiorization schemes in proton computed tomography image reconstruction". *Medical Physics* **37**(11) (2010), 5887–5895. ISSN: 00942405. DOI: 10.1118/1.3504603. arXiv: 1010.1663.

[134] B. Schultze, Y. Censor, P. Karbasi, K. E. Schubert, and R. W. Schulte: "An Improved Method of Total Variation Superiorization Applied to Reconstruction in Proton Computed Tomography". *IEEE Transactions on Medical Imaging* **39**(2) (2020), 294–307. DOI: 10.1109/TMI.2019.2911482.

[135] D. F. DeJongh and E. A. DeJongh: "An Iterative Least Squares Method for Proton CT Image Reconstruction" (2020). arXiv: 2009.14263 [physics.med-ph].

[136] A. Wunderlich and F. Noo: "Image covariance and lesion detectability in direct fan-beam x-ray computed tomography". *Physics in Medicine & Biology* **53**(10) (2008), 2471–2493. ISSN: 0031-9155. DOI: 10.1088/0031-9155/53/10/002.

[137] J. C. Gore and P. S. Tofts: "Statistical limitations in computed tomography". *Physics in Medicine & Biology* **23**(6) (1978), 1176–1182. DOI: 10.1088/0031-9155/23/6/014.

[138] J. Dickmann, F. Kamp, M. Hillbrand, S. Corradini, C. Belka, R. W. Schulte, K. Parodi, G. Dedes, and G. Landry: "Fluence-modulated proton CT optimized with patient-specific dose and variance objectives for proton dose calculation". *Physics in Medicine & Biology* **66**(6) (2021), 064001. DOI: 10.1088/1361-6560/abe3d2.

[139] S. S. Hsieh and N. J. Pelc: "The piecewise-linear dynamic attenuator reduces the impact of count rate loss with photon-counting detectors". *Physics in Medicine & Biology* **59**(11) (2014), 2829–2847. ISSN: 0031-9155. DOI: 10.1088/0031-9155/59/11/2829.

[140] V. Giacometti, S. Guatelli, M. Bazalova-Carter, A. Rosenfeld, and R. Schulte: "Development of a high resolution voxelised head phantom for medical physics applications". *Physica Medica* **33** (2017), 182–188. ISSN: 11201797. DOI: 10.1016/j.ejmp.2017.01.007.

[141] J. Dickmann, C. Sarosiek, S. Götz, M. Pankuch, G. Coutrakon, R. P. Johnson, R. W. Schulte, K. Parodi, G. Landry, and G. Dedes: "An empirical artifact correction for proton computed tomography". *submitted for publication* (2021).

[142] G. Battistoni, J. Bauer, T. T. Boehlen, F. Cerutti, M. P. W. Chin, R. D. S. Augusto, A. Ferrari, P. G. Ortega, W. Kozłowska, G. Magro, A. Mairani, K. Parodi, P. R. Sala, P. Schoofs, T. Tessonnier, and V. Vlachoudis: "The FLUKA Code: An Accurate Simulation Tool for Particle Therapy". *Frontiers in Oncology* **6** (2016). ISSN: 2234-943X. DOI: 10.3389/fonc.2016.00116.

[143] J. Perl, J. Shin, J. Schümann, B. Faddegon, and H. Paganetti: "TOPAS: An innovative proton Monte Carlo platform for research and clinical applications". *Medical Physics* **39**(11) (2012), 6818–6837. ISSN: 00942405. DOI: 10.1118/1.4758060.

[144] S. Jan, G. Santin, D. Strul, S. Staelens, K. Assié, D. Autret, S. Avner, R. Barbier, M. Bardiès, P. M. Bloomfield, D. Brasse, V. Breton, P. Bruyndonckx, I. Buvat, A. F. Chatziioannou, Y. Choi, Y. H. Chung, C. Comtat, D. Donnarieix, L. Ferrer, S. J. Glick, C. J. Groiselle, D. Guez, P. F. Honore, S. Kerhoas-Cavata, A. S. Kirov, V. Kohli, M. Koole, M. Krieguer, D. J. van der Laan, F. Lamare, G. Largeron, C. Lartizien, D. Lazaro, M. C. Maas, L. Maigne, F. Mayet, F. Melot, C. Merheb, E. Pennacchio, J. Perez, U. Pietrzyk, F. R. Rannou, M. Rey, D. R. Schaart, C. R. Schmidtlein, L. Simon, T. Y. Song, J. M. Vieira, D. Visvikis, R. V. de Walle, E. Wieërs, and C. Morel: "GATE: A simulation toolkit for PET and SPECT". *Physics in Medicine & Biology* **49**(19) (2004), 4543–4561. ISSN: 00319155. DOI: 10.1088/0031-9155/49/19/007.

[145] A. Resch, G. Landry, F. Kamp, G. Cabal, C. Belka, J. Wilkens, K. Parodi, and G. Dedes: "Quantification of the uncertainties of a biological model and their impact on variable RBE proton treatment plan optimization". *Physica Medica* **36** (2017), 91–102. ISSN: 11201797. DOI: 10.1016/j.ejmp.2017.03.013.

[146] S. Schmid, G. Landry, C. Thieke, F. Verhaegen, U. Ganswindt, C. Belka, K. Parodi, and G. Dedes: "Monte Carlo study on the sensitivity of prompt gamma imaging to proton range variations due to interfractional changes in prostate cancer patients". *Physics in Medicine & Biology* **60**(24) (2015), 9329–9347. ISSN: 0031-9155. DOI: 10.1088/0031-9155/60/24/9329.

[147] J. Dickmann, G. Landry, P. Wesp, S. Rit, M. Pankuch, R. Johnson, V. Bashkirov, R. Schulte, K. Parodi, and G. Dedes: "Modelling of contributions to image variance in proton CT for application in low-dose fluence-modulated imaging". *Conference of the Particle Therapy Cooperative Group (PTCOG), Manchester, United Kingdom* (2019).

[148] J. Dickmann, G. Landry, P. Wesp, S. Rit, M. Pankuch, R. Johnson, V. Bashkirov, R. Schulte, K. Parodi, and G. Dedes: "Monte Carlo study of image noise contributions of a prototype proton computed tomography scanner". *Jagiellonian Symposium on Fundamental and Applied Subatomic Physics, Krakow, Poland* (2019).

[149] J. B. Birks: "Scintillations from Organic Crystals: Specific Fluorescence and Relative Response to Different Radiations". *Proceedings of the Physical Society. Section A* **64**(10) (1951), 874–877. ISSN: 0370-1298. DOI: 10.1088/0370-1298/64/10/303.

[150] C. G. Broyden: "The Convergence of a Class of Double-rank Minimization Algorithms". *IMA Journal of Applied Mathematics* **6**(3) (1970), 222–231. ISSN: 0272-4960. DOI: 10.1093/imamat/6.3.222.

[151] M. D. Harpen: "A simple theorem relating noise and patient dose in computed tomography". *Medical Physics* **26**(11) (1999), 2231–2234. ISSN: 00942405. DOI: 10.1118/1.598778.

[152] S. A. Graham, D. J. Moseley, J. H. Siewerdsen, and D. A. Jaffray: "Compensators for dose and scatter management in cone-beam computed tomography". *Medical Physics* **34**(7) (2007), 2691–2703. ISSN: 00942405. DOI: 10.1118/1.2740466.

[153] V. I. Tretyak: "Semi-empirical calculation of quenching factors for ions in scintillators". *Astroparticle Physics* **33**(1) (2010), 40–53. DOI: 10.1016/j.astropartphys.2009.11.002. arXiv: 0911.3041.

[154] L. Reichhart, D. Y. Akimov, H. M. Araújo, E. J. Barnes, V. A. Belov, A. A. Burenkov, V. Chepel, A. Currie, L. Deviveiros, B. Edwards, V. Francis, C. Ghag, A. Hollingsworth, M. Horn, G. E. Kalmus, A. S. Kobyakin, A. G. Kovalenko, V. N. Lebedenko, A. Lindote, M. I. Lopes, R. Lüscher, P. Majewski, A. S. J. Murphy, F. Neves, S. M. Paling, J. P. D. Cunha, R. Preece, J. J. Quenby, P. R. Scovell, C. Silva, V. N. Solovov, N. J. Smith, P. F. Smith, V. N. Stekhanov, T. J. Sumner, C. Thorne, and R. J. Walker: "Quenching factor for low-energy nuclear recoils in a plastic scintillator". *Physical Review C - Nuclear Physics* **85**(6) (2012). ISSN: 1089490X. DOI: 10.1103/PhysRevC.85.065801.

[155] J. M. Schippers: "Beam-Transport Systems for Particle Therapy". *Proceedings of the CAS-CERN Accelerator School: Accelerators for Medical Applications* (2017), 241–252. DOI: 10.23730/CYRSP-2017-001.241.

[156] J. Dickmann, S. Rit, M. Pankuch, R. P. Johnson, R. W. Schulte, K. Parodi, G. Dedes, and G. Landry: "An optimization algorithm for dose reduction with fluence-modulated proton CT". *Medical Physics* **47**(4) (2020), 1895–1906. ISSN: 0094-2405. DOI: 10.1002/mp.14084.

[157] J. Dickmann, P. Wesp, S. Rit, M. Pankuch, R. P. Johnson, V. A. Bashkirov, R. W. Schulte, K. Parodi, G. Dedes, and G. Landry: "Method for Fluence Field Optimization to Achieve Nonconvex Image Noise Prescriptions with Fluence-Modulated Proton CT". *Annual Meeting of the American Association of Physicists in Medicine (AAPM), San Antonio, USA, proceedings published in Medical Physics* **46**(6) (2019), E477–E477. ISSN: 24734209. DOI: 10.1002/mp.13589.

[158] J. Dickmann, P. Wesp, S. Rit, M. Pankuch, R. Johnson, V. Bashkirov, R. Schulte, K. Parodi, G. Dedes, and G. Landry: "Dosiseinsparung durch fluenzmodulierte Protonen-Computertomographie mit Anwendung in der Protonentherapie". *Jahrestagung der Deutschen Gesellschaft für Medizinische Physik (DGMP), Stuttgart, Germany* (2019).

[159] J. Dickmann, C. Sarosiek, G. Coutrakon, S. Rit, N. Detrich, V. Rykalin, M. Pankuch, R. P. Johnson, R. W. Schulte, K. Parodi, G. Dedes, and G. Landry: "PO-1706: Low dose fluence-modulated proton CT: simulation study and first experimental results". *Conference of the European Society for Radiotherapy and Oncology (ESTRO), Vienna, Austria (online), proceedings published in Radiotherapy and Oncology* **152**(S1) (2020), S941. DOI: 10.1016/S0167-8140(21)01724-2.

[160] T. Toth, Z. Ge, and M. P. Daly: "The influence of patient centering on CT dose and image noise". *Medical Physics* **34**(7) (2007), 3093–3101. ISSN: 00942405. DOI: 10.1118/1.2748113.

[161] S. A. Graham, J. H. Siewerdsen, and D. A. Jaffray: "Intensity-modulated fluence patterns for task-specific imaging in cone-beam CT". *Proceedings of SPIE* **6510** (2007). Ed. by J. Hsieh and M. J. Flynn, 651003. DOI: 10.1117/12.713724.

[162] S. Bartolac, S. Graham, J. Siewerdsen, and D. Jaffray: "Fluence field optimization for noise and dose objectives in CT". *Medical Physics* **38**(7) (2011), S2–S17. ISSN: 00942405. DOI: 10.1118/1.3574885.

[163] S. Bartolac and D. Jaffray: "Compensator models for fluence field modulated computed tomography". *Medical Physics* **40**(12) (2013), 121909. ISSN: 00942405. DOI: 10.1118/1.4829513.

[164] S. S. Hsieh and N. J. Pelc: "Control algorithms for dynamic attenuators". *Medical Physics* **41**(6) (2014), 061907. ISSN: 00942405. DOI: 10.1118/1.4875727.

[165] G. J. Gang, J. H. Siewerdsen, and J. W. Stayman: "Task-Driven Optimization of Fluence Field and Regularization for Model-Based Iterative Reconstruction in Computed Tomography". *IEEE Transactions on Medical Imaging* **36**(12) (2017), 2424–2435. ISSN: 0278-0062. DOI: 10.1109/TMI.2017.2763538.

[166] G. J. Gang and J. W. Stayman: "Joint optimization of fluence field modulation and regularization for multi-task objectives". *Proceedings of SPIE* **10573** (2018), 1057313. DOI: 10.1117/12.2294950.

[167] A. Mao, J. W. Stayman, W. Shyr, and G. J. Gang: "Dynamic beam filtering for miscentered patients". *Proceedings of SPIE* **10573** (2018), 105730U. DOI: 10.1117/12.2293696.

[168] J. W. Stayman, A. Mathews, and R. Levinson: "Method for optimizing radiation beam intensity profile shape using dual multiple aperture devices". *U.S. Patent Application No. 15/639,044* (2018).

[169] G. J. Gang, A. Mao, W. Wang, J. H. Siewerdsen, A. Mathews, S. Kawamoto, R. Levinson, and J. W. Stayman: "Dynamic fluence field modulation in computed tomography using multiple aperture devices". *Physics in Medicine & Biology* **64**(10) (2019), 105024. ISSN: 1361-6560. DOI: 10.1088/1361-6560/ab155e.

[170] W. Wang, G. J. Gang, J. H. Siewerdsen, and J. W. Stayman: "Volume-of-interest imaging using multiple aperture devices". *Proceedings of SPIE* **10948** (2019), 1094823. DOI: 10.1117/12.2513427.

[171] T. P. Szczykutowicz and C. A. Mistretta: "Design of a digital beam attenuation system for computed tomography. Part I. System design and simulation framework". *Medical Physics* **40**(2) (2013), 21905. ISSN: 00942405. DOI: 10.1118/1.4773880.

[172] T. P. Szczykutowicz and C. A. Mistretta: "Experimental realization of fluence field modulated CT using digital beam attenuation". *Physics in Medicine & Biology* **59**(5) (2014), 1305–1326. ISSN: 00319155. DOI: 10.1088/0031-9155/59/5/1305.

[173] T. P. Szczykutowicz, J. Hermus, M. Geurts, and J. Smilowitz: "Realization of fluence field modulated CT on a clinical TomoTherapy megavoltage CT system". *Physics in Medicine & Biology* **60**(18) (2015), 7245–7257. ISSN: 13616560. DOI: 10.1088/0031-9155/60/18/7245.

[174] J. W. Stayman, A. Mathews, W. Zbijewski, G. Gang, J. Siewerdsen, S. Kawamoto, I. Blevis, and R. Levinson: "Fluence-field modulated x-ray CT using multiple aperture devices". *Proceedings of SPIE* **9783** (2016), 97830X. ISSN: 16057422. DOI: 10.1117/12.2214358. arXiv: 15334406.

[175] S. M. Huck, K. Parodi, and K. Stierstorfer: "First Experimental Validation of a Novel Concept for Dynamic Beam Attenuation in CT". *Proceedings of the 5th International Conference on Image Formation in X-ray Computed Tomography*. 2018, 24–27.

[176] S. M. Huck, G. S. K. Fung, K. Parodi, and K. Stierstorfer: "Technical Note: Sheet-based dynamic beam attenuator – A novel concept for dynamic fluence field modulation in x-ray CT". *Medical Physics* **46**(12) (2019), 5528–5537. DOI: 10.1002/mp.13690.

[177] S. Huck, G. Fung, K. Parodi, and K. Stierstorfer: "Optimized intensity modulation for a dynamic beam attenuator in x-ray computed tomography". *Proceedings of SPIE* **10948** (2019), 1094824. DOI: 10.1117/12.2511704.

[178] P. Shunhavanich, N. J. Pelc, S. S. Hsieh, and N. R. Bennett: "Implementation of a piecewise-linear dynamic attenuator". *Proceedings of SPIE* **10573** (2018), 105730T. DOI: 10.1117/12.2293525.

[179] P. Shunhavanich, S. S. Hsieh, and N. J. Pelc: "Fluid-filled dynamic bowtie filter: Description and comparison with other modulators". *Medical Physics* **46**(1) (2019), 127–139. ISSN: 00942405. DOI: 10.1002/mp.13272.

[180] J. A. Nelder and R. Mead: "A Simplex Method for Function Minimization". *The Computer Journal* **7**(4) (1965), 308–313. ISSN: 0010-4620. DOI: 10.1093/comjnl/7.4.308.

[181] P. M. Joseph: "An Improved Algorithm for Reprojecting Rays through Pixel Images". *IEEE Transactions on Medical Imaging* **1**(3) (1982), 192–196. ISSN: 0278-0062. DOI: 10.1109/TMI.1982.4307572.

[182] J. Dickmann, C. Sarosiek, V. Rykalin, M. Pankuch, S. Rit, N. Detrich, G. Coutrakon, R. Johnson, R. W. Schulte, K. Parodi, G. Landry, and G. Dedes: "Prescribing Image Noise Using Dynamic Fluence Field Optimization: Experimental Results Using a Pre-Clinical Proton CT Scanner". *Joint Annual Meeting of the American Association of Physicists in Medicine and the Canadian Organization of Medical Physics (AAPM/COMP), Vancouver, Canada (online), proceedings published in Medical Physics* **47**(6) (2020), 2580–2674. ISSN: 24734209. DOI: 10.1002/mp.14315.

[183] J. Dickmann, C. Sarosiek, G. Coutrakon, S. Rit, N. Detrich, V. Rykalin, M. Pankuch, R. P. Johnson, R. W. Schulte, K. Parodi, G. Landry, and G. Dedes: "Dynamic Fluence Modulation using Proton CT for Low-dose Imaging in Particle Therapy". *International Conference on Image Formation in X-Ray Computed Tomography (CT Meeting), Regensburg, Germany (online)* (2020).

[184] J. Dickmann, F. Kamp, R. Schulte, K. Parodi, G. Dedes, and G. Landry: "Joint Dose Minimization and Variance Optimization for Fluence-Modulated Proton CT". *IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), Boston, USA (online)* (2020).

[185] T. Bortfeld: "Optimized Planning Using Physical Objectives and Constraints". *Seminars in Radiation Oncology* **9**(1) (1999), 20–34. DOI: 10.1016/S1053-4296(99)80052-6.

[186] J. Markman, D. A. Low, A. W. Beavis, and J. O. Deasy: "Beyond bixels: Generalizing the optimization parameters for intensity modulated radiation therapy". *Medical Physics* **29**(10) (2002), 2298–2304. ISSN: 00942405. DOI: 10.1118/1.1508799.

[187] F. Kamp, D. J. Carlson, and J. J. Wilkens: "Rapid implementation of the repair-misrepair-fixation (RMF) model facilitating online adaption of radiosensitivity parameters in ion therapy". *Physics in Medicine & Biology* **62**(13) (2017), N285–N296. ISSN: 13616560. DOI: 10.1088/1361-6560/aa716b.

[188] C. Zhu, R. H. Byrd, P. Lu, and J. Nocedal: "Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization". *ACM Transactions on Mathematical Software (TOMS)* **23**(4) (1997), 550–560.

[189] G. Guennebaud, B. Jacob, et al.: *Eigen v3*. 2010. URL: http://eigen.tuxfamily.org.

[190] G. A. Ezzell, J. W. Burmeister, N. Dogan, T. J. Losasso, J. G. Mechalakos, D. Mihailidis, A. Molineu, J. R. Palta, C. R. Ramsey, B. J. Salter, J. Shi, P. Xia, N. J. Yue, and Y. Xiao: "IMRT commissioning: Multiple institution planning and dosimetry comparisons, a report from AAPM Task Group 119". *Medical Physics* **36** (11 2009), 5359–5373. ISSN: 00942405. DOI: 10.1118/1.3238104.

[191] J. H. Song, M. J. Kim, S. H. Park, S. R. Lee, M. Y. Lee, D. S. Lee, and T. S. Suh: "Gamma analysis dependence on specified low-dose thresholds for VMAT QA". *Journal of Applied Clinical Medical Physics* **16** (6 2015), 263–272. ISSN: 15269914. DOI: 10.1120/jacmp.v16i6.5696.

[192] D. Mackin, X. R. Zhu, F. Poenisch, H. Li, N. Sahoo, M. Kerr, C. Holmes, Y. Li, M. Lii, R. Wu, K. Suzuki, M. T. Gillin, S. J. Frank, D. Grosshans, and X. Zhang: "Spot-Scanning Proton Therapy Patient-Specific Quality Assurance: Results from 309 Treatment Plans". *International Journal of Particle Therapy* **1**(3) (2014), 711–720. ISSN: 2331-5180. DOI: 10.14338/ijpt-14-00017.1.

[193] L. Koivula, L. Wee, and J. Korhonen: "Feasibility of MRI-only treatment planning for proton therapy in brain and prostate cancers: Dose calculation accuracy in substitute CT images". *Medical Physics* **43**(8) (2016), 4634–4642. ISSN: 00942405. DOI: 10.1118/1.4958677.

[194] P. Yepes, A. Adair, D. Grosshans, D. Mirkovic, F. Poenisch, U. Titt, Q. Wang, and R. Mohan: "Comparison of Monte Carlo and analytical dose computations for intensity modulated proton therapy". *Physics in Medicine & Biology* **63**(4) (2018). ISSN: 13616560. DOI: 10.1088/1361-6560/aaa845.

[195] R. E. Shore: "Radiation and cataract risk: Impact of recent epidemiologic studies on ICRP judgments". *Mutation Research/Reviews in Mutation Research* **770** (2016), 231–237. ISSN: 13835742. DOI: 10.1016/j.mrrev.2016.06.006.

[196] U. Schneider and R. Hälg: "The Impact of Neutrons in Clinical Proton Therapy". *Frontiers in Oncology* **5** (2015). ISSN: 2234-943X. DOI: 10.3389/fonc.2015.00235.

[197] J. Dickmann, C. Sarosiek, V. Rykalin, M. Pankuch, G. Coutrakon, R. P. Johnson, V. Bashkirov, R. W. Schulte, K. Parodi, G. Landry, and G. Dedes: "Proof of concept image artifact reduction by energy-modulated proton computed tomography (EMpCT)". *Physica Medica* **81** (2021), 237–244. ISSN: 11201797. DOI: 10.1016/j.ejmp.2020.12.012.

[198] J. Dickmann, P. Wesp, S. Rit, M. Pankuch, R. Johnson, V. Bashkirov, R. Schulte, K. Parodi, G. Landry, and G. Dedes: "Understanding image artifacts for a prototype proton computed tomography scanner via Monte Carlo simulations". *International Conference on Monte Carlo Techniques for Medical Applications (MCMA), Montréal, Canada* (2019).

[199] M. Saito and S. Sagara: "A simple formulation for deriving effective atomic numbers via electron density calibration from dual-energy CT data in the human body". *Medical Physics* **44**(6) (2017), 2293–2303. ISSN: 00942405. DOI: 10.1002/mp.12176.

[200] C. Ordoñez, N. Karonis, K. Duffin, W. J., E. DeJongh, D. DeJongh, G. Coutrakon, N. Myers, M. Pankuch, and W. J.: "Fast in situ image reconstruction for proton radiography". *Journal of Radiation Oncology* **8**(2) (2019), 185–198. DOI: 10.1007/s13566-019-00387-x.

[201] M. Kachelrieß, K. Sourbelle, and W. A. Kalender: "Empirical cupping correction: A first-order raw data precorrection for cone-beam computed tomography". *Medical Physics* **33**(5) (2006), 1269–1274. ISSN: 00942405. DOI: 10.1118/1.2188076.

[202] J. Park and I. W. Sandberg: "Universal Approximation Using Radial-Basis-Function Networks". *Neural Computation* **3**(2) (1991). ISSN: 0899-7667. DOI: 10.1162/neco.1991.3.2.246.

[203] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, bibinitperiodI. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors: "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python". *Nature Methods* **17** (2020), 261–272. DOI: 10.1038/s41592-019-0686-2.

[204] T. E. Plautz, V. Bashkirov, V. Giacometti, R. F. Hurley, R. P. Johnson, P. Piersimoni, H. F. Sadrozinski, R. W. Schulte, and A. Zatserklyaniy: "An evaluation of spatial resolution of a prototype proton CT scanner". *Medical Physics* **43**(12) (2016), 6291–6300. ISSN: 00942405. DOI: 10.1118/1.4966028.

LIST OF FIGURES

LIST OF TABLES