Advancing computational and imaging methods to reduce uncertainties in ion beam therapy

Dr. rer. nat. Georgios Dedes

München 2021

Advancing computational and imaging methods to reduce uncertainties in ion beam therapy

Dr. rer. nat. Georgios Dedes

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

vorgelegt von
Dr. rer. nat. Georgios Dedes
aus Salamis, Griechenland

München, den 7. Oktober 2021
Date of the Habilitation lecture: July 6th 2022

Fachmentorat:

Prof. Dr. Katia Parodi, Department of Medical Physics, LMU Munich

Prof. Dr. Jochen Weller, University Observatory, LMU Munich

Prof. Dr. Claus Belka, Department of Radiation Oncology, LMU Hospital

Contents

ΤΕΥΚΡΟΣ: . . . ἐς γῆν ἐναλίαν Κύπρον,
οὗ μ’ ἐθέσπισεν οἰκεῖν Ἀπόλλων,
ὄνομα νησιωτικὸν Σαλαμῖνα
θέμενον τῆς ἐκεῖ χάριν πάτρας.
ΕΥΡΙΠΙΔΗΣ, ΕΛΕΝΗ

TEUKROS: . . . zum Meereseiland Kypern,
wo Apollo mich heißt siedeln
und den Inselnamen Salamis der Gründung geben,
meiner Heimat eingedenk.
EURIPIDES, HELENA

Abstract

Research on the treatment of cancer is a topic on which mankind has continuously and persistently focused its efforts and resources since many decades and it will continue being on the forefront of science for the foreseeable future. In our endeavor to understand and treat cancer, numerous scientific fields have been interwoven and combined: medicine, epidemiology, biology, chemistry, engineering, mathematics, artificial intelligence and physics. Radiation therapy is one of the main tools for offering safe and efficient treatments to cancer patients. Based, among others, on our comprehension of particles' interaction with matter and on our capability of precisely controlling and guiding radiation to malignant tumors, it naturally became a very fertile research area in the field of physics.

Radiation therapy using heavy charged particle beams, often simply called particle therapy, has emerged already since the 1950s as a promising technique for treating cancer. Due to the finite range of charged particles in matter, which allows to achieve a very conformal dose distribution to the tumor and spare healthy tissues, combined with often enhanced radiobiological effects, has rendered particle therapy a clinically established treatment modality offered nowadays by hospitals worldwide to thousands of patients. This inherently precise and effective technique is nonetheless very sensitive to a broad spectrum of uncertainties. Anatomical uncertainties about the patient specific treatment site, uncertainties about the exact biological response of living cells to radiation, positioning uncertainties of the patient prior to the treatment delivery and organ motion, are a few examples of different effects that may limit the potential of particle therapy for sparing healthy tissues during irradiation.

This manuscript presents published work that addresses various of the above mentioned uncertainties impacting the quality of treatment in the case of particle therapy. It starts with the basic concepts related to cancer as a disease, the relevant definitions and statistical facts as well as a review of radiation therapy and its imaging. Subsequently, the main physics of particle therapy is outlined. Next, a direct comparison of two imaging modalities, one based on photons and one based on protons, is presented, concluding that both are capable of reducing proton range uncertainties at the treatment planning level below the desired 1%. Continuing to anatomical and positioning uncertainties that have to be accounted for at the day of the treatment, a new concept of acquisition for proton computed tomography that allows for low dose – highly accurate image guidance is described. The theoretical foundations are laid and the potential of this technique is investigated both in detailed simulations and experiments performed in a clinical facility. It is demonstrated

that this new technique retains the high image quality of conventionally acquired proton computed tomography, while reducing dose to healthy tissues by as much as 60%. Considering biological uncertainties, their impact on treatment plan optimization is investigated. The findings indicate that including these uncertainties in the plan optimization procedure can reduce overdosage of healthy tissues. Real–time verification of the delivered dose to the patient is explored by means of secondary prompt photon detection. The currently established nuclear models predicting secondary prompt photon emission are assessed and improved and thereafter the sensitivity of the method in detecting range uncertainties due to anatomical changes is investigated on a number of simulated patients. It is shown that ± 2 mm range shifts inside the patient are detectable with this method. The manuscript closes with the radiation protection study of CALA, the facility for laser–based acceleration of electron, proton and heavy ions for novel biomedical applications.

Chapter 1

General introduction

The aim of this chapter is to introduce the main aspects of cancer as a disease and set the framework for the rest of the manuscript. In the subsequent pages the reader will be presented with: basic definitions of cancer and of radiation oncology, a brief historical review, recent facts and statistics about cancer and a summary of imaging tools and therapeutic strategies. As a transition to the following chapters the basic concepts of radiation therapy with heavy charged particles will be outlined. The chapter concludes with the scientific goals of this work.

1.1 Definitions and brief historical review

According to the World Health Organization (WHO) (WHO 2020), cancer is a large group of diseases that can start in almost any organ or tissue of the body when abnormal cells grow uncontrollably, go beyond their usual boundaries to invade adjoining parts of the body and/or spread to other organs. The latter process is called metastasizing and is a major cause of death from cancer. Other common names for cancer are neoplasm and malignant tumor.

Cooper (2019) provides an informative description of different classifications of types of tumors and cancers. There are more than a hundred distinct types of cancer, which can vary substantially in their behavior and response to treatment. A basic distinction in cancer pathology is categorizing tumors as benign or malignant. Benign tumors remain confined to their original location. Malignant tumors are capable of both invading surrounding normal tissue and metastasizing throughout the body. Only malignant tumors are specified as cancers.

Cancers can be classified according to the type of cell from which they arise and fall into one of three main groups: carcinomas, sarcomas, and leukemias or lymphomas. Carcinomas, account for approximately 90% of human cancers, are malignancies of epithelial cells, namely cells at the outer surfaces of organs and blood vessels. Leukemias and lymphomas, which account for approximately 8% of human cancers, arise from the blood–forming cells and from cells of the immune system, respectively. Finally sarcomas, are solid tumors of

connective tissues, such as muscle, bone, cartilage, and fibrous tissue. Tumors are also characterized according to the organ they are located in, for example lung or breast carcinomas.

Although cancer has become one of the leading causes of death only in the 20$^{\text{th}}$ century, it is by no means a modern disease. The ancient Egyptians have observed malignant tumors and the first description of breast cancer by an Egyptian physician is dated back to 1600 BC. Growths suggestive of bone cancer have been seen in mummies. Cancer received further attention during the classical antiquity and between 460–370 BC Hippocrates coined the term *carcinos* for malignant tumors. This term was later translated in Latin as *cancer* by the Roman physician Celsus (28 BC–50 AD). The Greek physician Galenus (130–200 AD) used the word *oncos* (Greek for swelling or volume) to describe tumors. Breast amputations as a breast cancer treatment were advocated for and reported already in the 16$^{\text{th}}$ century (Sakorafas & Safioleas 2009). In what is regarded as one of the earliest documented epidemiological works, Bernardino Ramazzini observed in 1713 the frequency of occurrence of specific types of cancers in nuns and correlated them to specific behavioral patterns with respect to the general female population. In 1838 Müller was the first to relate cancer to anomalies of the cells (Triolo 1965). In 1896, only one year after the discovery of x–rays by Röntgen, Grubbé possibly made the first application of radiation for cancer treatment (Grubbé 1933). The medical use of ionizing radiation as part of cancer treatment to control malignant cells is nowadays defined as radiation oncology (Gunderson 2016). It is a field that saw tremendous development in the 20$^{\text{th}}$ century, evolving from the early irradiation techniques of Grubbé to sophisticated modern irradiation methods such intensity modulated radiation therapy and particle therapy, the latter being the conceptual basis of this manuscript.

1.2 Cancer facts and statistics

As shown in figures 1.1(a) and (d), in 2020 the estimated total number of new cancer cases worldwide exceeded 19 millions and the mortality approached 10 millions. Due to the anatomical differences between the two genders, some cancer types are common to both genders while others (for example breast and prostate cancer) are gender specific.

In terms of incidence, the most frequently occurring cancer types in adults are in the breast and in the lung, with about 11% incidence rate for each. For males the most frequent indication is lung cancer with approximately 14% (figure 1.1(b)) and for females it is breast cancer with about 25% (figure 1.1(c)). It is noteworthy mentioning the large diversity of cancer types, as about 40% of incidence rates concern what is called "other cancers" (figures 1.1(a)–(c)). Concerning mortality rates, lung cancer is identified as the most frequent cause of death related to cancer with about 20% in the whole population (figure 1.1(d)). For females the leading cause of mortality due to cancer is breast cancer (16% – figure 1.1(f)) while for males it is lung cancer (22% – figure 1.1(e)). It is important to note that age plays a significant role in the above mentioned data. When analyzing the data for patients younger than 14 years of age (pediatric patients – data not shown),

both the most frequently occurring cancer type and the highest mortality are related to leukemia (exceeding 30%), regardless of the patient gender.

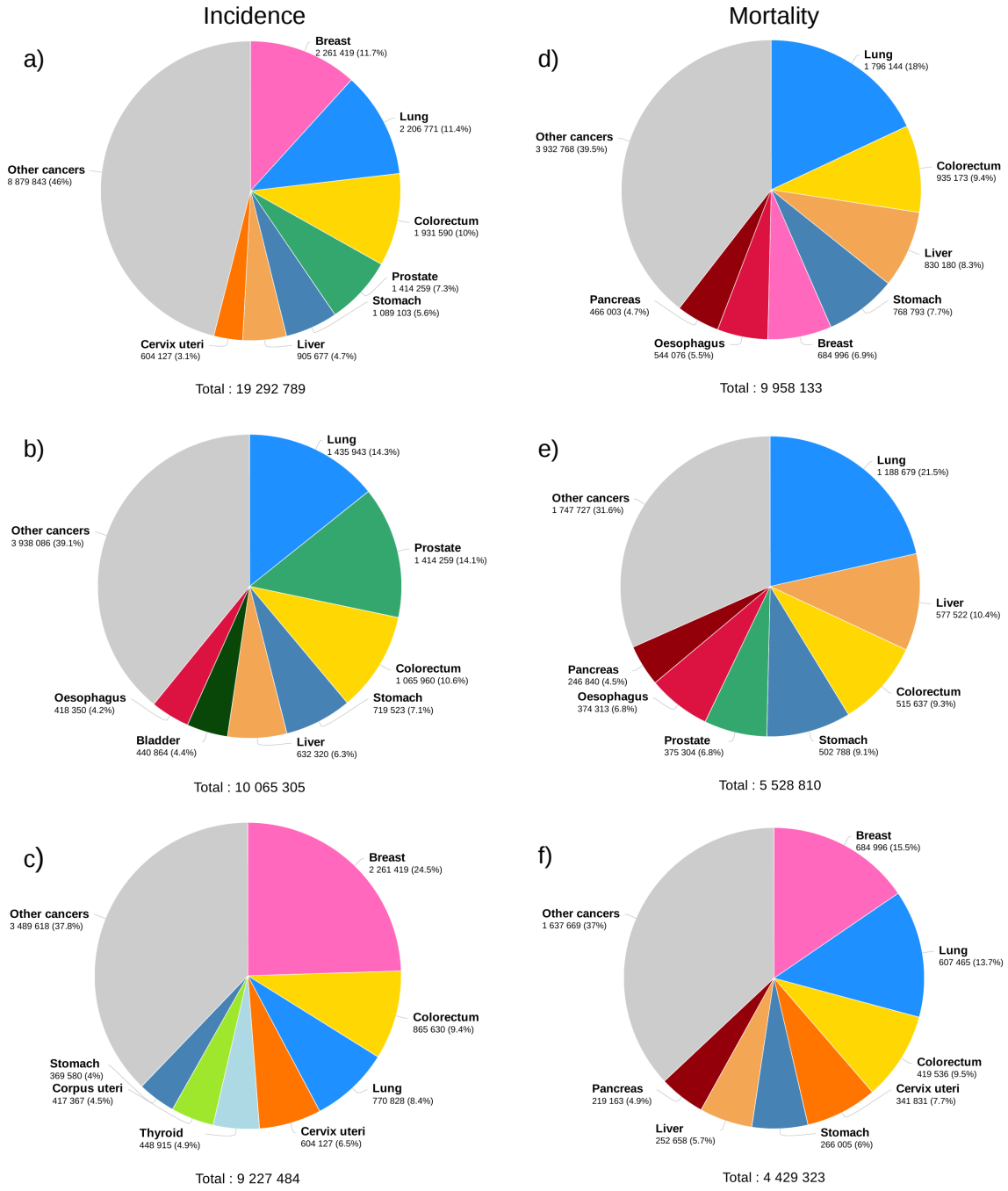

Figure 1.1: Worldwide cancer incidence and mortality estimates for the year 2020, integrated over all patient ages. Graphs were produced using WHO data and the tool Globocan (http::/gco.iarc.fr). The first row shows data for all genders, the second and the third rows for male and female patients, respectively.

According to the German Centre for Cancer Registry Data (Zentrum für Krebsregisterdaten) (ZfKD 2016) a total of 492.090 incidences were registered in 2016. Of these, approximately 47% occurred in women and 53% in men. About half of these incidences concerned the mammary gland (68.900), the prostate (58.800), the large intestine (58.300) and the lungs (57.500). Gender standardized incidence rates are summarized in figure 1.2. During the same year, the number of deaths in Germany related to cancer was estimated to be 229.827. Between 2006 and 2016, the absolute number of new cases of cancer increased by around 2% in men and 5% in women. This increase was less than would have been expected if age–specific incidence rates had remained at the level found in 2006. Since the risk of developing nearly all types of cancer increases with age, population aging would currently be expected to lead to an increase in cancer incidence of around 1% per year.

The German Cancer Research Center (DKFZ) estimates that at least 37% of all new cancer cases in Germany are linked to behavioral or dietary factors, hence avoidable or reducible. The dominant contribution to these risk factors is tobacco consumption, as 19% of annual cancer incidence in Germany is attributable to smoking.

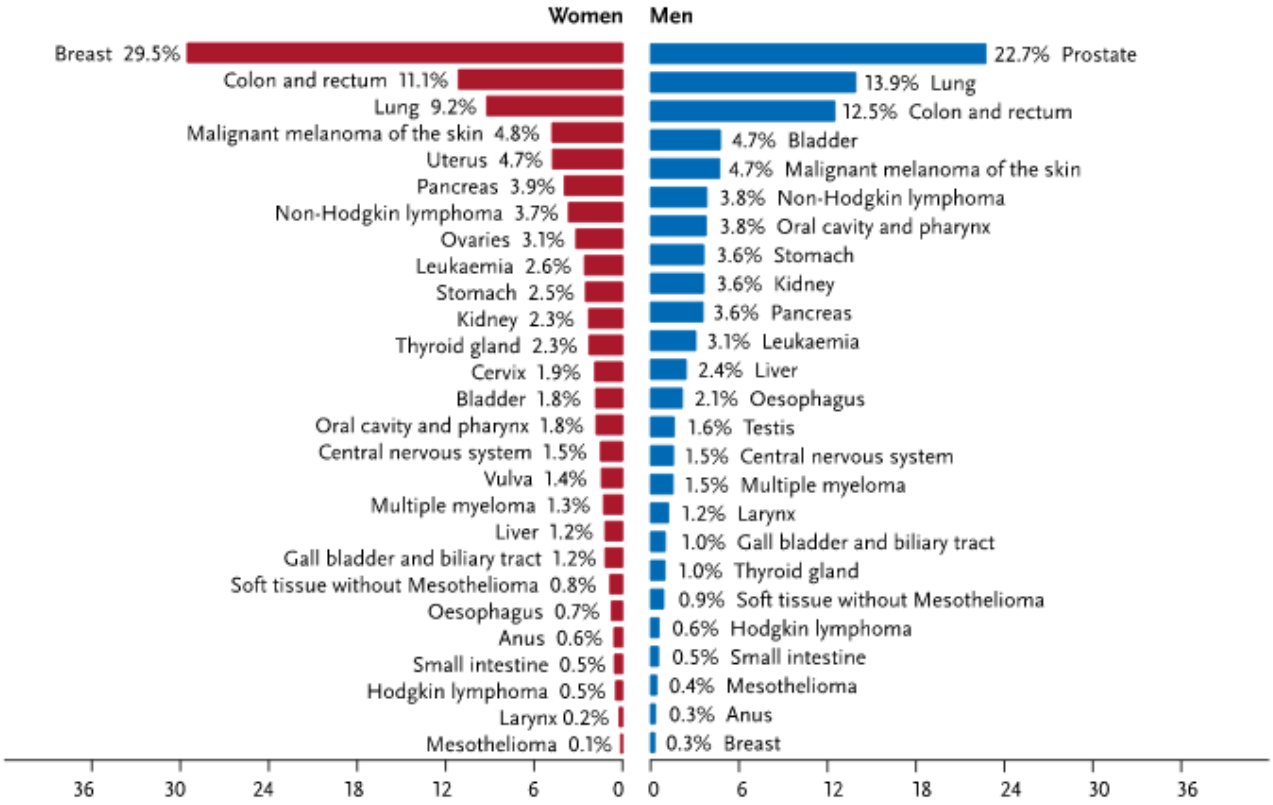

Figure 1.2: Most frequent tumor sites as a percentage of all new cancer cases in Germany 2016, excluding non–melanoma skin cancer (ZfKD 2016).

1.3 Imaging tools and therapeutic strategies

Diagnosing cancer at the earliest possible stage is often a decisive factor for the outcome of the treatment. Due to the diversity in cancer types, in the anatomical sites where they can occur and in the manifestation of symptoms, highly advanced techniques are employed for detection and accurate staging of cancer. There is a wide spectrum of diagnostic tools ranging from computed tomography imaging and biopsy to radiogenomics and artificial intelligence. Many of these tools are applicable beyond diagnosis and staging, to the treatment procedure and treatment response assessment.

Imaging forms an essential part of cancer clinical protocols, providing anatomical, metabolic and functional information. Combining different imaging techniques allows fusing of complementary information, for example functional together with morphological, and can significantly improve staging and therapy planning. The role of imaging throughout the whole clinical workflow, from diagnosis to therapy is schematically outlined in figure 1.3. Most clinically established imaging systems are based on the interaction of electromagnetic radiation with body tissues, with the exception of ultrasound which is based on the propagation of acoustic waves. An excellent review of biomedical imaging techniques applicable to cancer diagnosis and therapy is provided by Fass (2008).

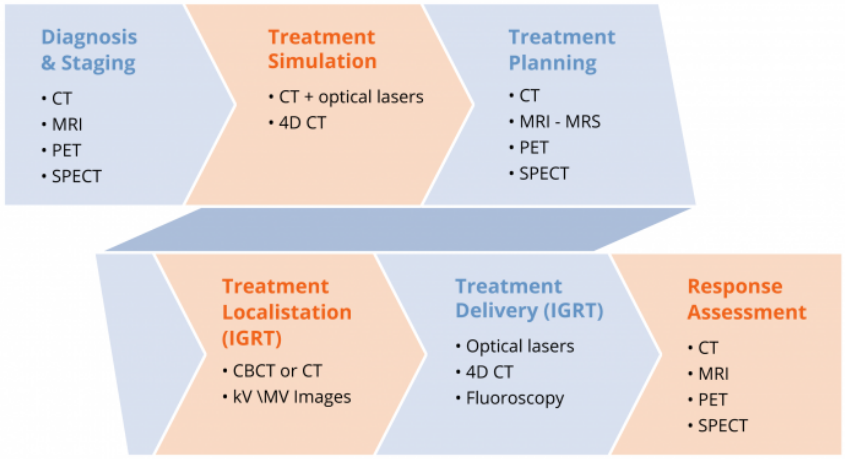

Figure 1.3: The role of imaging (excluding molecular imaging) in the clinical workflow from diagnosis to treatment response (http://www.radiationoncology.com.au/).

Large part of this manuscript refers to novel aspects of computed tomography (CT) imaging applicable to particle therapy. The history of CT comprises development in physics, mathematics and computing intertwined with each other. Seminal role played the discovery of x–rays by Röntgen in 1895 (Röntgen 1898) and the formulation of the Radon theorem (Radon 1917). The basic idea behind tomographic reconstruction is that by obtaining a set of measurements from different angles of radiation passing through an object, yielding line integrals of a property of that object, we can finally reconstruct the spatial distribution of that property. For example, an x–ray radiography consists of mea-

surements of the integrated photon attenuation coefficient along lines connecting the source and particular positions on a detector. From a set of such radiographies it is possible to reconstruct the spatial distribution of the attenuation coefficient in the object. The first experiments on medical applications of this type of reconstruction were performed by Cormack (1963). It is here worth noting that in his pioneering work Cormack proposed that also protons could be used for obtaining tomographic images for radiological applications. The first successful demonstration of an x–ray CT scanner coupled with tomographic reconstruction was achieved by Hounsfield (1973), who is now generally recognized as the inventor of CT. An excellent review of the history and of the technical developments of x–ray CT was published by Kalender (2006).

There are two large classes of tomographic reconstruction algorithms and a recent review of them was published by Willemink & Noël (2018). The first class is the so called direct method or analytical reconstruction. Algorithms belonging to that class are derived from the analytic relationships between functions and their line integrals. An example of such an algorithm is the well established filtered backprojection. A brief description of a filtered backprojection algorithm applied to proton computed tomography is provided in chapter 6. The second class, dominant in nowadays commercial CT scanners, is the iterative reconstruction methods. The basic concept starts by discretizing the volume of the image object. Then the measured values are considered as the result of a matrix multiplication of the unknown discretized object values, with a matrix modelling the measurement procedure and the physical interactions of radiation with matter. The result is a huge set of simultaneous linear equations that can be solved to compute the CT image. In principle the system of equations can be solved with standard methods. In practice, due to the immense number of equations and of the properties of the problem (such as sparsity), we resort to iterative methods. For comprehensive review of the various reconstruction algorithms employed for tomographic reconstruction and their mathematical formulations, the reader is referred to chapter 26 (Leahy et al. 2009) from the book of Bovik (2009).

Following the diagnostic phase which is heavily relying on imaging, a therapeutic strategy is planned. The main treatment options are: surgery, radiation therapy, and systemic treatment which includes chemotherapy, targeted therapy, hormonal therapy, and immunotherapy. Surgery is an invasive procedure with the aim of removing the entirety or part of a tumor and the branch of medicine associated to it is called surgical oncology (a reference for specialized medical practitioners is (Poston 2007)). Radiation therapy, which is the contextual basis of this manuscript, means the use of ionizing radiation for killing cancer cells and the corresponding branch of medicine is called radiation oncology (see section 1.4 for more details). Chemotherapy consists of administering to the patient pharmacological agents mainly aiming at inhibiting mitosis (cell division) or inducing DNA damage (DeVita & Chu 2008). Targeted therapy is also a drug based type of treatment aiming at interfering with specific proteins involved in tumorigenesis, by focusing on specific molecular changes which are unique to a particular cancer (Baudino 2015). Hormonal therapy is a non–toxic therapy used to mainly reduce the size of the primary cancer by removing specific hormones from the body or blocking their effects (Abraham & Staffurth 2016). Finally, immunotherapy, the most recent treatment technique, is based on

the concept of stimulating the immune system or blocking immune checkpoints, thus allowing the immune system to better recognize and attack cancer cells (Waldman et al. 2020). The branch of medicine encompassing all different systemic treatment options is named medical oncology.

Due to the very large diversity in cancer types it is difficult to summarize in a definitive and meaningful way the frequency with which different treatment options are used. Furthermore, treatment options vary depending on the exact stage of a particular cancer type and very often more than one treatment options are used for optimal treatment outcomes. In general the most common treatments are surgery, chemotherapy, and radiation. A handful of examples highlighting this variability, taken from the American Cancer Society report from year 2019 on cancer treatment and survivorship statistics (Miller et al. 2019), are listed here. The most common treatment among women with early–stage (stage I or II) breast cancer is surgery combined with radiation therapy (49%). By comparison, more than two–thirds (68%) of patients with stage III breast cancer disease undergo surgery, most of whom also receive chemotherapy. The majority of patients with stage I and II colon cancer undergo surgery without chemotherapy (84%), whereas approximately two-thirds of patients with stage III disease (as well as some patients with stage II disease) also receive chemotherapy. Chemotherapy is the standard treatment for most leukemias. For the most common type of lung cancer (non–small cell lunc cancer), only 18% of patients with stage III undergo surgery, whereas most (62%) are treated with chemotherapy and/or radiation.

1.4 Particle therapy and its imaging

Before delving into the field of particle therapy, a brief review of the broad field of radiation therapy might be informative to the reader. As described in previous sections (1.1 and 1.3), radiation therapy consists in utilizing ionizing radiation in order to kill cancer cells. From the discovery of x–rays (Röntgen 1898) and the first application of radiation for cancer treatment (Grubbé 1933), both at the end of the 19$^{\text{th}}$ century, radiation therapy experienced more than a century of rapid technological advancements. Comprehensive historical reviews can be found in the work of Bernier et al. (2004) and Connell & Hellman (2009). These technological advances, alongside with developments in medical imaging and the corresponding clinical research, resulted in elaborate techniques which give us today the capability to personalize treatments for accurate radiation dose delivery based on clinical, biological and anatomical information.

Radiation therapy can be split in two categories, depending on the position of the ionizing radiation source relative to the patient. Brachytherapy is the category of radiation therapy in which the source is directly placed into or adjacent to the tumor, hence inside the patient. It started with surface application of radium in moulds or other applicators (Paterson & Parker 1934) and continued in this form until the 70s. A drastic change in this form of radiation therapy came with the development of the technique to plant ra-

dioactive capsules or seeds in or near the tumor. External beam radiation therapy on the other hand signifies that the source (usually from an accelerator but can also originate from a radioactive substance) is positioned outside the patient and that radiation is directed to the specific anatomical site to be treated. The main workhorse of modern external beam radiation therapy is the electron linear accelerator (LINAC). Electrons are accelerated and hit a target, producing x–rays of a broad energy spectrum. The first such accelerator designed for radiotherapy was developed by 1948 by Fry et al. (1948) and produced 4 MV x–rays. The broad energy spectrum of x–rays from LINACs, is characterized by the accelerating potential in which electrons travel before hitting the target that produces x–rays, mainly via Bremsstrahlung. Hence the unit used is that of an electric potential, usually in MV (mega–Volt).

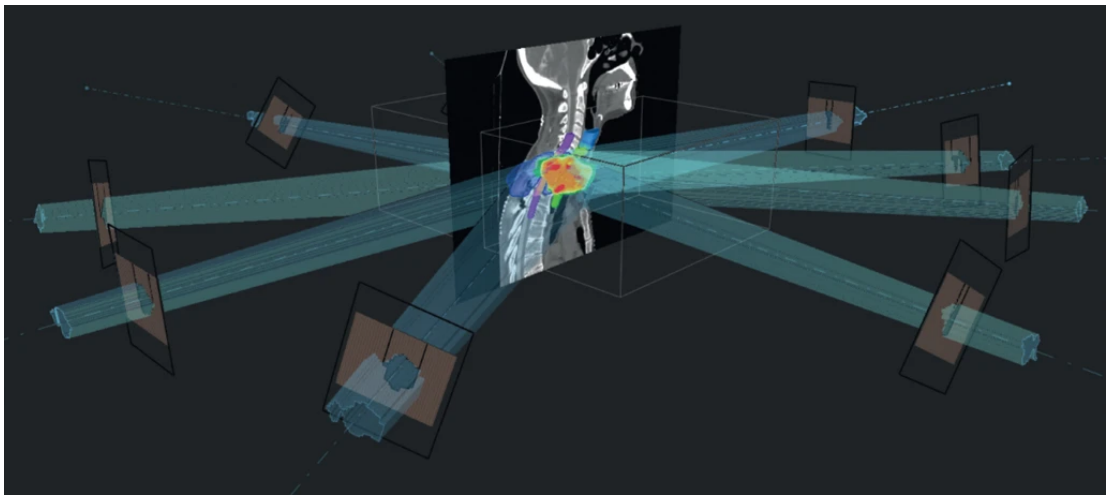

Figure 1.4: Schematic representation of dose delivery in IMRT (Baumann et al. 2016). A treatment session is split into a nmber of angles/fields. In within each field the use of multileaf collimators allows the shaping of the beam and the modulation of the intra–field fluence.

The state–of–the–art external beam radiation therapy with x-rays is the intensity modulated radiation therapy, abbreviated as IMRT (Brahme 1987) – an informative topical review is provided by Bortfeld (2006). An illustration of the concept is shown in figure 1.4 (Baumann et al. 2016). The dose is delivered to the tumor from a number of different angles using a LINAC with the capability to rotate around the patient and equipped with dynamically controlled multileaf collimators (MLCs). At each angle the radiation field can be shaped conformally to the tumor and the intra–field fluence is modulated in order to achieve the desired dose distribution (Bortfeld et al. 1994). Superimposing a large number of fields allows to reach the prescribed dose in the tumor, while sparing radiosensitive healthy organs. The optimization of the dose delivery/fluence patterns for each field is obtained via inverse treatment planning (Xiao et al. 2013). The procedure is based on a high quality x–ray CT image acquired prior to the treatment, containing a three dimensional photon attenuation coefficient map of the patient. The tumor and several adjacent organs

are delineated on that patient x–ray CT and dose prescriptions and limits are defined for the tumor and for the healthy organs. For a given set of treatment angles, a computer algorithm determines the optimal field shape and the intra–field fluence by minimizing a cost function. The resulting treatment plan is one of the possible solutions that satisfy the dose prescription and limits, for that given set of angles. A step further in the direction of delivering the radiation from many different angles are techniques which employ the concept of rotation with continuous beam delivery. A main benefit of these methods called arc IMRT is the reduction of the time needed to complete a treatment session. Examples of arc IMRT include tomotherapy (Mackie et al. 1993), intensity–modulated arc therapy (IMAT) (Yu 1995) and volumetric modulated arc therapy (VMAT) (Otto 2007).

The denomination particle therapy is nowadays utilized to describe external beam radiation therapy with heavy charged particles, mainly discriminating from photon therapy. The ambiguous term heavy refers to particles heavier than electrons. This definition therefore excludes electrons and neutrons, occasionaly used in radiation therapy. The former are used for treating superficial tumors (Hogstrom & Almond 2006) due to their favorable dose distribution as a function of depth in the patient, while the latter are well suited for treating slowly growing tumors that are resistant to x–rays (Jones & Wambersie 2007) because of their increased biological effects. Although a detailed description of biological effects of radiation is out of the scope of this work, specific aspects of the biological effects of protons used for the therapy of cancer are investigated later in this manuscript.

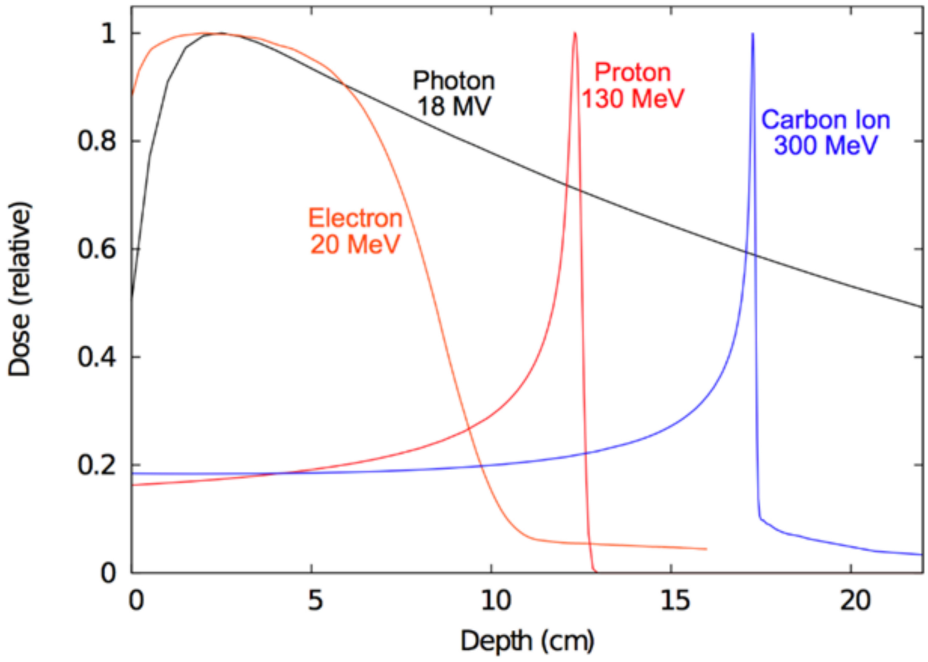

Figure 1.5: Dose distributions normalized to their respective maximum, as a function of depth in water shown for various clinical radiation beams: 20 MeV electrons, 18 MV photons, 130 MeV protons and 300 MeV/u carbon–ions (Kaiser et al. 2019)

Figure 1.5 depicts the dose distributions of different particle beams as a function of depth in water. The photon dose distribution displays an initial secondary electron build–up region, followed by a maximum in the dose deposit and then a slowly decreasing tail governed by the exponential attenuation of the photon beam (Beer–Lambert law). For comparison, the electron dose distribution with its maximum at approximately the same location has a much steeper dose fall–off. For both particle species, increasing the energy yields a broader maximum of the dose distribution, but also creates a slower decreasing fall–off. Thus for deep seated tumors (at $10 - 20\,\mathrm{cm}$ water equivalent depth), a significant volume of healthy tissue receives high dose, which might actually be even higher than what is deposited in the tumor when considering a single treatment field. Heavy charged particles on the other hand, due to fundamentally different interactions in matter compared to photons (see chapter 2) and reduced scattering compared to electrons, yield a much more favorable dose distribution, which is called Bragg curve, named after William Henry Bragg who discovered it in 1903 (Bragg & Kleeman 1905). The Bragg curve (see figure 1.5 proton and carbon ions dose curves) is characterized by a distinct peak at a depth determined by the initial beam energy, a shallow plateau upstream and a very steep fall–off downstream. For ions heavier than protons a fragmentation tail is also present after the Bragg peak. Therefore the main motivation for use of heavy charged particles in external beam radiation therapy is the possibility to place the dose maximum region in to the tumor and spare healthy tissues from high dose irradiation. To overcome the limitations of the photon dose distribution and achieve similar results to heavy ions, a larger number of fields is required (see also figure 1.4). Superimposing photon fields from different angles results in a dose maximum in the tumor, at the expense though of spreading the dose to a larger volume around it, when compared to protons as shown in figure 1.6.

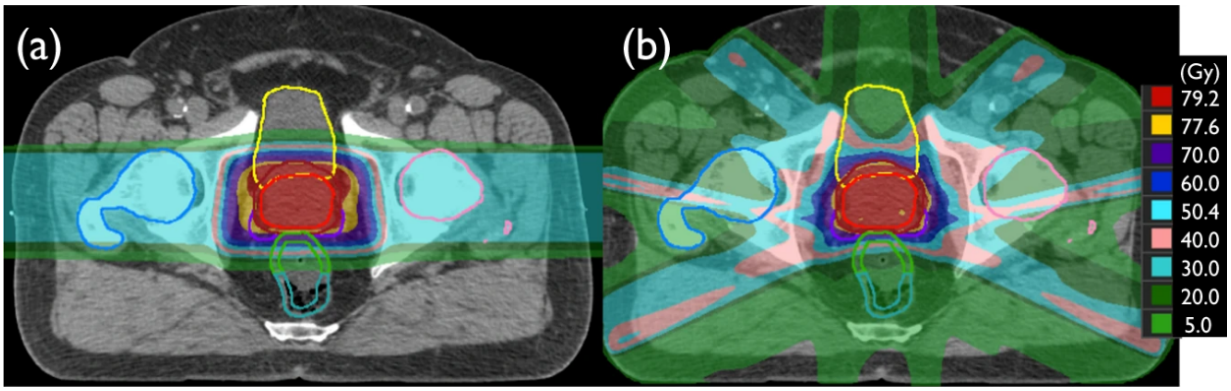

Figure 1.6: Example of the dose distribution for (a) proton beam treatment and (b) intensity–modulated radiation treatment (photons) of the prostate (delineated with a red line). The photon creates a “low–dose bath”, whereas the proton plan effectively spares more of the normal tissues from receiving the lower–dose radiation (Kamran et al. 2019)

Robert Wilson in his seminal paper in 1946 (Wilson 1946) predicted the potential applications of fast protons in radiation therapy. The famous 184–inch (467.36 cm) Berkeley cyclotron, capable of producing high–energy proton, deuterons and helium ions was utilized

by Tobias et al. (1955) to investigate and eventually confirm the predictions of Wilson. In the 1950s, heavier ion therapy with carbon, nitrogen, argon, neon, and silicon ions was assessed. The Berkeley cyclotron continued playing a pivotal role in particle therapy until its shutdown in 1987, and was succeeded by Bevatron/Bevalac which was used for helium ions particle therapy, until its shutdown in 1993. The biomedical research and patient treatments continued in a few research centers (Larsson et al. 1958, Kjellberg et al. 1962, Suit et al. 1975). The first clinical facility dedicated proton therapy was built at the Loma Linda University in California in 1990 (Slater et al. 1992, Slater 2007).

There are currently 91 exclusively proton and 12 proton and heavy ion therapy facilities worldwide, with a large number of new facilities in the planning or building phase. In Germany there are five particle therapy facilities, compared to a total of about 400 photon based radiation therapy facilities. Three of them (Berlin, Dresden, Essen) offer proton therapy and two of them (Heidelberg and Marburg) are capable of proton and heavier ion therapy. These five facilities, together with the decommissioned proton center in Munich, have treated so far about 20.000 patients. On a global basis, approximately three hundred thousand cancer patients have been treated so far with protons or heavier ions (see figure 1.7).

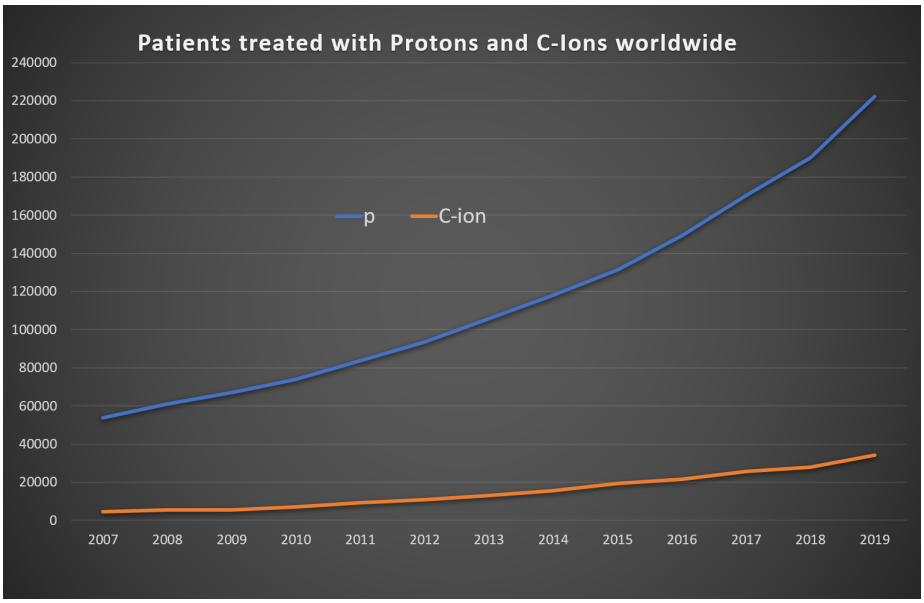

Figure 1.7: Cumulative number of patients treated with protons and carbon ions worldwide, up to the year 2019. Data from the Particle Therapy Co–operative Group (PTCOG) (www.ptcog.ch).

In terms of the technique of delivering the prescribed dose to the patient, particle therapy is further split to passive beam delivery and active beam scanning. Passive beam delivery consists in spreading an initially narrow beam using scattering foils to a size that can cover laterally an area at least as large as the tumor. Patient–specific milled collimators are employed to conform the beam to the lateral shape of the tumor. Longitudinally, the

beam is spread by intercepting it with a rotating modulator wheel, whose varying thickness results in a polyenergetic beams of particles stopping in different depths in the patient, thus covering the tumor. Unlike passive proton beam delivery, in active beam scanning a narrow quasi–monoenergetic beam (approximately 1 cm full width at half maximum) is magnetically deflected to cover laterally the tumor. Such small sized beams are call pencil beams. Longitudinally the tumor is covered by choosing beams of different energies, thus moving the position of the Bragg peak in depth.

Imaging is an integral part of modern radiation therapy, hence of particle therapy as well. In this manuscript the focus is on imaging for treatment planning, treatment localisation and treatment monitoring. For treatment planning, the information required is not only the morphology of the treated anatomical site, but also a volumetric map of the stopping power of charged particles in that region (see chapter 2). Instead of merely stopping power, which is strongly dependent on the energy of the particle, the nearly energy independent relative stopping power to that of water (RSP) is used. The RSP map allows to choose the energy of the beam in order to fully cover the tumor region, without placing the Bragg peak in healthy tissues. The RSP map is currently obtained indirectly from an x–ray CT image of the patient. The latter provides an attenuation coefficient map, which is then converted to RSP via a semi–empirical calibration, for example the well established stoichiometric calibration proposed by Schneider et al. (1996). This procedure results in RSP errors of up to 3.5% in biological material (Paganetti 2012, Yang et al. 2012). These errors can be significantly reduced by using dual energy x–ray CT (DECT). The patient is imaged with two well separated energies which allow the estimation of the RSP of the tissues (for detailed information read 3.6.1). The potential of DECT for reducing the RSP error to about 1% was demonstrated in several studies (Yang et al. 2010, Hünemohr et al. 2013, Bourque et al. 2014, Hudobivnik et al. 2016, Möhler et al. 2016, Han et al. 2016, Taasti et al. 2016, Lalonde et al. 2017, Saito & Sagara 2017a, Almeida et al. 2018).

Proton CT (pCT) is considered as a candidate for improving RSP accuracy, proposed in 1963 by Cormack (1963). A pCT scanner concept that could reconstruct RSP by detecting the positions, directions and residual energies of individual protons was described in (Huesman et al. 1975). The main idea is that protons of sufficiently high energy for penetrating the patient could be used to acquire proton radiographies from different angles. Knowing the initial energy and measuring the residual energy for each proton one can determine the water equivalent path length (WEPL) traversed in the patient. The WEPL is defined as the length in water required to yield the same energy loss measured for protons traversing the patient, and it is the line integral of the RSP, in analogy to the line integral of the attenuation coefficient in x–ray radiography. Tomographic reconstruction techniques (Penfold et al. 2009, Rit et al. 2013, Poludniowski et al. 2014, Hansen et al. 2016) can be then applied in order to obtain three dimensional RSP maps of the patient. Contrary to photons which travel mostly along straight paths, protons scatter frequently (described in chapter 2) resulting in inherently lower spatial resolution images when compared to x–ray CTs. To improve the spatial resolution of pCT images the proton paths can be approximated by curved lines and this has to be accounted for in the image reconstruction. This is done by measuring the position and direction of each proton before and after

the imaged object, allowing for the construction of a curved path estimate. A few pCT scanners being currently at the level of prototype are used for preclinical studies (Dedes et al. 2020) and have achieved RSP accuracies ranging from 1.6% to 0.6% (Esposito et al. 2018, Giacometti et al. 2017a, Dedes et al. 2019).

Following treatment planning, the optimized therapeutic dose is delivered to the patient in a number of treatment sessions (usually around 30 to 40) each on a different day, called fractions. Prior to the delivery of each fraction of the treatment dose the patient has to be positioned exactly as on the treatment planning session. Especially in the case of particle therapy, due to the very steep dose gradients, any significant anatomical changes have to be identified and potentially accounted for by adapting the initial treatment plan to the new anatomy. Large remaining uncertainties in positioning or anatomical changes may compromise the dose conformity to the tumor and the planned sparing of healthy organs (Landry & Hua 2018). For these reasons, image guidance, the use of imaging prior to each fraction, is of increasing importance in modern radiation therapy (Verellen et al. 2007). Although 2D imaging (radiography) may be adequate for patient positioning (Bel et al. 1993, Pisani et al. 2000), volumetric imaging of the patient performed in the treatment room (in–room imaging) and at the treatment position (on–isocenter imaging) would bring the benefit of accurate plan adaptation (Jagt et al. 2020). Cone beam x–ray CT (CBCT) is nowadays utilized by most photon therapy facilities as a good compromise of daily volumetric in–room and on–isocenter imaging at low dose. Proton therapy facilities are currently catching up in terms of image guidance. CBCT has certain limitations when applied in particle therapy, which have been identified and partially corrected for (Landry et al. 2015a, Landry et al. 2015b, Thing et al. 2016). Most of these limitations can be circumvented by the use of pCT, which is capable of producing more accurate RSP maps of the patient at lower imaging doses. The pCT imaging dose in healthy tissues can be pushed to even lower values by applying techniques that allow for spatially varying image quality and imaging dose, and are one of the main topics covered in detail in this manuscript.

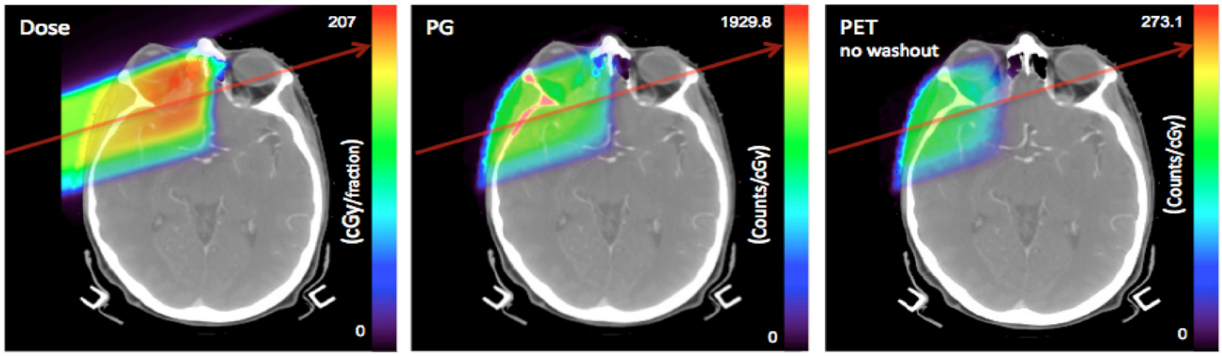

Figure 1.8: Example of the correlation of dose (left) with prompt–gamma (middle) and with PET activity (right) in patient anatomy. The image was adapted from the simulation study of Moteabbed et al. (2011).

As previously described, the higher dose conformity of particle therapy brings the drawback of increased sensitivity to anatomical variations, patient positioning and RSP errors. In addition to or combined with image guidance, one could exploit secondary radiation generated during treatment by the therapeutic beam. Certain types of that secondary irradiation are strongly correlated with the therapeutic ion range in the patient (see figure 1.8) and can be used to perform in–vivo ion range monitoring, without incurring additional imaging dose to the patient. Different ion range monitoring techniques have been proposed (Parodi & Polf 2018), with the majority of them being under development or on the preclinical investigation level. The two most mature techniques that have already been applied to patient irradiation are the positron emission tomography (PET) and the prompt–gamma (PG) based range verification. An early comparison between PET and PG for range monitoring on the simulation level was presented by Moteabbed et al. (2011). PET is based on detecting the 511 keV annihilation photons originating from the β^+ emitting isotopes created in the patient by the therapeutic beam (Parodi et al. 2001, Enghardt et al. 2004). Its clinical application was demonstrated by Parodi et al. (2007). An alternative method is based on the detection of prompt photons emitted by nuclei undergoing de–excitation following nuclear interactions (Stichelbaut F 2003, Min et al. 2006), therefore called prompt–gammas. These MeV energy scale photons are emitted and reach the detecting devices within nanoseconds or less from the time of interaction and their spatial distribution is correlated to the ion beam path and the location of the Bragg peak in the patient. A first clinical application was described by Richter et al. (2016). In this manuscript, the nuclear modelling aspect relevant to PG emission and the applicability of PG based range monitoring to prostate cancer patients are investigated. Recently a series of studies took a step further (Tian et al. 2018, Tian et al. 2020, Tian et al. 2021), by taking into account the aspect of range monitoring in the process of creating a treatment plan for the patient.

1.5 Goals of the habilitation project

The scientific goal of this habilitation project was to pursue solutions to some of the open questions pertaining to radiation therapy of cancer using heavy ions, in particular proton therapy. Research in this field is by its nature multidisciplinary, combining physics with biomedical engineering, experimental work with advances on the theoretical background, particle physics with imaging techniques. The following chapters are arranged in a conceptual rather than chronological order of investigation, highlighting that problems representative of the whole particle therapy workflow, ranging from treatment planning to treatment verification, were addressed.

As a prelude, the next chapter (chapter 2) deals with the basic physics describing heavy charged particle interactions with matter, in the energy regime relevant to radiation therapy. Chapter 3 presents a direct experimental comparison between pCT and DECT, the first of this kind, addressing the since long standing question of the achievable RSP accuracy for treatment planning by these two promising imaging modalities. Chapter 4

explores a potential solution to the problem of image quality vs. imaging dose, on the idealized simulation level, by employing fluence–modulated pCT (FMpCT), thus opening the door for using pCT not only for treatment planning but also for daily low dose image guidance and plan adaptation. The proof–of–concept FMpCT was put to the test experimentally in a clinical proton therapy facility, as described in chapter 5. For FMpCT to be applicable as an image guidance technique, the imaging quality has to be well understood. This is covered in chapter 6, where the basic theory describing noise formation in proton imaging and means to control it are outlined. In chapter 7 that theoretical background was validated and combined with a fully realistic pCT scanner simulation platform which was benchmarked against experimental data. FMpCT reaches its full potential when spatially varying image quality can be controlled to reach specific prescriptions. This was developed and demonstrated in chapter 8. With the complete workflow in place, chapter 9 describes the experimental realization of the full concept of FMpCT. Assuming an accurate RSP patient map of the anatomy of the treatment day, acquired in a dose efficient way, optimal treatment plans can be calculated. Treatment planning for proton therapy is based on specific radiobiological assumptions. The effect of the radiobiological uncertainties on the dose delivered to the patient is studied in chapter 10 for simulated treatments of a nasopharyngeal and a prostate cancer patient. Verification of the delivered treatment in the form of prompt–gamma range monitoring is heavily relying on Monte Carlo simulations. Chapter 11 assesses nuclear modeling of these processes and proposes improvements. The sensitivity of PG range monitoring simulated on prostate cancer patients is explored in chapter 12. Finally, in the context of novel particle accelerators for particle therapy, a radiation protection study for CALA, the laser–ion accelerator facility of Munich, is presented in chapter 13.

Chapter 2

Basic physics of particle therapy

This chapter reviews the basic physics behind particle therapy. Excellent dedicated reviews on the topic have been published by Chu et al. (1993), Newhauser & Zhang (2015), Durante & Paganetti (2016), Gottschalk (2018) and others. A thorough examination of the theoretical background of charged particle interactions with matter is beyond the scope of this chapter. The aim is therefore to present in a brief and concise fashion the main interaction mechanisms relevant to the energy regime utilized in particle therapy and expose their impact on therapeutic and imaging particle beams, with a focus on protons.

2.1 Review of proton interaction mechanisms

The purpose of this section is to provide a summary of the main interaction mechanisms of protons with matter, for the energy range below a few hundreds of MeV's. The three main types mechanisms are schematically described in figure 2.1.

When a proton travels trough matter, it experiences a large number of inelastic Coulomb interactions, which transfer small amounts of energy from the proton to the electrons of the medium (see figure 2.1(a)) via ionization and excitation of the atoms. For example, the number of primary ionizations induced by a 10 MeV proton in liquid water is of order $10^2/\mu$m, each of them reducing the proton's energy by a few tens of keV's (Perris & Zarris 1989). Due to the fact that the proton mass to electron mass ratio is approximately 1836, the deflection of the proton from its initial trajectory is considered negligible. Therefore the net effect of these interactions is energy loss. Bremsstrahlung, which is a radiative process for energy loss, although still possible for protons, is considered negligible due to its large mass. Deflections of the proton direction mainly occur when it passes near an atomic nucleus (see figure 2.1(b)), experiencing a repulsive Coulomb force in conjunction with the usually larger mass of the nucleus. A myriad of these scattering events sum up to non-negligible change of direction of the proton when it traverses adequate material thickness, but still elastic scattering mainly yields small angle deflections. As an example for quantifying this, the scattering angle distribution of 158.6 MeV protons after traversing 1 cm

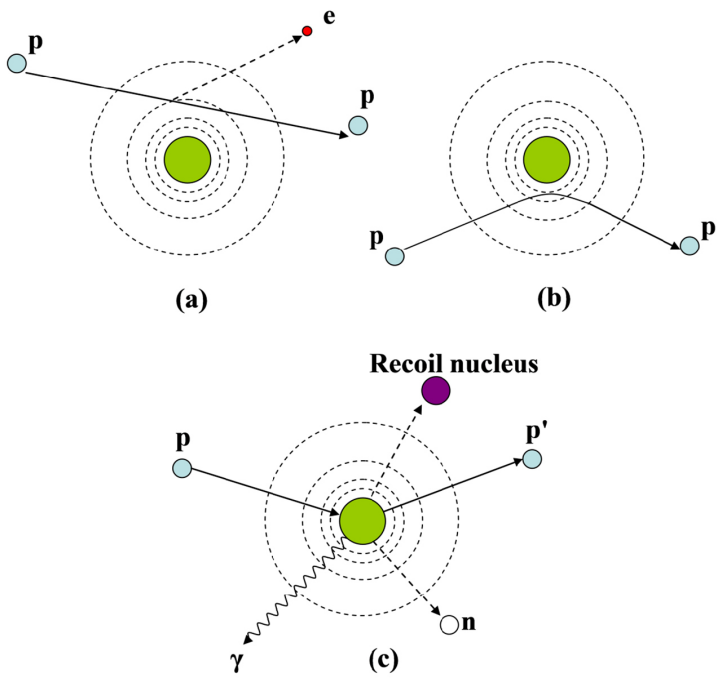

Figure 2.1: Illustration of proton interaction mechanisms (Newhauser & Zhang 2015): (a) energy loss via inelastic Coulomb interactions with atomic electrons, (b) deflection of proton trajectory due to repulsive Coulomb elastic scattering with a nucleus, (c) non-elastic nuclear interaction and creation of secondary particles.

of water is nearly Gaussian, with more than 99% of the protons experiencing a direction change of less than 1° (Gottschalk 2018). Finally there are non–elastic nuclear interactions between protons and the atomic nuclei of the absorber medium (see figure 2.1(c)). The projectile proton is absorbed and the excited nucleus de–excites via emission of secondaries such as photons, protons, neutrons, deuterons, tritons, alpha particles and heavier nuclei. Nuclear interactions might also result to radioactive nuclei with long decay times. In the radiation therapy energy range regime (80 MeV - 250 MeV), proton beams attenuate approximately by 1% per cm in water (Paganetti 2016).

2.2 Energy loss

2.2.1 Stopping power

Although proton energy loss mainly occurs via frequent ionizations and excitations, it can be regarded as a continuous loss along the proton's trajectory. This assumption is called the continuous slow down approximation (CSDA) and a way to characterize then the energy loss is the linear stopping power, which is the energy loss rate per unit path length:

$$S = -\frac{\mathrm{d}E}{\mathrm{d}x} \tag{2.1}$$

where $\mathrm{d}E$ is the energy loss along the proton path length $\mathrm{d}x$. It is important to note that the energy loss is a stochastic effect, meaning that not every proton with the same initial energy, impinging the same material will undergo the exact same interactions yielding the exact same energy loss. Therefore, equation 2.1 and all subsequent equations of stopping power are only meaningful for a particle bean and they quantify the mean energy loss.

For a material of given chemical composition, the probability of interactions depends on the number of atoms, hence on the mass density of the material. To factor out this effect, the stopping power of a particle in a material can be defined independently of the material mass density via the mass stopping power:

$$\frac{S}{\rho} = -\frac{\mathrm{d}E}{\rho \mathrm{d}x} \tag{2.2}$$

where ρ is the mass density of the material.

A simple mathematical expression that parametrizes the mass stopping power of a particle species as a function of its energy was formulated by Bragg & Kleeman (1905):

$$\frac{S}{\rho} = -\frac{E^{1-p}}{\rho \alpha p} \tag{2.3}$$

where α is a material dependent constant, E is the energy of the beam and p is a dimensionless constant depending on the particle type and the beam energy.

In the early 1930's, Bethe (1930) developed a more precise formula for stopping power of heavy charged particles, taking into account quantum mechanical effects and using the assumption of unbound stationary electrons in the atom:

$$\frac{S}{\rho} = -\frac{\mathrm{d}E}{\rho \mathrm{d}x} = 4\pi N_\mathrm{A} r_\mathrm{e}^2 m_\mathrm{e} c^2 \frac{Z}{A} \frac{z^2}{\beta^2} \left[\ln \frac{2 m_\mathrm{e} c^2 \gamma^2 \beta^2}{I(1-\beta^2)} - \beta^2 - \frac{\delta}{2} - \frac{C}{Z} \right] \tag{2.4}$$

where N_A is Avogadro's number, r_e and m_e are the classical electron radius and the electron mass, c is the speed of light, Z and A are the atomic and mass numbers of the material, I is the mean excitation potential of the material in eV which signifies the averaged excitation potential per electron in the target atom, z is the charge of the projectile, β is the ratio of the projectile's speed to the speed of light and γ is the Lorentz factor. The term $\frac{C}{Z}$ is the so called shell correction which is needed for protons of energy below approximately 100 MeV as at these energies Bethe's initial requirement that the projectile's velocity by far exceeds the atomic electron's velocity is not valid. Omitting the shell correction results in overestimation of the stopping power at low projectile energies. The term $\frac{\delta}{2}$ is the density

correction, which adds to the Bethe formula the effect of shielding of remote electrons by the electrons nearer to the projectile. This correction becomes important for highly relativistic projectiles and high density absorber media and omitting it results in overestimation of the stopping power at high projectile energies.

According to equation 2.4, we can deduce a few main dependencies of the stopping power. The first is that the linear stopping power is proportional to the electron density of the material ($\rho N_{\mathrm{A}} \frac{Z}{A}$). It is worth noting here that for biological materials the ratio $\frac{Z}{A}$ does not deviate much from 0.5, which means that the dependency on the electron density is dominated by the mass density. The stopping power also depends on the I value of the material. The latter monotonically increases with Z and varies from approximately 19 eV for oxygen to approximately 820 eV for lead (Newhauser & Zhang 2015). Due to the fact that the I value appears in the formula in the logarithmic term, its effect is small in the energy range regime relevant for particle therapy. Concerning the projectile, the stopping power is proportional to its charge squared (z^2). It is also highly dependent on the velocity (hence kinetic energy) of the projectile. For low energies ($\beta\gamma < 3$) the stopping power increases as a function of $\frac{1}{\beta^2}$. For $\beta\gamma \approx 3 - 3.5$ the stopping power as a function of the projectile energy or velocity reaches a minimum. This $\beta\gamma$ range is independent of the particle species. Particles with such $\beta\gamma$ are called minimum ionizing particles and their mass stopping power is approximately $2\,\mathrm{MeV} \cdot \mathrm{g/cm^2}$. Above that $\beta\gamma$ area the logarithmic term becomes increasingly important and the stopping power increases with increasing β^2. For protons, $\beta\gamma \approx 3 - 3.5$ corresponds to kinetic energies of $2 - 2.4\,\mathrm{GeV}$, which are far above the maximum therapeutic proton energies of about 250 MeV. In the proton energy range of $50 - 250\,\mathrm{MeV}$ the stopping power in water drops from 12.45 MeV/cm to 3.91 MeV/cm (Seltzer 1993). As already mentioned, the energy transfer from protons to atomic electrons happens in rather small amounts, due to the small mass of the electron with respect to that of the proton. The maximum possible energy transfer from an ion of mass M to an unbound stationary electron of mass m_{e} is:

$$\Delta_{max} = 2m_{\mathrm{e}}c^2\beta^2\gamma^2 \left[1 + 2\frac{m_{\mathrm{e}}}{M}\gamma + \left(\frac{m_{\mathrm{e}}}{M}\right)^2\right] \tag{2.5}$$

where c is the speed of light, β is the ratio of the ion velocity to the speed of light and γ is the Lorentz factor. Even very energetic protons, for example with a kinetic energy of 200 MeV, will transfer a maximum energy of 500 keV to an atomic electron (Newhauser & Zhang 2015).

2.2.2 Absorbed dose

The absorbed dose is a quantity applicable to ionizing radiation. It is defined as the energy e imparted to matter of a finite volume and of mass m:

$$D = \frac{e}{m} \tag{2.6}$$

The unit of absorbed dose is the gray (Gy) equal to one Joule per kilogram (J/kg). In the case of indirectly ionizing radiation, energy is imparted to matter in a two step process (Podgorsak 2016). In a first step, the indirectly ionizing radiation transfers energy to secondary charged particles in the form of kinetic energy. In a second step, these charged particles deposit some of their kinetic energy to the medium (resulting in absorbed dose) and lose some of their energy in the form of radiative losses (bremsstrahlung, annihilation in flight). Assuming charged particle equilibrium, the absorbed dose can be also expressed as a function of the fluence of a beam of particles $\Phi(\mathbf{r})$ at a given point $\mathbf{r}$:

$$D(\mathbf{r}) = \Phi(\mathbf{r}) \cdot \frac{S}{\rho} \tag{2.7}$$

where S/ρ is the mass stopping power of that particle type in the traversed medium. In the general case of a charged particle beam described at a given point $\mathbf{r}$ by a fluence spectrum $\Phi(\mathbf{r}, E)$, the dose becomes:

$$D(\mathbf{r}) = \int_0^\infty \Phi(\mathbf{r}, E) \cdot \frac{S(E)}{\rho} \, \mathrm{d}E \tag{2.8}$$

where $S(E)$ denotes the energy dependent stopping power.

2.2.3 Range

The range of a particle beam is defined as the depth in the absorber where half of the particles have come to rest. Since the energy loss is a stochastic effect, a definition of range for a single particle would not be meaningful. As shown in figure 2.2, the particle beam attenuates while traversing the medium due to nuclear interactions. Therefore, the range is defined rather as the depth at which half of the particles that reached near the end–of–range come to rest, instead of half of the initial number of particles. So far the range has been defined as a depth in the medium. Given that charged particles and specifically protons do not travel along straight lines because of scattering, a given depth in the absorber is not necessarily the same in physical length as the actual path length particles travelled to reach that depth. Hence the depth should be more precisely called the projected range, while the path length is called the CSDA range. The ratio between the projected range and the CSDA range of a particle beam is the detour factor. For protons of 50 and 250 MeV the detour factor is 0.9985 and 0.9989, respectively (Seltzer 1993).

The CSDA range for a given initial energy E_0 can be calculated numerically from the linear stopping power $\frac{\mathrm{d}E}{\mathrm{d}x}$:

$$R(E_0) = \int_{E_0}^0 \left(\frac{\mathrm{d}E}{\mathrm{d}x}\right)^{-1} \mathrm{d}E \approx \sum_{E_0}^0 \left(\frac{\mathrm{d}E}{\mathrm{d}x}\right)^{-1} \cdot \Delta E \tag{2.9}$$

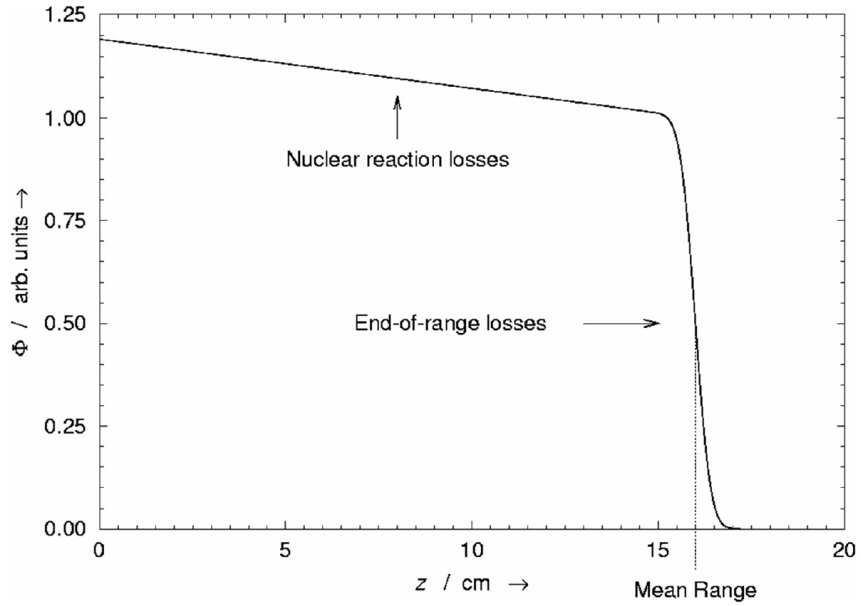

Figure 2.2: Relative fluence Φ of a proton beam as a function of depth z in water (Newhauser & Zhang 2015). Ratio of 1 corresponds to the number of protons just before the end–of–range. The gradual attenuation of protons up to that depth is caused by nuclear interactions. The steep reduction at the end of range is caused by ions coming to rest. The sigmoid shape of that area is caused by stochastic fluctuations in the energy loss of individual protons.

The summation is an accurate numerical approximation when performed over small discrete energy steps ΔE, during which the stopping power remains nearly constant. The CSDA range for protons in liquid water as a function of energy was calculated from the NIST database (Seltzer 1993) and is shown in figure 2.3. The CSDA range from about 1 MeV to 300 MeV seems to follow a power law. This behavior was utilized by Bragg and Kleemann who formulated the Bragg-Kleemann rule (Bragg & Kleeman 1905) parameterizing the CSDA range as a function of the initial energy of a particle beam:

$$R(E) = \alpha E^p \tag{2.10}$$

where as in equation 2.3 α is a material dependent constant, E is the initial energy of the particle beam and p is a dimensionless constant depending on the particle type and the beam energy. For proton beams with energy below 200 MeV traversing water, $\alpha \approx 2.2 \times 10^{-3}$ cm/(MeV)^p and $p \approx 1.77$ (Bortfeld 1997).

2.2.4 Energy – Range straggling

As mentioned in the previous sections the energy loss is a stochastic effect and equations 2.4 and 2.9 provide mean values of energy loss and of range for a beam of particles. The

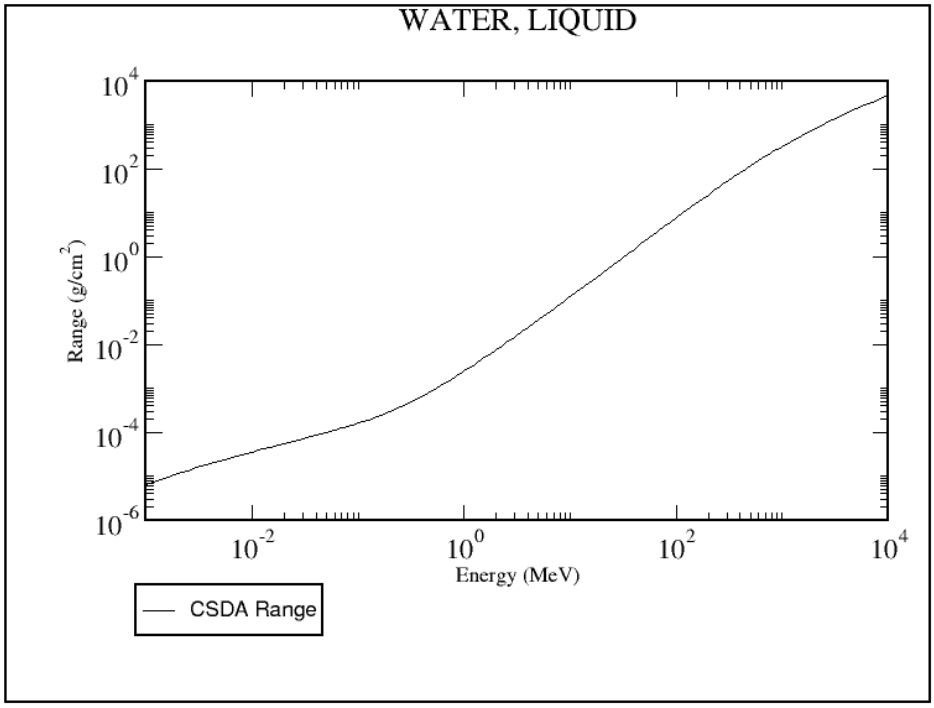

Figure 2.3: Proton CSDA range in liquid water, expressed in g/cm^2 (length multiplied by the mass density, in this case $1\,g/cm^3$) for a wide range of energies in MeV. The figure was obtained using the NIST database (Seltzer 1993).

fluctuations of the energy loss within a beam lead to each particle of the beam travelling a slightly different length until coming to rest and are responsible for the sigmoid shape of the beam fluence at the end of range shown in figure 2.2. These fluctuations are called energy straggling or range straggling.

The shape of probability density function (PDF) of these fluctuations depends on the absorber thickness. Bohr (1915), Landau (1944) and Vavilov (1957) have formulated theories describing the energy straggling PDF for thick, intermediate and thin absorbers, respectively. For most radiation therapy applications we are concerned with thick absorbers, therefore the energy straggling distribution as described by Bohr's theory will be briefly presented.

For thick absorbers the energy straggling is distributed according to a Gaussian:

$$f(\Delta E)\mathrm{d}\Delta E = \frac{1}{\sigma_{\Delta E}\sqrt{2\pi}} \exp \frac{-\left(\Delta E - \overline{\Delta E}\right)^2}{2\sigma^2_{\Delta E}} \tag{2.11}$$

where ΔE is the energy loss in a single interaction with an atom, $\overline{\Delta E}$ is the mean energy loss in the absorber. For relativistic ions the variance $\sigma^2_{\Delta E}$ of the distribution is given by:

$$\sigma^2_{\Delta E} = 2\pi r_e^2 m_e c^2 \rho N Z x \frac{z^2}{\beta^2} \Delta E_{max} \tag{2.12}$$

where r_e is the classical electron radius, m_e is the mass of the electron, c is the speed of light, ρ is the mass density of the material, N is the number of atoms per unit volume of material, Z is the atomic number of the absorber, x is the thickness of the absorber, z is the charge of the projectile, β is the projectile's velocity to the speed of light ratio and ΔE_{max} is the maximum possible energy transfer in one interaction.

A simple parametrization of the range straggling as a function of the ion range has been proposed by Chu et al. (1993):

$$\sigma_{\Delta R} = k R^m A^n \tag{2.13}$$

where R is the ion range, A the ion mass number and k, m and n are empirically determined constants. For protons the constant n vanishes as $A = 1$. Bortfeld (1997) determined for protons in water $k = 0.012$ and $m = 0.935$ which results in the empirical rule of proton range straggling (one standard deviation) in water being approximately 1% of the range.

2.3 Coulomb scattering

As briefly mentioned in section 2.1, protons can be elastically scattered by nuclei in the absorber, which results in change of direction and negligible energy loss. Adopting the description of Newhauser & Zhang (2015), elastic Coulomb scattering can be classified by the number of scattering interactions N_s happening in an absorber. For $N_\mathrm{s} = 1$ (single scattering) and adequate description can be obtained via Rutherford's scattering theory. For $1 < N_\mathrm{s} < 20$ (plural scattering) theoretical modelling becomes complicated. For $N_\mathrm{s} \geq 20$ (multiple scattering) which is relevant to particle therapy applications, the net effect of a large number single scattering interactions can be modelled using a statistical approach.

Figure 2.4 shows schematically the concept of multiple Coulomb scattering in two dimensions. From the passage in an absorber of thickness x, the myriad of single scattering interactions (solid curved line) sum up to a net angular deflection θ with respect to the initial direction and a net displacement y from the entry point in the absorber. To characterize the net angular deflection we define the quantity of scattering power:

$$T = \frac{\mathrm{d} < \theta^2 >}{\mathrm{d}x} \tag{2.14}$$

The statistical description of multiple Coulomb scattering aims at modelling the distribution of these net deflection angles for a large number of particles. Therefore we are

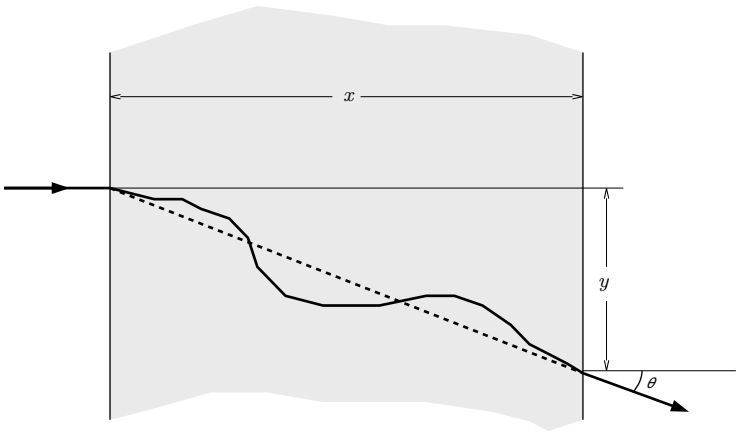

Figure 2.4: Schematic representation of a particle trajectory in an absorber of thickness x. The particle experiences a large number of single scattering events which result in a net deflection θ and displacement y.

interested in the statistical moments of that distribution, hence the scattering power is related to the mean deflection angle over a large number of particles. In addition, as the mean is expected to be at zero, in order to have a meaningful characterization of the deflection angle distribution we rather use the mean of the square of the deflection angles.

A complete theory of multiple Coulomb scattering was formulated by Molière (1948), under the assumption of small angle scattering ($sin\theta \approx \theta$). The probability $P(\theta)$ for scattering with angle θ into a solid angle $\mathrm{d}\Omega$ is:

$$P(\theta)\mathrm{d}\Omega = \eta \mathrm{d}\eta \left(2 \cdot e^{-\eta^2} + \frac{F_1(\eta)}{B} + \frac{F_2(\eta)}{B^2} + \ldots \right) \tag{2.15}$$

where $\eta = \theta / \left(\theta_1 B^{1/2}\right)$, $\theta_1 = 0.3965 \cdot (Z \cdot z/p\beta)\sqrt{\rho \delta x / A}$, z is the charge of the particle, p is the momentum of the particle, β is the ratio of the velocity of the particle to the speed of light, Z and A are the atomic and mass numbers of the absorber, ρ is the mass density of the absorber and δx is the traversed thickness. The functions $F_{\mathrm{i}}(\eta)$ and the parameter B have been tabulated in published literature.

A simplified model was developed by Gottschalk et al. (1993) approximating the deflection angle distribution with a Gaussian functional form:

$$P(\theta) \approx \frac{2\theta}{< \theta^2 >} \exp\left(\frac{-\theta^2}{< \theta^2 >} \right) \mathrm{d}\theta \tag{2.16}$$

Further work on simplifying Molière's theory (Gottschalk 2010) led to a differential approximation and the following widely used expression for the scattering power:

$$T = \frac{\mathrm{d} < \theta^2 >}{\mathrm{d}x} = f_{\mathrm{dM}}(pv, p_0 v_0) \cdot \left(\frac{E_{\mathrm{S}}}{pv}\right)^2 \cdot \frac{1}{X_{\mathrm{S}}} \tag{2.17}$$

where $E_{\mathrm{S}} = 15\,\mathrm{MeV}$, p_0 and v_0 are the initial momentum and velocity of the particle, p and v are the momentum and velocity of the particle after traversing thickness x of the absorber and X_{S} is the scattering length of the material in cm given by:

$$\frac{1}{X_{\mathrm{S}}} = \rho \alpha N_{\mathrm{A}} r_e^2 \frac{Z^2}{A} \left[2 \ln(33219 \cdot (AZ)^{-1/3}) - 1\right] \tag{2.18}$$

where ρ is the mass density of the absorber, α is the fine structure constant, N_{A} Avogadro's number, r_e is the electron radius, Z and A are the atomic and mass numbers of the absorber.

The term f_{dM} in equation 2.17 is a material-independent factor given by:

$$\begin{aligned} f_{\mathrm{dM}} = 0.5244 + 0.1975 \log_{10} &\left[1 - \left(\frac{pv}{p_0 v_0}\right)^2\right] \\ &+ 0.2320 \cdot \log_{10}\left(\frac{pv}{\mathrm{MeV}}\right) - 0.00981 \cdot \log_{10}\left(\frac{pv}{\mathrm{MeV}}\right) \cdot \log_{10}\left[1 - \left(\frac{pv}{p_0 v_0}\right)^2\right] \end{aligned} \tag{2.19}$$

2.4 Nuclear interactions

Non–elastic nuclear reactions cause fragmentation of the target nucleus in the case of proton beams, and of the target nucleus and the projectile in case of heavier ion beams. This results in a gradual attenuation of the beam before particles reach their CSDA range (see also figure 2.2). Figure 2.5 (Durante & Paganetti 2016) highlights this effect for different ion species in water, all having the same CSDA range of approximately 20 cm. It is worth noting that beam attenuation which is already non–negligible for protons (approximately 80% of the initial protons reach near their CSDA range of 20 cm), becomes a significant effect for heavier ions (approximately 50% of the initial ^{12}C-ions reach near their CSDA range of 20 cm).

Several non–elastic nuclear cross section (σ_{R}) parameterizations have been proposed in published literature. One of them assuming simple geometric approximations is that of Bradt & Peters (1950):

$$\sigma_{\mathrm{R}} = \pi r_0^2 \left(A_{\mathrm{p}}^{1/3} + A_{\mathrm{T}}^{1/3} - b\right)^2 \tag{2.20}$$

where r_0 is the nucleon radius, A_{p} and A_{T} the mass number of the projectile and the target, and b a correction factor (overlapping factor).

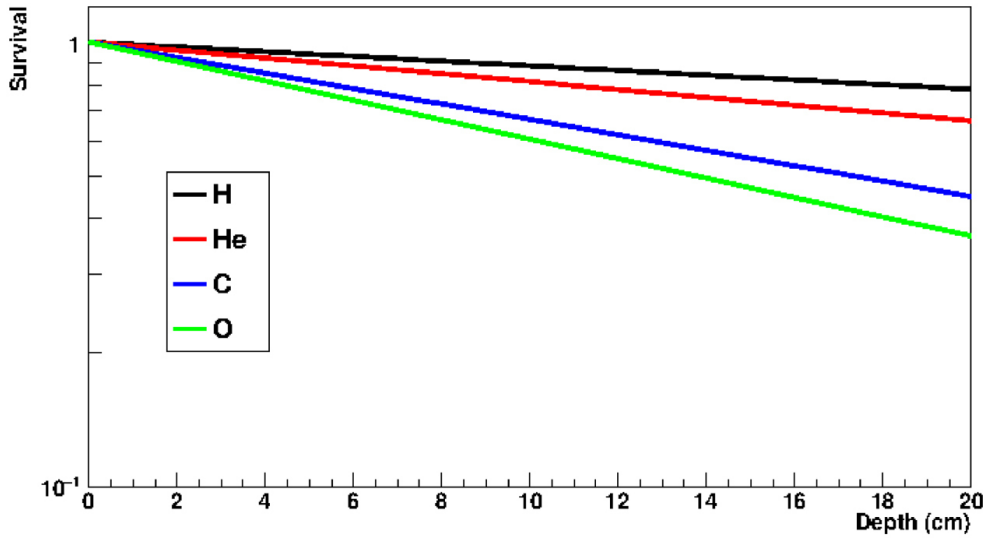

Figure 2.5: Beam attenuation as a function of depth in water, for different ions of the same CSDA range of approximately 20 cm (Durante & Paganetti 2016).

For proton beams of particle therapy relevant energies an energy dependent parameterization derived by Sihver & Mancusi (2009) is:

$$\sigma_{\mathrm{R}}(E) = \sigma_0 \cdot f(E, Z_{\mathrm{T}}) \tag{2.21}$$

where σ_0 is a Bradt–Peters factor determined as:

$$\sigma_0 = \pi r_0^2 \left[1 + A_{\mathrm{T}}^{1/3} - b_0 \left(1 + A_{\mathrm{T}}^{-1/3}\right)\right]^2 \tag{2.22}$$

where b_0 is called the transparency parameter and is a polynomial expansion of the target's mass number:

$$b_0 = 2.247 - 0.915 \left(1 + A_{\mathrm{T}}^{1/3}\right) \tag{2.23}$$

For proton energies higher than 200 MeV the factor $f(E, Z_{\mathrm{T}})$ is very close to one and can be omitted, but for lower energies it needs to be accounted for and the parameterization can take complicated functional forms.

For protons traversing biological relevant materials the non–elastic nuclear cross section displays a threshold at approximately 8 MeV and it increases to a maximum at about 20 MeV, the latter corresponding to about 0.4 mm residual range in water. It thereafter decreases up to approximately 100 MeV, equivalent to about 7.7 mm residual range in water. From there and on it stays almost constant as a function of energy with a minor

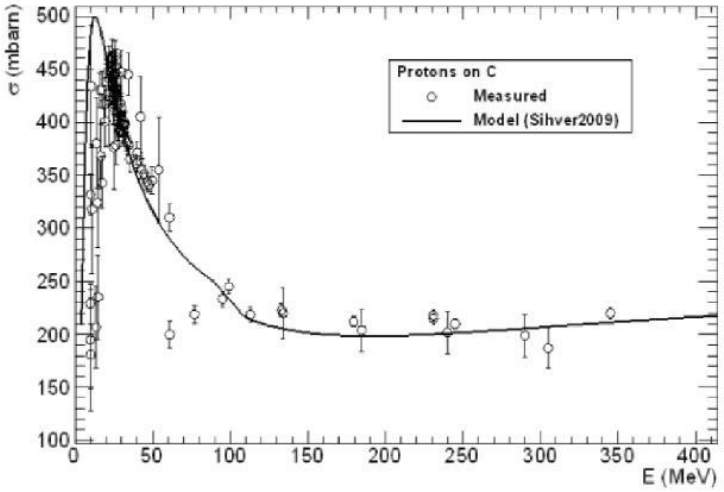

Figure 2.6: Non–elastic cross section of of protons on carbon as a function of energy (Durante & Paganetti 2016).

positive slope. An example of the proton non–elastic cross section in carbon as a function of energy is shown in figure 2.6 (Durante & Paganetti 2016).

Secondary fragments from non–elastic nuclear interactions contribute to the dose deposited by the therapeutic beam. Figure 2.7 (Grassberger & Paganetti 2011) shows an example of the most important dose components stemming from a 160 MeV proton beam in water. Dose from secondaries sums up to approximately 10% of the total dose, with the most significant contributions originating from secondary protons and followed by α–particles. Except for charged secondaries, photons and neutrons are also created as products of non–elastic nuclear reactions. Their contribution to the therapeutic dose is typically negligible in the treated area. Nevertheless, as their range is considerably larger than that of charged secondaries, they can deposit dose in healthy tissues far from the treated site. Especially neutrons have an energy dependent biological effectiveness up to 20 times higher than that of protons (ICRP 2007) and can increase the risk of second cancer due to radiation therapy (Newhauser & Durante 2011).

2.5 Characteristics of the Bragg curve

All the aspects discussed in the previous sections of this chapter contribute in their own way to the distinct shape of the dose as a function of depth. The so–called Bragg curve was first observed by Bragg & Kleeman (1905) for α–particles. In this section the main characteristics of the Bragg curve will be discussed.

Figure 2.8 (Newhauser & Zhang 2015) shows the absorbed dose of a monoenergetic proton beam as a function of depth in the absorber. The longitudinal dose distribution has this distinct shape as a function of depth when plotted along the central axis of a

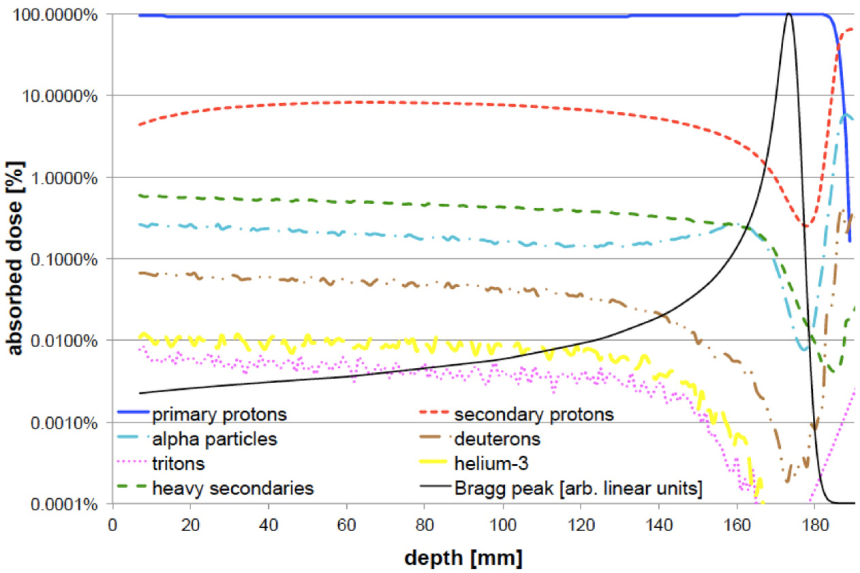

Figure 2.7: GEANT4 simulation of the absorbed dose as a function of depth of a 160 MeV proton beam in water (Grassberger & Paganetti 2011). The different contributions to the dose are shown as a percentage of the absorbed dose at that depth.

broad beam or when integrating the dose laterally at every depth. This happens due to the equilibrium of lateral scattering away from the central axis exactly compensated by scattering towards it. Otherwise the dose maximum at the end of range is severely reduced and might not even be the maximum of the distribution. The general shape of the Bragg curve as depth increases (from left to right in the figure) is characterized by a small buildup at shallow depths, a rather flat plateau area and a pronounced dose maximum followed by a sharp fall–off.

The first region of the Bragg curve is the *electronic buildup*. It is located near the entrance to the absorber, at the beginning of the dose curve. It is formed due to the fact that the proton beam will cause there the first ionizations and the produced δ-electrons have enough energy (tens to hundreds of keV) to travel deeper in the absorber, further ionize and deposit dose. The presence of small amounts of material before the absorber is enough to create charged particle equilibrium and disguise the electronic buildup.

The *protonic buildup*, observed near the entrance of the beam to the absorber is the region where an increase in the dose is observed due to the buidlup of secondary protons. These secondary protons originate from non–elastic nuclear interactions.

The region extending from the entrance to the absorber, or from the electronic/protonic buildup, up to the proximal of the peak is called *sub–peak* region or *plateau*. In this area of the longitudinal dose profile the dose increases slowly and this effect is governed primarily by the increase of stopping power with decreasing energy ($S \propto \frac{1}{\beta^2}$). To a lesser extent, the dose is increased vs. depth due to the contribution from secondaries originating from non–elastic nuclear interactions. On the other hand, the gradual attenuation of the proton beam contributes negatively to the dose buildup.

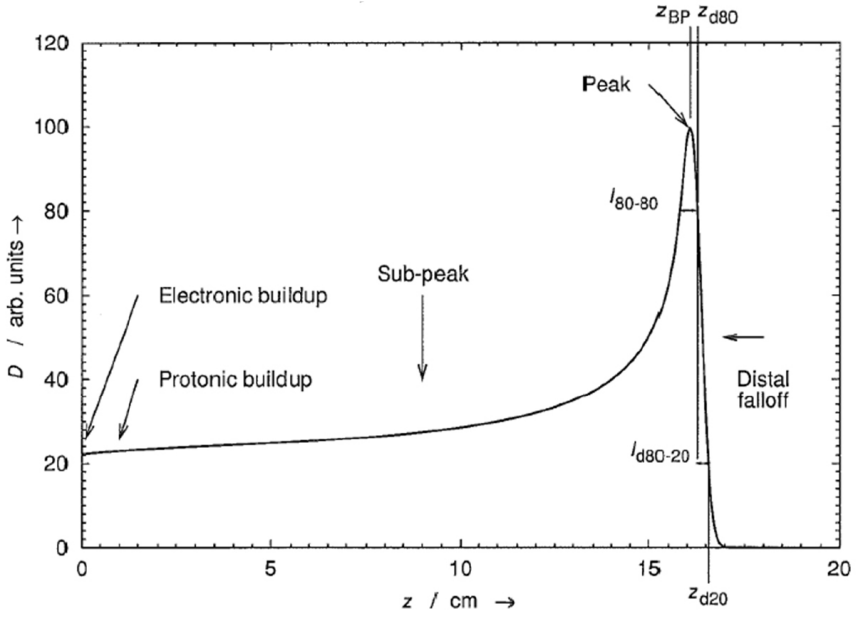

Figure 2.8: Absorbed dose D as a function of depth z in water from a monoenergetic 154 MeV broad proton beam (Newhauser & Zhang 2015). The various regions labelled are explained in the text.

The dose maximum near the end of range is the *pristine Bragg peak*. The increase of stopping power with decreasing energy and the beam attenuation due to non–elastic nuclear interactions determine the exact location and magnitude of the peak for a given beam energy. The width of the peak is governed by the energy straggling. The depth where the dose maximum occurs is denoted in figure 2.8 as z_{BP}. The width of the peak is often quantified by the distance between the proximal and the distal 80% of the dose maximum (I_{80-80}).

Beyond the pristine Bragg peak follows a sharp dose negative gradient named the *distal fall–off*. The steepness of this falloff is determined by the initial beam energy and by the energy straggling. It is usually quantified by the distance between two distal dose levels, for example the distance between distal 80% to 20% of the maximum dose ($I_{\mathrm{d80-20}}$). For heavier ions than protons a significant *fragmentation tail* falling slowly to zero follows (Kempe et al. 2006). It is caused by charged secondaries from non–elastic nuclear interactions which are lighter than the primary ions.

Chapter 3

Comparison of proton CT and dual energy x–ray CT

This chapter contains the work published in Physics for Medicine and Biology, volume 64, issue 16, page 165002, in August 2019, with the title *Experimental comparison of proton CT and dual energy x–ray CT for relative stopping power estimation in proton therapy*, by Dedes et al. (2019).

3.1 Introduction

The increased use of protons for external beam radiation therapy of cancer offers the potential of sparing of healthy tissues by achieving highly conformal dose distributions to the tumor. Protons gradually slow down when traversing matter and their initial energy can be chosen such that the therapeutic proton beam stops in the tumor. The range of the protons in the patient is dependent on their energy and on the stopping power of the tissue they traverse. The latter is commonly expressed as relative to water and denoted as relative stopping power (RSP). A highly accurate RSP map of the patient in treatment position is crucial for the calculation of optimal proton therapy treatment plans. Any inaccuracies in the RSP map will be translated to proton range prediction errors, thus compromising treatment plan quality by introducing range uncertainties, for which safety margins are required (Paganetti 2012).

The current practice for obtaining RSP images is based on converting x–ray linear attenuation coefficients, acquired with single energy x–ray CT (SECT) imaging, to RSP. This procedure involves a stoichiometric calibration (Schneider et al. 1996) which can result in up to 3.5% errors in the determination of RSP (Paganetti 2012, Yang et al. 2012). Proton CT (pCT) is considered as a candidate for improving RSP accuracy. Already in 1963, Cormack (1963) proposed that protons could be used for tomographic imaging. Huesman et al. (1975) described a pCT scanner concept relying on the registration of individual protons' positions and directions, as well as the residual energy behind the

patient. pCT scanners effectively measure RSP line integrals which can be processed by dedicated reconstruction algorithms (Penfold et al. 2009, Penfold et al. 2010, Rit et al. 2013, Poludniowski et al. 2014, Hansen et al. 2016) to yield three–dimensional RSP maps.

Prototype pCT scanners have been designed and constructed (Takada et al. 1988, Coutrakon et al. 2013, Sadrozinski et al. 2016, Johnson et al. 2016, Pettersen et al. 2016, Esposito et al. 2018). In recent studies, the RSP accuracy achieved by some of these prototypes has been reported to be better than 1.6% for three inserts (Esposito et al. 2018) and 1.39% for seven inserts (Giacometti et al. 2017a). For the same seven inserts, Volz et al. (2018) achieved accuracy better than 1% using helium ions with a pCT prototype scanner.

An alternative for obtaining RSP images of high accuracy is dual energy x–ray CT (DECT) (Yang et al. 2010), where several studies (Hünemohr et al. 2013, Bourque et al. 2014, Hudobivnik et al. 2016, Möhler et al. 2016, Han et al. 2016, Taasti et al. 2016, Lalonde et al. 2017, Saito & Sagara 2017b, Almeida et al. 2018) have demonstrated the potential of reaching an RSP accuracy of about 1%. DECT methods for RSP estimation have been recently validated using biological tissue samples (Taasti et al. 2017, Bär et al. 2018, Xie et al. 2018, Möhler et al. 2018) and have been found to consistently outperform SECT in terms of RSP accuracy. DECT scanners are currently making their way into proton therapy clinics (Wohlfahrt et al. 2017a) and may impact clinical proton range calculation (Hudobivnik et al. 2016, Wohlfahrt et al. 2017b).

While a comparison between ideal simulated pCT and DECT measurements in terms of RSP accuracy has suggested that pCT may achieve superior results (Hansen et al. 2015), there has so far been no direct comparison between the two modalities. This study aims at filling this gap.

The current study is split in two. The main part presents for the first time an experimental comparison of pCT and DECT in terms of RSP accuracy. For that purpose, two different phantoms containing a total of 13 tissue equivalent inserts of known RSP have been scanned in the phase II preclinical pCT prototype scanner and in a commercially available DECT scanner. This allowed a direct comparison of the accuracy of the resulting RSP images.

Since pCT technology is at a much earlier stage of development than state–of–the–art dual source DECT scanners, we expect that pCT images may suffer from artifacts and other effects absent in DECT. In the second part of our study, we thus supplemented our experimental work with ideal and realistic pCT detector simulations. Validation of the realistic simulations against experimental results allowed us to assess whether pCT accuracy may be impacted by fundamental limitations of using protons for imaging, or from design aspects of the pCT prototype we used. Additionally, the simulations were used to pinpoint sources of image artifacts in the pCT scanner design.

3.2 Material and methods

3.2.1 Experimental aspects

Known-RSP phantoms

Two phantoms containing plastic tissue equivalent inserts were used in this study (see drawings in figure 3.2). The first phantom is the CTP404 module from the commercial multislice Catphan®600 phantom (The Phantom Laboratory, New York, USA), denoted for simplicity in the rest of the manuscript as CTP404 phantom. The CTP404 phantom is of cylindrical shape with a diameter of 150 mm and height of 25 mm, made of polystyrene. The phantom contained eight radially placed cylindrical inserts of 12.2 mm radius and 25 mm height. It also contained other smaller inserts of various materials (tungsten carbide wires, air and Teflon small cylindrical inserts and acrylic spherical inserts of various radii) which are not relevant to this study. During the imaging scans, two of the inserts were filled with air, while the rest contained materials whose RSP is summarized in table 3.4. The reference RSP was obtained by means of variable water column measurements with a 4.4 mm FWHM 310.82 MeV/u carbon ion beam (Giacometti et al. 2017a).

The second phantom is a custom–made acrylic (PMMA) cylinder of 130 mm diameter and 227 mm height, called henceforth LMU phantom. The phantom body contained seven cylindrical holes of 30 mm diameter and of 55 mm height, in which different plastic tissue equivalent materials were inserted. The seven inserts were arranged in three different layers along the axis of the cylinder of the phantom body. In two of the layers one insert was placed centrally and one off–center and in a third layer one insert was placed centrally and two off–center. Due to the fact that the LMU phantom was in the vertical direction larger than the scanner's field of view, its three insert layers were scanned in three different runs. Therefore, the LMU phantom is presented as three different images, hereafter called Top Supremum (*TopSup*), Top Infimum (*TopInf*) and Bottom (*Bot*). The reference RSP of the seven inserts were obtained by means of variable water column measurements with a 4.4 mm FWHM 310.82 MeV/u carbon ion beam (Berndt 2016, Hudobivnik et al. 2016) and are listed in table 3.4.

Proton CT scanner

The proton CT scans of this study were performed at the Northwestern Medicine Chicago Proton Center using the phase II preclinical pCT prototype scanner of the Loma Linda University and the of University of California Santa Cruz. The scanner hardware is described in detail in Johnson et al. (2016). It comprises two tracking modules and an energy detector for the determination of the water–equivalent path length (WEPL). The data acquisition system is capable of acquiring broad beam proton events at a sustained rate in excess of 1 MHz.

The front and rear tracking modules (upstream and downstream of the scanned object) contain in total 32 single–sided silicon strip detectors (SSD) with a strip pitch of 228 µm and a thickness of 400 µm. In each tracking module, the silicon strip detectors are arranged

in four layers, two measuring the horizontal coordinates and two measuring the vertical coordinates. The tracking system provides the capability of a four–point measurement for each proton, allowing the estimation of the curved proton path using the measured positions and calculated directions.

The WEPL detector (Bashkirov et al. 2016) consists of five polystyrene scintillator stages (RSP $\approx$ 1.038), each with a thickness of 51 mm and a lateral area of $10\times40\,\mathrm{cm}^2$. Each scintillator stage is wrapped with 65 µm thick reflective material (Vikuit$^{\mathrm{TM}}$ ESR film). The WEPL of a proton is deduced from the energy detector signal of the stage in which the proton stopped, using a stage–specific energy–to–WEPL calibration curve. The calibration procedure is described in section 3.2.3.

The WEPL calibration, as well as the conversion of the digitized signal from the tracker (strip number) to physical coordinates, and the assembly of proton events from information registered by the different parts of the detector are performed by dedicated raw data processing software. The output of the processing software is fed to the reconstruction algorithm (described in section 3.2.4) which produces a voxelized RSP map of the scanned object.

Both the data acquisition required for the WEPL calibration and the imaging runs were performed with a wobbled proton beam of 40 mm FWHM size, which was magnetically deflected to sweep the field of view (FOV) of the scanner. The nominal proton energy was 200 MeV as defined in the accelerator library and the variation was found to be less than 0.2 MeV between runs. Finally, the phantoms used for imaging were placed on a rotating stage and data were acquired while continuously rotating. Data were subsequently binned in 360 projections, spaced at 1 degree steps. pCT scans took approximately 6 min.

Dual energy CT scanner and conversion to RSP

RSP estimates based on DECT images were calculated following a method proposed by Saito & Sagara (2017a) and Saito & Sagara (2017b). For this purpose a PMMA calibration phantom with 150 mm diameter and a central bore housing inserts from an electron density phantom (RMI 467, Gammex, Middleton, USA, part of the Sun Nuclear Corporation) was scanned using a dual–source DECT scanner (SOMATOM Definition FORCE, Siemens Healthineers, Forchheim, Germany) at peak tube voltages of 90 kVp and 150 kVp with tin filtration. The scan pitch was 0.7 and exposures of 168 mAs were set for both x–ray tubes with automatic exposure control disabled. The CT dose index ($\mathrm{CTDI_{vol}}$) was 35.7 mGy. Images were reconstructed using the Q32\3s image reconstruction kernel with the vendor's iterative reconstruction (ADMIRE) on a $0.39\,\mathrm{mm}\times0.39\,\mathrm{mm}\times3\,\mathrm{mm}$ grid. Both known–RSP phantoms were scanned the same way. DECT scans took 17 sec.

Using the mean CT numbers HU_k for the high-energy ($k = \mathrm{H}$) and low-energy ($k = \mathrm{L}$) scans of the calibration phantom materials, scanner specific calibration parameters were obtained by least–square fitting of the known electron densities relative to water ρ_e and the ratio of effective atomic number Z_eff to the following functions:

$$\rho_\mathrm{e} = a\frac{(1+\alpha)\mathrm{HU_H} - \alpha\mathrm{HU_L}}{1000} + b \tag{3.1}$$

$$\left(\frac{Z_{\text{eff}}}{Z_{\text{eff,w}}}\right)^{3.3} - 1 = \gamma_{\text{L}} \left(\frac{u_{\text{L}}}{\rho_{\text{e}}} - 1\right) + \gamma_0 \tag{3.2}$$

with the effective atomic number of water $Z_{\text{eff,w}} = 7.4774$ and reduced CT number $u_{\text{L}} = \text{HU}_{\text{L}}/1000 + 1$. The offset γ_0 in equation (3.2) was added to Saito's original proposal in order to improve the accuracy of the fit. The calibration of the mean excitation energy I relative to that of water I_{w} was done separately for soft tissues ($Z_{\text{eff}} < 8.8$) and bone tissues ($Z_{\text{eff}} > 8.8$) by fitting equation (3.3).

$$\ln \frac{I}{I_{\text{w}}} = C_1 \left[\left(\frac{Z_{\text{eff}}}{Z_{\text{eff,w}}}\right)^{3.3} - 1\right] - C_0 \tag{3.3}$$

The resulting fit parameters are summarized in table 3.1. The figures showing the fits on the data are shown in section 1 of the *Supplementary Material*.

Table 3.1: Calibration parameters (CP) required to convert DECT images to RSP, together with their 95% confidence level (CL).

	CP	CL		
α	0.3452	0.02		
a	0.9928	0.01		
b	0.9929	0.004		
γ_{L}	9.0814	0.3		
γ_0	-0.0941	0.08		
	soft tissues		bone tissues	
	CP	CL	CP	CL
C_1	0.2020	0.1	0.0662	0.009
C_0	0.0821	0.03	0.0945	0.03

Applying these calibration parameters to the scans of the known-RSP phantoms, voxelized maps of relative proton stopping power values could be obtained according to the Bethe equation:

$$\text{RSP} = \rho_{\text{e}} \frac{\ln\left(\frac{2m_{\text{e}}c^2\beta^2}{I(1-\beta^2)}\right) - \beta^2}{\ln\left(\frac{2m_{\text{e}}c^2\beta^2}{I_{\text{w}}(1-\beta^2)}\right) - \beta^2} \tag{3.4}$$

with rest electron mass m_{e}, speed of light c, and proton speed relative to the speed of light β. In this work, we used an I_{w} value of 78 eV according to the latest ICRU report (Sigmund et al. 2009) and a β value of 0.4282, corresponding to a proton kinetic energy of 100 MeV.

3.2.2 Proton CT Monte Carlo simulations

A GEANT4 (version 10.03.p01) based simulation application, modelling in detail all parts of the pCT phase II prototype scanner, has been developed and published by Giacometti et al. (2017a). The physics models used were the `G4EmLivermorePhysics` for the electromagnetic physics, the `G4HadronPhysicsQGSP_BIC_HP` for the inelastic interactions of hadrons, the `G4HadronElasticPhysicsHP` for the elastic interaction of hadrons and the `G4IonBinaryCascadePhysics` for the inelastic interactions of ions. In some particular cases nuclear/hadronic interactions were switched off, so as to investigate their effect on the RSP accuracy. Whenever this was the case, it is explicitly stated in the text, otherwise the full set of physics was used. The maximum step length was chosen to be 6 µm for the energy detector stages (one tenth of the wrapping material thickness) and 1 mm in the rest of the simulation geometry. For this study an amended version of the simulation code was used. In that version, the simulation includes parameterization of the non–linear response of the scintillator to the deposited energy (Birks' effect). The Birks' factor $k_B = 0.0887$ (Dickmann et al. 2019) was used to modify the deposited energy $\mathrm{d}E$ per step $\mathrm{d}x$ according to

$$\frac{\mathrm{d}E'}{\mathrm{d}x} = \frac{\mathrm{d}E/\mathrm{d}x}{1 + k_B \cdot \mathrm{d}E/\mathrm{d}x}. \tag{3.5}$$

The spatial dependence of the scintillator response (Bashkirov et al. 2016), related to the position of the hit with respect to the location of the photomultiplier tube was also implemented.

In addition, the simulation code emulates the digitization process of the real scanner, yielding raw data in the same format as the actual scanner. This means, for each proton the simulation outputs the tracker hits as strip numbers and the energy deposit in each stage of the WEPL detector as analog to digital converter (ADC) numbers. This allowed the use of a unified software workflow for experimental and simulated scans. The output of the simulation (calibration and imaging runs) is processed in the same way as the experimental data.

It is important to mention here that the four silicon strip detector modules comprising each layer of a tracking module are arranged in a 1×4 matrix configuration, therefore leaving three gaps in every layer. These gaps are 0.6 mm wide vertical stripes of insensitive areas. Gaps are offset horizontally from layer to layer in order to reduce the probability that a single proton crosses more than one of them. Special care was taken to minimize these gaps, and it was estimated that, at the interface between two silicon detectors there was an opening of about 0.1 to 0.2 mm, that was partially filled with glue. Nevertheless, in the simulations, the whole 0.6 mm wide insensitive area of every gap was modelled as air, which had implications on the reconstructed image quality, as shown later.

Besides the full detector simulation resulting in raw data aiming at closely modeling experimental data, the simulation is also capable of producing idealized data. In this case, the proton's exact position, direction and energy are scored before and after the object at planes coinciding with the trackers.

For simplicity, pCT scans were simulated in step–and–shoot mode, in contrast to the continuous acquisition in experiments. 360 projections at 1 degree steps were simulated for all phantom cases.

An additional water phantom was simulated in order to investigate specific image artifacts. It was modelled after an existing water phantom consisting of a cylindrical PMMA container with outer diameter of 150.5 mm and a height of 40 mm. The wall thickness of the PMMA container was 6.35 mm and the container was filled with distilled water. For the purposes of the current study, a larger version of the phantom (1.5 times larger in diameter) with an outer diameter of 225.75 mm was also simulated. Simulations were performed with both water cylinders centered at the imaging isocenter. To further investigate the artifact related to the tracker gap, a set of three realistic simulations of the 150.5 mm diameter water phantom was performed. In the first simulation, the phantom was centered at the isocenter and the tracker gaps were assumed to be filled with air, as in the case of all previous simulations. In the second simulation, the water phantom was placed with a lateral offset of 40 mm to the isocenter and the tracker gaps were assumed to be filled with air. Finally, in the third simulation, the water phantom was located 40 mm off the isocenter and the tracker gaps were assumed to be filled with silicon.

3.2.3 Proton CT scanner calibration

The signal from the five–stage detector for each proton is converted to WEPL via a calibration procedure. The concept is described in Bashkirov et al. (2016) and a detailed update based on the current calibration phantom is given in Piersimoni et al. (2017). The calibration phantom is made of a polystyrene wedge which provides a WEPL gradient due to the slopes of the wedge, as well as four polystyrene blocks, which when combined with the wedge allow sampling of the entire WEPL range in the five–stage plastic scintillator detector. The calibration runs, namely a run without any object and five runs (wedge alone and wedge plus 1–4 blocks), are performed at the beginning of a scanning session, all with 200 MeV protons. The run without any object is used in order to map the spatial dependence of the energy detector response as well as to provide a conversion from ADC counts (digitized energy detector signal) to energy. The runs with the wedge phantom are utilized for the creation of a look–up table, associating the known WEPL a proton traversed in the calibration runs to the energy deposit in the stopping stage of the five–stage detector. The known WEPL is obtained by calculating the length of the proton trajectory in the calibration phantom from the tracker information, assuming straight paths and knowing the calibration phantom's RSP and geometry. The look–up table, referred to as WEPL calibration, contains a WEPL value for each of the 340 energy deposit bins in each of the five stages (in total 1700 energy bins of 0.25 MeV bin width). To obtain an optimal calibration, fits and interpolations can be applied in regions where either a lack of sufficient data or geometrical effects distort the calibration. The standard practice so far was to make an attempt of correcting also for detector effects. This approach was based on the expectation that the calibration curve should be continuous and smooth. In this study we investigate whether this assumption is valid by toggling corrections.

All these effects and the corresponding corrections are explained in section 3.2.6. Subsequent imaging runs are processed using the look–up table and energy deposits of protons in the stopping stage are converted to WEPLs.

Calibration curve variants

Alternative calibration curves (table 3.2) were generated to assess whether the calibration procedure may contribute to image artifacts and decrease RSP accuracy. The calibration obtained from experimental data as described previously and including all corrections is referred to as *ExpCalib1*. A variant called *ExpCalib2* was derived by omitting all applied corrections (fits and interpolations). A last experimental calibration named *ExpCalib3* was derived by omitting only the corrections related to the stage interfaces. We generated additional calibrations using the simulation platform. The *ExpCalib1*–equivalent calibration from simulations is called *SimCalib1*. In *SimCalib2*, corrections related to the stage interfaces were omitted.

Table 3.2: Variants of the WEPL calibration with their main parameters.

Calibration name	Type of data	Interpolation correction	Stage interface correction
ExpCalib1	Experimental	Yes	Yes
ExpCalib2	Experimental	No	No
ExpCalib3	Experimental	Yes	No
SimCalib1	Simulation	Yes	Yes
SimCalib2	Simulation	Yes	No

3.2.4 Proton CT image reconstruction

The algorithm used to reconstruct the pCT images was a filtered back projection (FBP) implementation that accounts for the curved proton paths. A detailed description of the underlying principles of the algorithm is given in Rit et al. (2013). The path of every proton, curved due to multiple Coulomb scattering, is approximated by a most likely path (MLP) formulation, introduced by Schulte et al. (2008). For the determination of the curved path, the position and direction information from the tracking modules is necessary. To eliminate protons stemming from nuclear interactions, cuts on the energy and angular distributions were applied. The cuts rejected protons whose energy or angle were outside three standard deviations around their median energy and angle. For the cuts, protons were grouped together in $2\,\mathrm{mm} \times 2\,\mathrm{mm}$ pixels according to their position at the front tracker module. The resulting proton–per–proton data were binned in projection images with $1\,\mathrm{mm} \times 1\,\mathrm{mm}$ pixels. The projections were then filtered and back–projected. Finally, the pCT images were reconstructed as RSP maps in a grid of $1\,\mathrm{mm} \times 1\,\mathrm{mm} \times 1\,\mathrm{mm}$. The reconstruction was applied to both experimental and simulated data without change of parameters.

3.2.5 RSP accuracy quantification

The reconstructed pCT images, from simulations and experimental data as well as from the DECT experimental scans, were compared in terms of RSP accuracy. The latter was quantified for the cylindrical inserts of the phantoms as follows: in a cylindrical region of interest (ROI) in the image, concentric with the inserts, the mean RSP ($\mathrm{RSP}_{\mathrm{mean}}$) of the voxels in that region was calculated. The accuracy is then the difference of the mean RSP from the reference RSP ($\mathrm{RSP}_{\mathrm{ref}}$) in percentage:

$$\mathrm{RSP}_{\mathrm{acc}} = 100 \cdot \frac{\mathrm{RSP}_{\mathrm{mean}} - \mathrm{RSP}_{\mathrm{ref}}}{\mathrm{RSP}_{\mathrm{ref}}} \quad \%. \tag{3.6}$$

The ROI radius was chosen to have 50% of the radius of the cylindrical inserts. Furthermore, the ROIs spanned across 15 slices (1 mm each) for the pCT scans and across 5 slices (3 mm each) for the DECT scans. In addition to the RSP accuracy for each insert, the mean absolute percentage error (MAPE) from all inserts and for each imaging modality was calculated according to:

$$\mathrm{MAPE} = \frac{\sum_{i=1}^{n} |\mathrm{RSP}_{\mathrm{acc},i}|}{n}. \tag{3.7}$$

where n is the total number of inserts and $\mathrm{RSP}_{\mathrm{acc},i}$ is the accuracy in percent for every insert i as calculated from equation 3.6.

3.2.6 Influence of problematic WEPL intervals on image artifacts

The RSP value of a voxel following image reconstruction is obtained from potentially very different WEPL values, since it corresponds to the average of the projections' values over all projection angles. This is because protons intersecting a given voxel, but from different acquisition angles, traverse very different paths in the object. In the special case of cylindrical objects these paths are, in two dimensions, circle's chords. The length of these chords ranges from a minimum length, depending on the radial location of the voxel and the acquisition angle, up to the object's diameter. For example, central voxels in a cylindrical homogeneous object are crossed only by protons of the maximum possible WEPL for that particular object, whereas voxels at the edge of the cylinder will see a wider WEPL distribution. If a certain WEPL interval is systematically distorted by the scanner, this will lead to artifacts in localized regions in the image, depending on the phantom. There are two types of calibration curve regions which may introduce image artifacts due to inaccurate WEPLs: stage interfaces and intra–stage calibration curve kinks. Their WEPL intervals are presented in table 3.3.

Protons stopping near the interface of two stages of the energy detector are of particular interest since ambiguities in their signal may lead to image artifacts. Additionally, they can distort the calibration curve. There is a number of corrections which can be applied

in order to obtain a smooth curve in the region between two adjacent stages. Due to a threshold of 1 MeV in the minimum energy required at the stopping stage, the first four energy bins for every stage contain no WEPL value. They can be arbitrarily set to the WEPL value of one of the next non–zero WEPL energy bins. Furthermore, the last few highest energy bins of every stage are populated by a small number of protons. In order to obtain a smooth curve there, an extrapolation correction can be applied, using the values of lower energy bins. After the aforementioned corrections, there is a WEPL discontinuity between the last energy bin of a stage and the first energy bin of the next stage. This can be removed by using the first energy of the next stage in the extrapolation correction described previously. To summarize, the fact that protons might deposit part of their energy in non–active detector material at the interfaces or split their energy in adjacent stages, in addition to the applied energy thresholds, results in inaccuracies and uncertainties in the corresponding part of the WEPL calibration.

Intra–stage calibration curve kinks correspond to a discontinuity observed in the calibration curves due to the calibration phantom geometry. This discontinuity is observed at about 60 MeV, in every stage except for the fifth (last) stage. The kink originates from the interplay between the geometry of the calibration phantom, the calibration procedure and the fact that the beam is divergent. The kink, which results in a severe artifact if left uncorrected, can be mitigated to a large extent by interpolation correction, using parts of the curve before and after the kink region. However residual WEPL inaccuracies may remain. As opposed to the stage interface correction described above, this correction is detector independent.

Table 3.3: WEPL ranges corresponding to either interpolation of calibration curve kinks, or stage interfaces.

	kinks	stage interfaces
WEPL region	WEPL range / mm	
1	46.5 – 56.7	37.6 – 40.9
2	97.2 – 107.3	90.1 – 92.2
3	147.9 – 158.7	141.7 – 144.9
4	199.0 – 209.9	193.0 – 196.2

To investigate the impact of the WEPL regions in table 3.3 on the accuracy in the image, we calculated heatmaps in image domain showing the percentage contribution of a given WEPL range to a voxel of the reconstruction volume. This was done by thresholding a given slice of the reconstruction volume from the experimental pCT images to the nearest known RSP value and calculating a forward projection in parallel beam geometry. The resulting sinogram was set to 1 if its value was within the given WEPL range and to 0 otherwise. Disregarding filtering, the binary sinogram was then backprojected and divided by the number of query points in each summation. This resulted in an image with values in the range $[0, 1]$, which are 0 if the voxel is backprojected from WEPLs that are strictly

outside the given WEPL range, and 1 if the voxel is backprojected from WEPLs that are exclusively inside the given WEPL range. This, however, is not to be taken quantitatively, as filtering for CT reconstruction was neglected.

3.2.7 Proton CT imaging dose estimation

In the case of the pCT scans, no direct dose estimation was possible. Therefore, the imaging dose was calculated with the Monte Carlo simulation code described in 3.2.2. A dose grid of 1 mm × 1 mm × 1 mm was defined and the dose to material was scored in every voxel and for each projection. The dose delivered in a single scan was obtained by simply summing the doses from all projections. The number of simulated protons per scan (2.7×10^8) was chosen to be approximately equal to that in the experimental scans. The exact dose estimation was obtained by scaling the simulated dose with the factor required to match the number of protons registered by the scanner in simulations to that in each experimental scan. Only physical dose was considered.

3.3 Results

3.3.1 Proton CT – Dual energy CT comparison

Proton CT calibration

In figure 3.1 the WEPL calibration is plotted for experimental data with all corrections *(ExpCalib1)*, experimental data without any corrections *(ExpCalib2)* and for comparison, for simulated data with corrections *(SimCalib1)*. For more details about the calibration parameters see table 3.2. The WEPL ranges to which corrections were applied are listed in table 3.3. The relative agreement between *ExpCalib1* and *SimCalib1* varied from approximately 1% or lower for stages 1-3, up to 6% for stages 4 and 5. A severe kink appeared in the calibration without interpolation correction *(ExpCalib2)*, at approximately 60 MeV and was to a large extent corrected via interpolation *(ExpCalib1* and *ExpCalib3)*. Removing the additional corrections concerning the stage interfaces and high energy deposits *(ExpCalib2)* also led to a dramatic distortion of the calibration curve between adjacent energy detector stages. Nevertheless, although the corrections related to stage interfaces yielded a smoother curve, they exacerbated artifacts and led to reduced RSP accuracy. The RSP MAPE achieved with pCT when using *ExpCalib1* was 0.87% for experimental data and 0.86% for simulations. Maximum errors exceeded 1.5%. The optimal calibration applied to experimental and realistic pCT simulations was obtained with interpolation at the kink region, but without stage interface corrections. Therefore, unless mentioned otherwise, *ExpCalib3* was used for experimental pCT data and *SimCalib2* for simulated pCT data. Detailed RSP accuracy results from this optimized calibration variant are presented in the following paragraph, together with results for DECT.

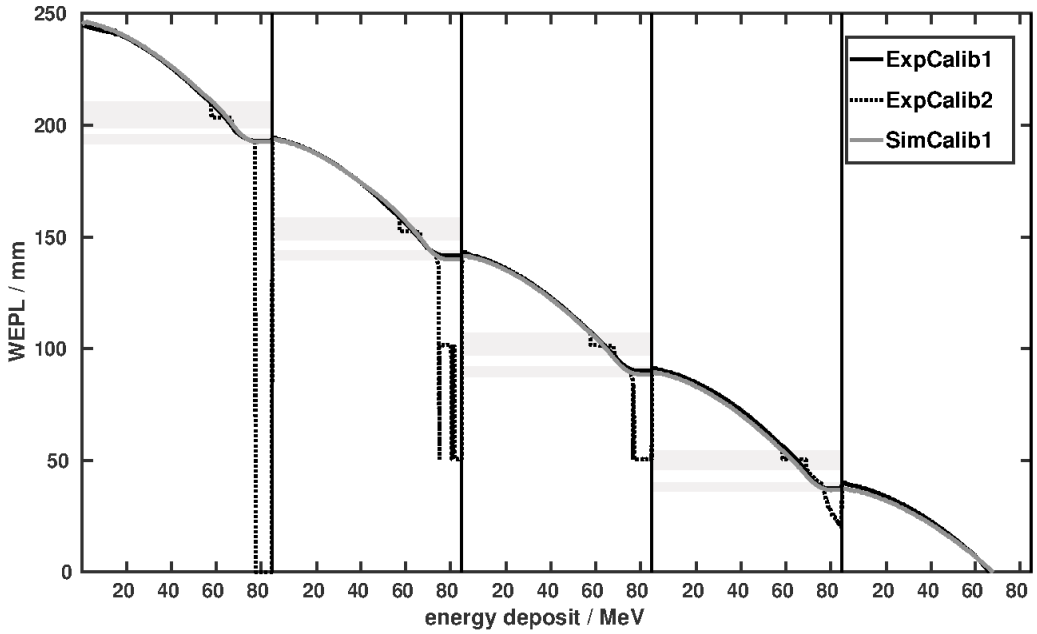

Figure 3.1: Calibration curves for the experimental (*ExpCalib1* and *ExpCalib2*) and simulated *(SimCalib1)* calibration runs. Vertical lines denote the transition between stages. Shaded areas indicate the WEPL ranges listed in table 3.3. The shallowest stage is on the left, and the deepest on the right. The step–like features located at approximately 60 MeV for the first four stages are the kink regions. *ExpCalib3* is not shown for clarity but would overlap with *ExpCalib2*, except for the kink regions where it would overlap with *ExpCalib1*.

RSP accuracy

The reconstructed RSP images from experimental pCT and experimental DECT are presented in figure 3.2. As explained in section 3.2.5, the RSP accuracy was quantified in cylindrical ROIs. The RSP accuracy as a function of the reference RSP is shown in figure 3.3 (top) for the experimental pCT and experimental DECT data, and in figure 3.3 (bottom) for the realistic simulation pCT and experimental DECT data. The experimental pCT RSP accuracy showed a tendency towards underestimation, which was well reproduced by the realistic simulation. For experimental pCT, inserts of PMP, Delrin and Teflon had errors exceeding 1% (1.08%, 1.16% and 1.31% respectively). In the case of realistically simulated pCT, all three central inserts had an RSP accuracy worse than 1% due to a detector modeling effect which will be discussed later. DECT values were more evenly distributed, and only cortical bone and Teflon had errors larger than 1% (1.17% and 2.38% respectively). These results are also summarized in table 3.4, in addition to the RSP accuracy obtained from ideal pCT simulations. The RSP MAPE achieved with the phase II preclinical prototype scanner was 0.55%. For realistic pCT simulations it was 0.69%, dominated by the central inserts. Without the central inserts, the realistic pCT

simulation MAPE was 0.50%, in good agreement with measurements. The RSP accuracy for DECT was 0.67%. In the case of ideal pCT simulations the RSP MAPE was below 0.2%.

The noise (one standard deviation) in a 25 mm diameter circular homogeneous ROI of the LMU phantom (outside of inserts) was 2.1×10^{-2} for experimental pCT and 5.0×10^{-3} for experimental DECT. For the CTP404 phantom, the standard deviations in a similar ROI were 2.6×10^{-2} and 5.1×10^{-3}.

Table 3.4: RSP accuracy from experimental pCT (pCT_{exp}), realistically simulated pCT (pCT_{sim}^{real}), ideally simulated pCT (pCT_{sim}^{ideal}) and experimental DECT scans. The inserts are ordered in increasing reference RSP values. The standard error of the mean was used to express the uncertainty on the estimated RSP accuracy. The ROI size for the CTP404 phantom was 319 pixels for pCT and 985 pixels for DECT. For the LMU phantom, it was 1773 pixels for pCT and 6285 pixels for DECT. The mean absolute percentage error (MAPE) is shown for each simulation mode and imaging modality.

Insert	Phantom	RSP_{ref}	pCT_{exp} (%)	pCT_{sim}^{ideal} (%)	pCT_{sim}^{real} (%)	DECT (%)
PMP	CTP404	0.88	1.08 ± 0.11	-0.07 ± 0.09	-0.22 ± 0.11	-0.64 ± 0.02
Adipose	LMU	0.97	-0.14 ± 0.04	-0.36 ± 0.03	-0.95 ± 0.04	-0.09 ± 0.01
LDPE	CTP404	0.98	-0.49 ± 0.11	-0.18 ± 0.08	-0.08 ± 0.10	-0.46 ± 0.02
Breast	LMU	0.99	-0.52 ± 0.04	0.05 ± 0.03	-0.39 ± 0.04	-0.25 ± 0.01
Polystyrene	CTP404	1.02	-0.04 ± 0.10	0.02 ± 0.08	-0.04 ± 0.10	0.43 ± 0.02
Muscle	LMU	1.06	-0.12 ± 0.04	-0.44 ± 0.03	-0.95 ± 0.03	-0.76 ± 0.01
Liver*	LMU	1.06	0.04 ± 0.03	-0.17 ± 0.03	-1.47 ± 0.03	-0.73 ± 0.01
Bone200*	LMU	1.11	-0.41 ± 0.03	-0.14 ± 0.03	-1.36 ± 0.03	0.48 ± 0.01
Acrylic	CTP404	1.16	-0.30 ± 0.10	-0.10 ± 0.07	-0.44 ± 0.09	0.49 ± 0.01
Bone400*	LMU	1.22	-0.84 ± 0.03	-0.44 ± 0.03	-1.11 ± 0.03	-0.50 ± 0.01
Delrin	CTP404	1.36	-1.16 ± 0.09	-0.01 ± 0.07	-0.45 ± 0.09	0.38 ± 0.02
Cort. Bone	LMU	1.69	-0.73 ± 0.02	-0.21 ± 0.02	-0.37 ± 0.02	1.17 ± 0.01
Teflon	CTP404	1.79	-1.31 ± 0.05	-0.06 ± 0.05	-1.11 ± 0.05	2.38 ± 0.01
MAPE %			0.55	0.17	0.69	0.67

* central insert

Proton CT imaging dose

The imaging dose in the pCT simulated scans was approximately 1.5 mGy for all phantoms, with variations less than 0.2 mGy in different inserts. This value was obtained considering that 7.5×10^5 protons were simulated per projection for a total of 2.7×10^8 protons in a scan with 360 projections. Scaling the dose calculated from simulations, as described in section 3.2.7, we estimated the dose in the experimental scans to vary from 1.5 mGy to 1.9 mGy, depending on the phantom. This is compatible to the dose measured with an ionization chamber during scans with the pCT phase II prototype scanner and reported by Johnson et al. (2017).

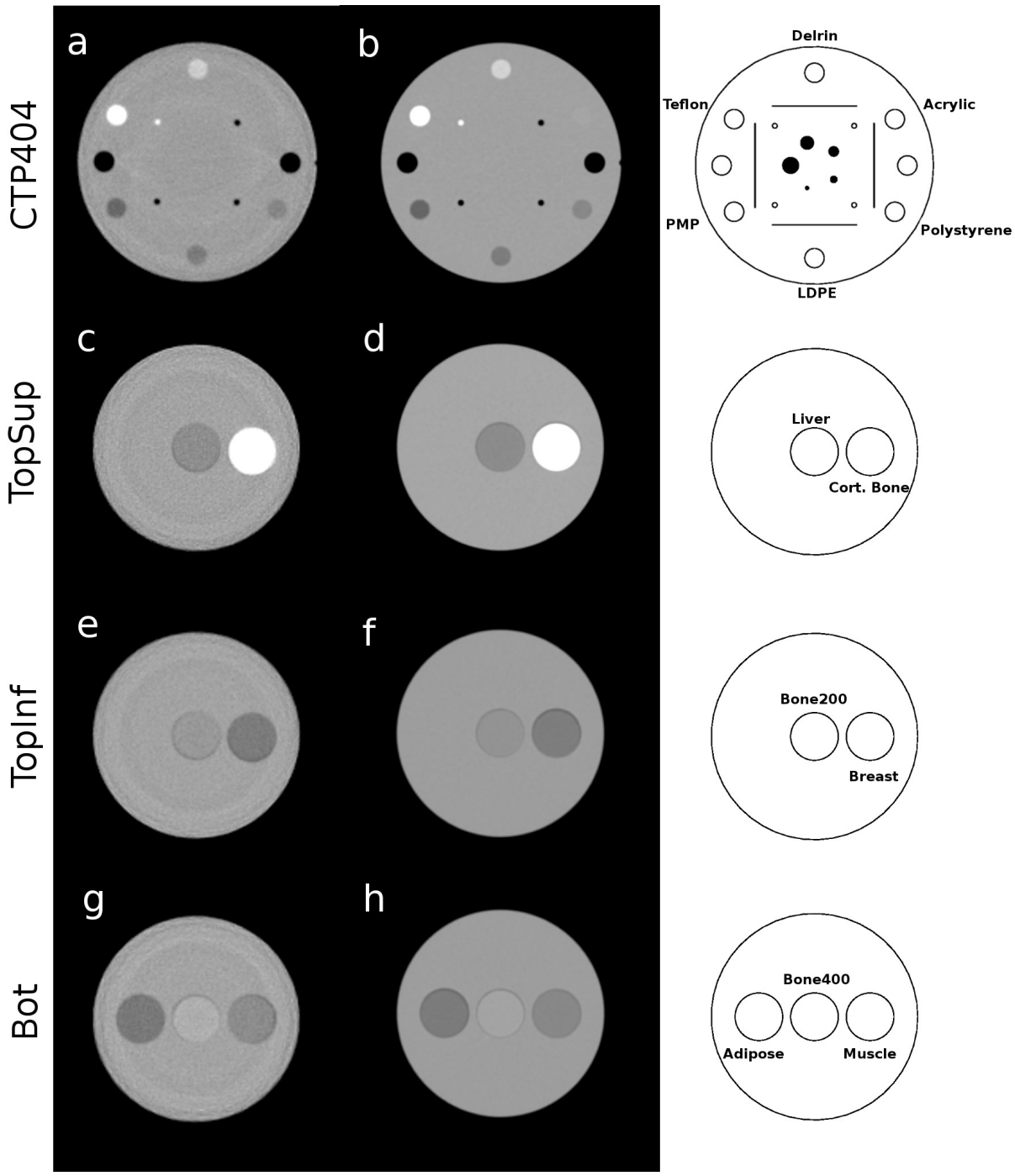

Figure 3.2: Reconstructed experimental pCT and DECT images. The left column contains images reconstructed from experimental pCT data and the middle column from measured DECT data. An RSP level of 1.0 and window of 1.5 were applied on the images for display purposes. The right column contains drawings of the phantoms with the insert materials labeled.

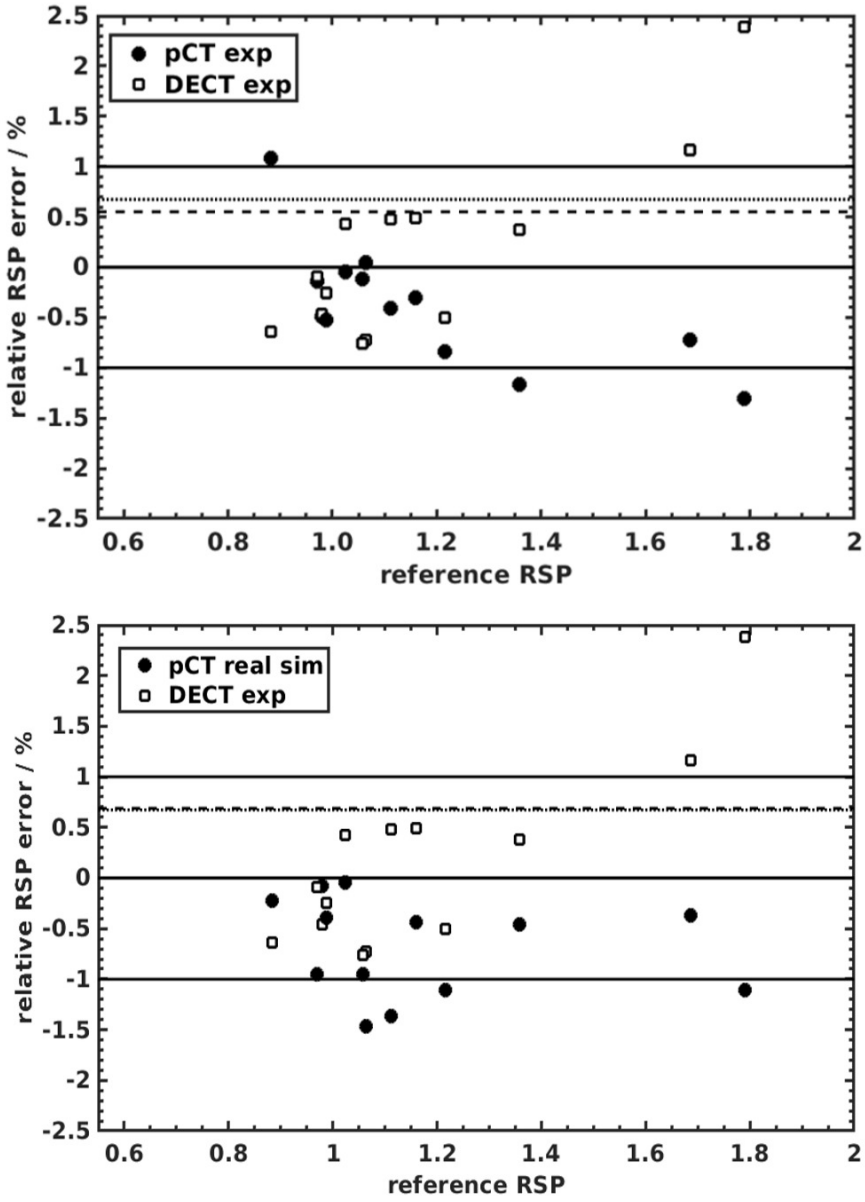

Figure 3.3: The RSP accuracy as a function of the reference RSP from (top) experimental pCT and (bottom) realistically simulated pCT is indicated with black circles. The RSP accuracy from experimental DECT is indicated with open square markers. The dashed and dotted horizontal lines indicate the MAPE for pCT and DECT, respectively.

3.3.2 Simulation investigations of pCT artifacts

Appearance of artifacts in proton CT images

In figure 3.4 the reconstructed pCT images of the CTP404 and the LMU phantom are shown. A narrow window (see figure caption) was chosen in order to highlight the image artifacts, along with averaging 15 slices to reduce noise. Similar ring artifacts were observed in experimental scans (figures 3.4 a, d, g, j) and realistic simulations (figures 3.4 b, e, h, k), but not in ideal simulations (figures 3.4 c, f, i, l). The right column of figure 3.4 shows a comparison of line profiles through the phantoms for the experimental and realistically simulated pCT scans. In the case of experimental scans the strongest artifacts exceeded 2% in RSP and appeared mostly as RSP overestimation. In the case of the realistic simulations, the strongest artifacts reached up to 2% in RSP and appeared mostly as RSP underestimation. In the CTP404 phantom, the artifacts appeared distorted by the inserts of higher or lower RSP, deviating from the appearance of conventional ring artifacts.

Nuclear interactions of protons in the scanned object can produce secondary protons or heavier ions, which have mostly lower energy than the primary protons traversing the same material and experience only electromagnetic interactions. These secondary particles, if not efficiently removed by the cuts, will result in an overestimation of the RSP. Switching off nuclear interactions in the simulations did not entirely remove these artifacts (results not shown), confirming our hypothesis that detector and calibration effects are the main source of artifacts.

Proton CT water phantom simulations

All images presented in this section were obtained with the realistic pCT simulation of water phantoms detailed in section 3.2.2 and were averaged over 15 slices. In figure 3.5 simulations of a pCT scan of the 150.5 mm diameter water phantom are shown for different modelling of the tracker gaps and placements of the phantom. This allowed to identify artifacts originating from the tracker. In the current implementation of the tracker geometry in the simulation, these gaps were overestimated by the assumption that they were filled solely with air. Protons traversing a gap experience a slightly lower WEPL (by approximately 0.8 mm). This results in their arrival to the five–stage energy detector with higher energy than nearby protons that lose some additional energy by going through an additional tracker layer. As shown in figure 3.5 (a) there is a lower RSP artifact (dark spot), located at the center of the water phantom when the latter is placed at the isocenter. When the water phantom was laterally displaced from the isocenter, as shown in figure 3.5 (b), the dark spot remained at the isocenter, not coinciding anymore with the center of the phantom. The dark spot almost disappears from figure 3.5 (c), when filling the tracker gaps with silicon instead of modelling air.

The images of the 150.5 mm and the 225.75 mm diameter water phantoms are shown in figure 3.6, where we observed that the location of the ring artifacts varied with the object diameter, an observation consistent with rings originating from specific, problematic WEPL ranges. Artifacts observed in the water phantom simulations were consistent in terms of

amplitude with these from the other phantoms, both in experimental and realistically simulated pCT scans. For example in figures 3.6 (a) and (b) they reached up to 1.5%.

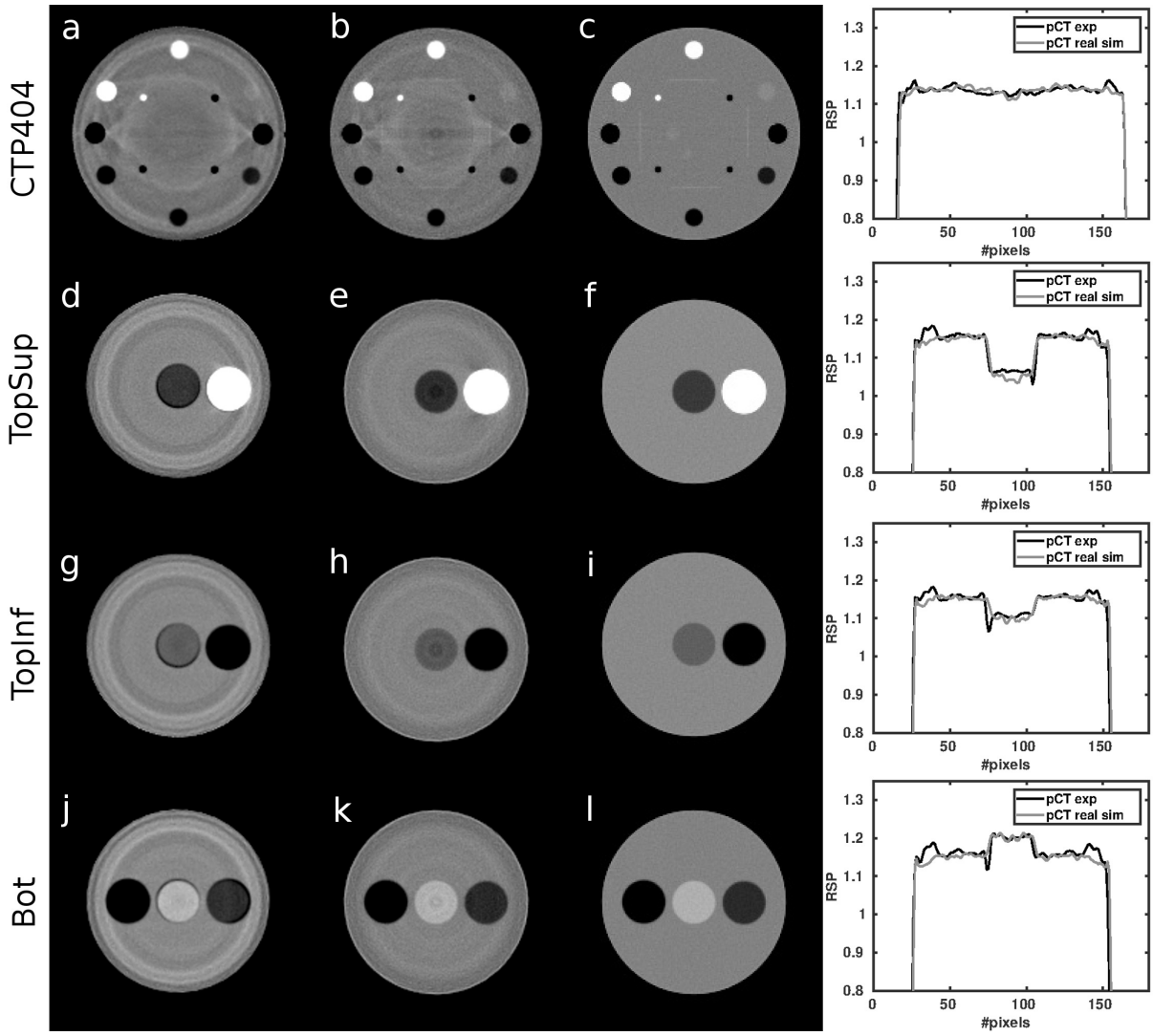

Figure 3.4: Reconstructed pCT images. The first column contains images reconstructed from experimental data, the second column from realistic simulations and the third column from ideal simulations. An RSP level of 1.15 and window of 0.3 were applied on the images in order to highlight the pCT image artifacts. The fourth column contains line profiles for the experimental and realistically simulated pCT images. For the LMU $TopSup$, $TopInf$ and Bot the profiles were obtained along the vertical diameter. For the CTP404 phantom the profile was obtained along the diameter that has a –30° angle with respect to the vertical and does not cross any visible insert. For all images and profiles 15 slices were averaged to better display artifacts.

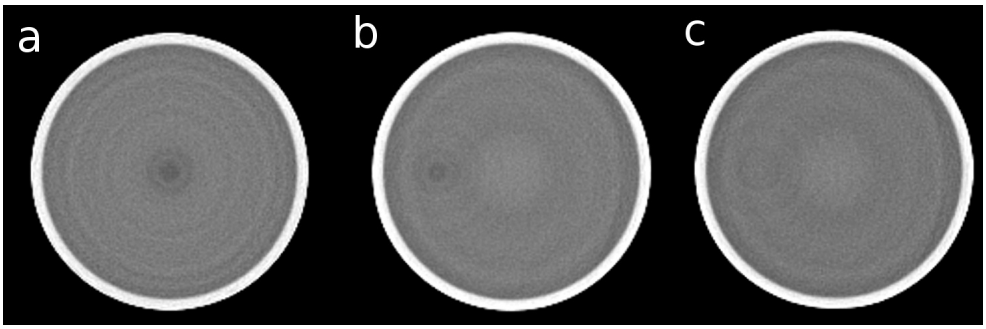

Figure 3.5: pCT images from realistic simulations processed with the *SimCalib2* calibration. In (a), the water phantom was aligned to the isocenter, while in (b) and (c) it was shifted by 40 mm with respect to the isocenter. In (a) and (b) the tracker layer gaps were filled with air. In (c), the tracker layer gaps were filled with silicon. An RSP level of 1.0 and window of 0.3 were applied on the images for display purposes.

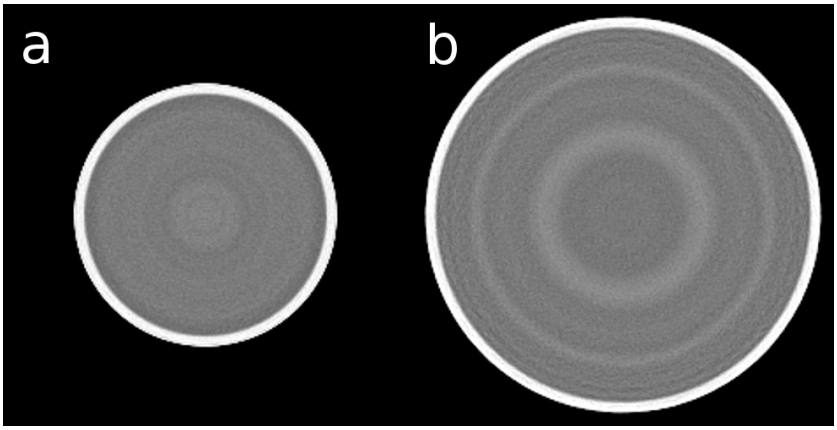

Figure 3.6: pCT images from realistic simulation processed with the *SimCalib2* calibration and tracker gaps filled with silicon. In (a) the 150.5 mm diameter water phantom and in (b) the 225.75 mm diameter water phantom. An RSP level of 1.0 and window of 0.3 were applied on the images for display purposes.

WEPL analysis

Following the methods described in section 3.2.6, we have identified WEPL value ranges (see table 3.3) that correspond to ambiguities and uncertainties in the calibration due to the four stage interfaces and to the kinks (see figure 3.1). Figure 3.7 contains maps displaying for each pixel what fraction of the total number of protons which intersected that pixel had WEPLs within the ranges listed in table 3.3. The value of 1 in the scale (bright yellow – "hot" regions) indicates pixels in which all protons from all projections had WEPLs in these ranges. As it can be deduced from figures 3.7 (a–d), areas with high fraction of lower accuracy WEPLs were overlapping with many of the inserts, and are in good qualitative agreement with the artifacts seen in figure 3.4 and 3.6.

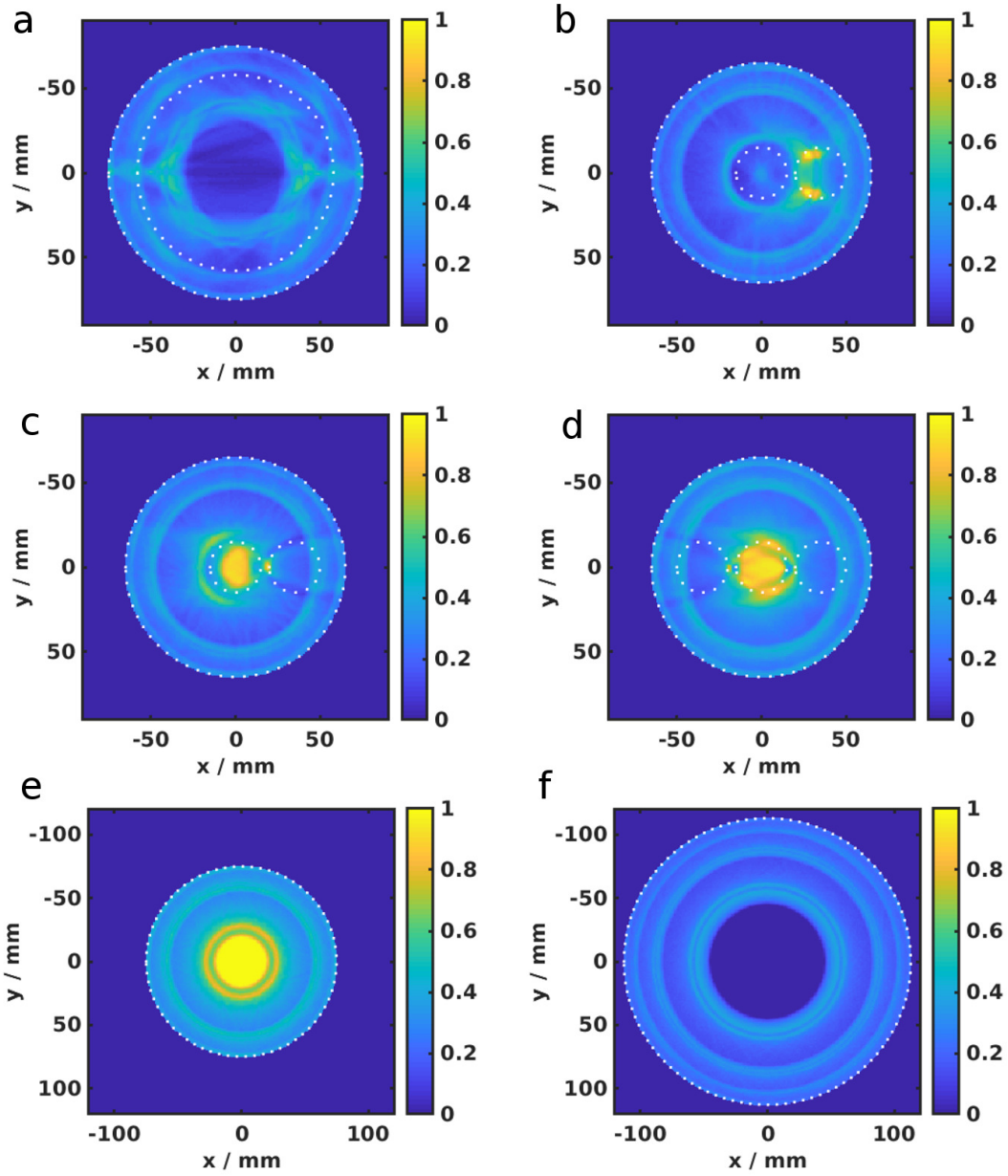

Figure 3.7: Fraction of the number of protons with certain WEPLs (see table 3.3) crossing a pixel for (a) CTP404 phantom, (b) LMU *TopSup*, (c) LMU *TopInf*, (d) LMU *Bot*, (e) water phantom 150.5 mm diameter and (f) water phantom 225.75 mm diameter. The outer dotted circles denote the hull of the cylindrical phantoms. The inner dotted circle in (a) indicates the radius at which the centers of the inserts of the CTP404 phantom are located. For (b–f), the inner dotted circles delineate the inserts of the LMU phantom. The images were produced by thresholding experimental pCT images to the nearest reference RSP values.

3.4 Discussion

3.4.1 Proton CT – Dual energy CT comparison

The RSP MAPE achieved with an optimized calibration of the pCT phase II preclinical prototype scanner was below 1%, at a physical dose of 1.5 mGy to 1.9 mGy, as summarized in table 3.4 and shown as a function of the reference RSP of the inserts in figure 3.3. In specific, for experimental pCT it was 0.55%. This dose to noise relationship is in agreement to Schulte et al. (2005) and Dickmann et al. (2019). The RSP MAPE accuracy from realistic pCT simulations was slightly worse, 0.69%. This is mainly due to the exaggerated tracker gaps in the simulation, resulting in high RSP errors for central inserts. When enforcing a smoother connection of calibration curves between adjacent stages, RSP MAPE deteriorated to 0.87%. At the stage interfaces, different effects such as low numbers of protons with very high energy deposits, sharing their energy to adjacent stages, depositing energy to inactive material and imposed minimum energy threshold can contribute to a higher ambiguity in that region of the calibration. Ignoring the interpolation of the kink region as in *ExpCalib2* led to markedly worse RSP MAPE. The RSP MAPE accuracy achieved with DECT was 0.67%, at an imaging dose of 35.7 mGy. The relatively high DECT imaging dose (about 20 times that of pCT) was used to provide a robust benchmark to compare pCT and resulted in noise in DECT being 4 to 5 times lower than in pCT. However, we did not aim at reducing the DECT imaging dose, and equivalent accuracy can be expected for lower exposures (Landry et al. 2016). Nevertheless, below a certain threshold noise is expected to impact DECT (Lee et al. 2019). Since the DECT reconstruction uses an iterative algorithm to reduce noise, while pCT uses an analytical reconstruction, it is outside the scope of this work to further discuss the noise differences. For both imaging modalities the maximum RSP error exceeded 1%. For pCT the maximum RSP error was 1.31% for the Teflon insert of the CTP404 phantom, while for DECT the highest RSP error was 2.38% also for the Teflon insert of the CTP404 phantom and the second highest was 1.17% for the cortical bone insert of the LMU phantom. Excluding the Teflon insert, the RSP MAPE for pCT and DECT were 0.49% and 0.53%, respectively.

3.4.2 Proton CT artifacts

For experimental and realistically simulated pCT scans, the RSP image contained artifacts whose amplitude in some cases exceeded 2% in RSP. As seen in figure 3.4 (a, d, g, j) and (b, e, h, k), realistic simulations and experimental scans both show artifacts in forms of rings and approximately at the same locations. Nevertheless, in experimental pCT scans the artifacts were expressed mostly as RSP overestimation, contrary to what happened in realistically simulated pCT scans. The presence of similar artifacts in realistic simulations and experiment was the result of the detailed modeling of the scanner geometry and detection effects (Birks' effect), as already manifested in the agreement of the respective calibration curves shown in figure 3.1.

In cylindrical phantoms, ring artifacts usually appear when the error in the WEPL

determination of protons that traverse specific WEPLs is higher than for other protons. For the case of the pCT phase II prototype scanner these can be attributed to the kink and stage interfaces regions. The ring artifacts were mostly expected to occur at specific WEPL value ranges (see table 3.3). Therefore they appeared at different radii in different phantoms, depending on the radial distance from the center at which protons traversed chord lengths corresponding to the previously mentioned WEPL ranges. This was confirmed by the realistic simulations of a water phantom with two different radii shown in figure 3.6. In the image of the 225.75 mm diameter phantom, the same ring artifacts are observed as in the 150.5 mm diameter phantom, but at larger radii.

Qualitative spatial maps of lower WEPL accuracy, using the WEPL ranges listed in table 3.3 were presented in figure 3.7. The image pixels which are sampled by a large fraction of the protons having low accuracy WEPLs were expected to suffer from lower RSP accuracy. This is indeed the case, as the areas with high fraction in each phantom seem to overlap well with the artifacts shown in figure 3.4 and 3.6, confirming our assumptions about the source of the problem. A relevant observation for this study is that several regions of low accuracy WEPLs were located inside the tissue mimicking inserts, thus affecting the achievable RSP accuracy. The strong correlation between the level of that fraction in an insert and the quantified RSP accuracy, can be appreciated by examining the example of the central inserts of the LMU phantom for experimental pCT. In the central insert of the LMU *TopSup* (Liver), the fraction was lower compared to that in the LMU *TopInf* (Bone200), and much lower than that in the LMU *Bot* (Bone400). The reverse trend, as we would expect, was observed in terms of RSP accuracy, which was in experimental pCT 0.04% for Liver, -0.41% for Bone200 and -0.84% for Bone400.

In addition to the above–mentioned artifacts occurring at specific WEPL ranges, other types of artifacts were also observed. The most prominent being a strong lower RSP artifact at the center of the images (dark spot) which was present in all images of realistic pCT simulations and is mostly visible in figures 3.4 (b, h and k). This artifact was caused by the gaps in each tracker layer, described in section 3.2.2. When the simulation models air in the tracker gaps, the position–fixed RSP artifact is visible and always located at the isocenter. When the tracker gaps are filled with the same material as the active areas (silicon), the main position–fixed artifact disappears. Less prominent position–fixed artifacts are still visible in the image. It is possible that some minor effect of the tracker gaps is also present in experimental pCT images, but given the fact that these inactive areas are mostly filled with silicon or glue, this effect should be small. The tracker gaps artifact (dark spot) is also responsible for the fact that in the realistically simulated pCT, where the gaps are exaggerated, the correlation between the fraction of protons with low WEPL accuracy and RSP accuracy was not preserved. In contrast to the experimental pCT results, for realistic simulations all three central inserts of the LMU phantom suffered, as expected, from a significant RSP underestimation. The RSP accuracy was -1.47%, -1.36% and -1.11% for Liver, Bone200 and Bone400. An example of the realistic pCT simulation with the tracker gaps filled is shown results in section 2 of the *Supplementary Material*.

As depicted in figure 3.4 (c, f, i, l), artifacts were not present in pCT images reconstructed from ideal detector simulations, i.e. from ideal proton energies and positions.

Therefore, it was concluded that they were not inherent to the applied FBP reconstruction algorithm. The overall RSP accuracy in ideal pCT simulations was better than 0.2%, on par with past ideal simulation based studies (Hansen et al. 2016), and suggesting that successful artifact mitigation is required to fully exploit the phase II pCT prototype's potential for high accuracy RSP estimation. For three inserts of the LMU phantom (Adipose, Muscle and Bone400) the RSP accuracy achieved with ideal pCT simulations was worse (approximately at 0.4%) than for all other inserts. These inserts, with relatively large differences amongst their reference RSP, were located in the same layer of the LMU phantom and moreover were aligned along one line. Therefore, we hypothesized that, in some projections, nuclear interaction and large angle scattering events might not be efficiently filtered from the data with the current cuts. To confirm this hypothesis, an ideal pCT simulation of that phantom, with the nuclear interaction physics switched off, was performed. The RSP accuracy of that simulation was below 0.1% for all three inserts, showing that for some material and geometry configurations more efficient filtering of nuclear interaction and large angle scattering events might be necessary.

3.5 Conclusion

In this first direct experimental comparison of RSP accuracy between a state–of–the–art DECT scanner and the phase II pCT prototype scanner, we have demonstrated that both modalities can currently achieve an RSP accuracy better than 1%. The pCT phase II prototype scanner yielded slightly better RSP MAPE (0.55%) than the commercial DECT scanner (0.67%). We could demonstrate, using a realistic simulation, that characteristic artifacts cause the ideal pCT RSP accuracy of 0.17% MAPE to be degraded to 0.55%. Mitigating these artifacts is thus critical to further improve pCT RSP accuracy.

Acknowledgements

This work was supported by the Bavaria–California Technology Center (BaCaTeC) project Nr. 28 [2015-2], by the German Research Foundation (DFG) project #388731804 "Fluence modulated proton computed tomography: a new approach for low–dose image guidance in particle therapy", by the DFG's Cluster of Excellence Munich–Centre for Advanced Photonics (MAP) and the DFG's Research Training Group GRK2274 Advanced Medical Physics for Image-Guided Cancer Therapy project B4. We thank Dr. Valentina Giacometti and Dr. Pierluigi Piersimoni for sharing their initial implementation of the pCT simulation platform.

3.6 Supplementary material

3.6.1 DECT

The fits mentioned in subsection 2.1.3 of the main manuscript were obtained from data acquired from the Gammex RMI 467 phantom, scanned at 90 kVp and at 150 kVp. Initially a fit was performed on the known relative electron density ρ_e of the inserts, according to:

$$\rho_e = a\frac{(1+\alpha)\mathrm{HU_H} - \alpha\mathrm{HU_L}}{1000} + b \tag{3.8}$$

where α, a and b are the fit parameters and $\mathrm{HU_H}$ and $\mathrm{HU_L}$ are the Hounsfield units (HU) obtained at 150 kVp and 90 kVp, respectively.

The results of the fit, together with the data points used for the fit (RMI phantom) are shown in figure 3.8. Additionally the data points from the inserts used in the LMU phantom are shown.

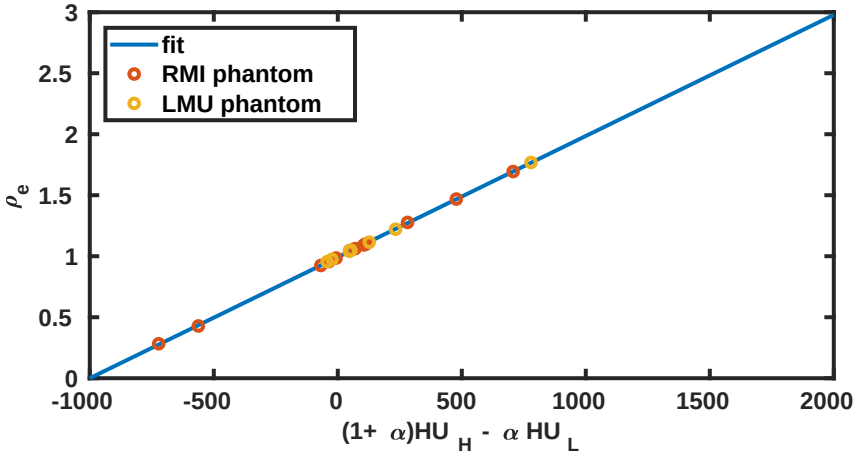

Figure 3.8: Fit on the known ρ_e. The data points of the RMI phantom were used for fitting. The data for the LMU phantom are also plotted but were not used for fitting.

In the next step a fit on the known effective atomic number Z_{eff} of the inserts was performed:

$$\left(\frac{Z_{\mathrm{eff}}}{Z_{\mathrm{eff,w}}}\right)^{3.3} - 1 = \gamma_{\mathrm{L}}\left(\frac{u_{\mathrm{L}}}{\rho_e} - 1\right) + \gamma_0 \tag{3.9}$$

where $Z_{\mathrm{eff,w}} = 7.4774$ is the effective atomic number for water, u_{L} is the scaled HU at 90 kVp according to $\mathrm{HU_L}/1000 + 1$. γ_{L} and γ_0 are fit parameters. The result of the fit is shown in figure 3.9.

Finally a fit was performed on the known ionization potential I according to the equation:

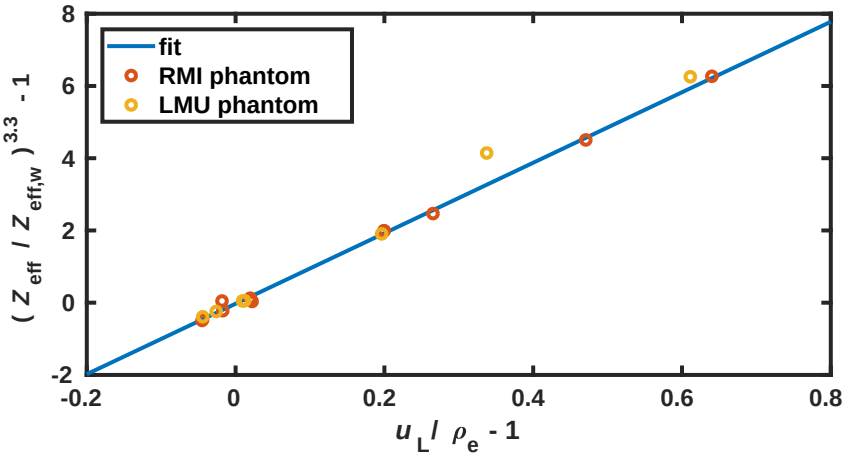

Figure 3.9: Fit on the known Z_{eff}. The data points of the RMI phantom were used for fitting. The data for the LMU phantom are also plotted but were not used for fitting.

$$\ln \frac{I}{I_{\text{w}}} = C_1 \left[\left(\frac{Z_{\text{eff}}}{Z_{\text{eff,w}}} \right)^{3.3} - 1 \right] - C_0 \tag{3.10}$$

where $I_{\text{w}} = 78\,\text{eV}$ is the ionization potential of water and C_0, C_1 the fit parameters. The resulting fit is shown in figure 3.10.

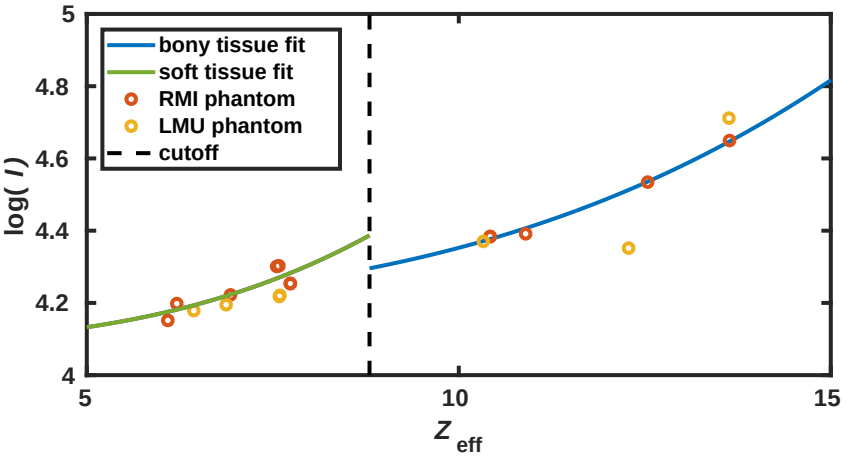

Figure 3.10: Fit on the known I. The data points of the RMI phantom were used for fitting. The data for the LMU phantom are also plotted but were not used for fitting. The vertical dashed line indicates the transition between the soft tissue and bony tissue fits.

3.6.2 pCT

As it was shown in the main part of the manuscript, modelling the tracker gaps material as silicon instead of air removed the central dark spot artifact in the water phantom. The central artifact in the LMU phantom coincided with the central inserts. Therefore a test simulation of that phantom with the tracker gap material modelled as silicon was performed. The result is shown in figure 3.11, where the effect of the tracker gap material (air vs silicon) becomes apparent. The RSP accuracy of the central insert also improved in the case of the tracker gaps being filled with silicon. For the TopSup part of the phantom the central insert (Liver) improved from -1.47% to -0.48%. For the TopInf part of the phantom the central insert (Bone200) improved from -1.36% to -0.84%. Finally, for the Bot part of the phantom the central insert (Bone400) improved from -1.11% to 0.30%.

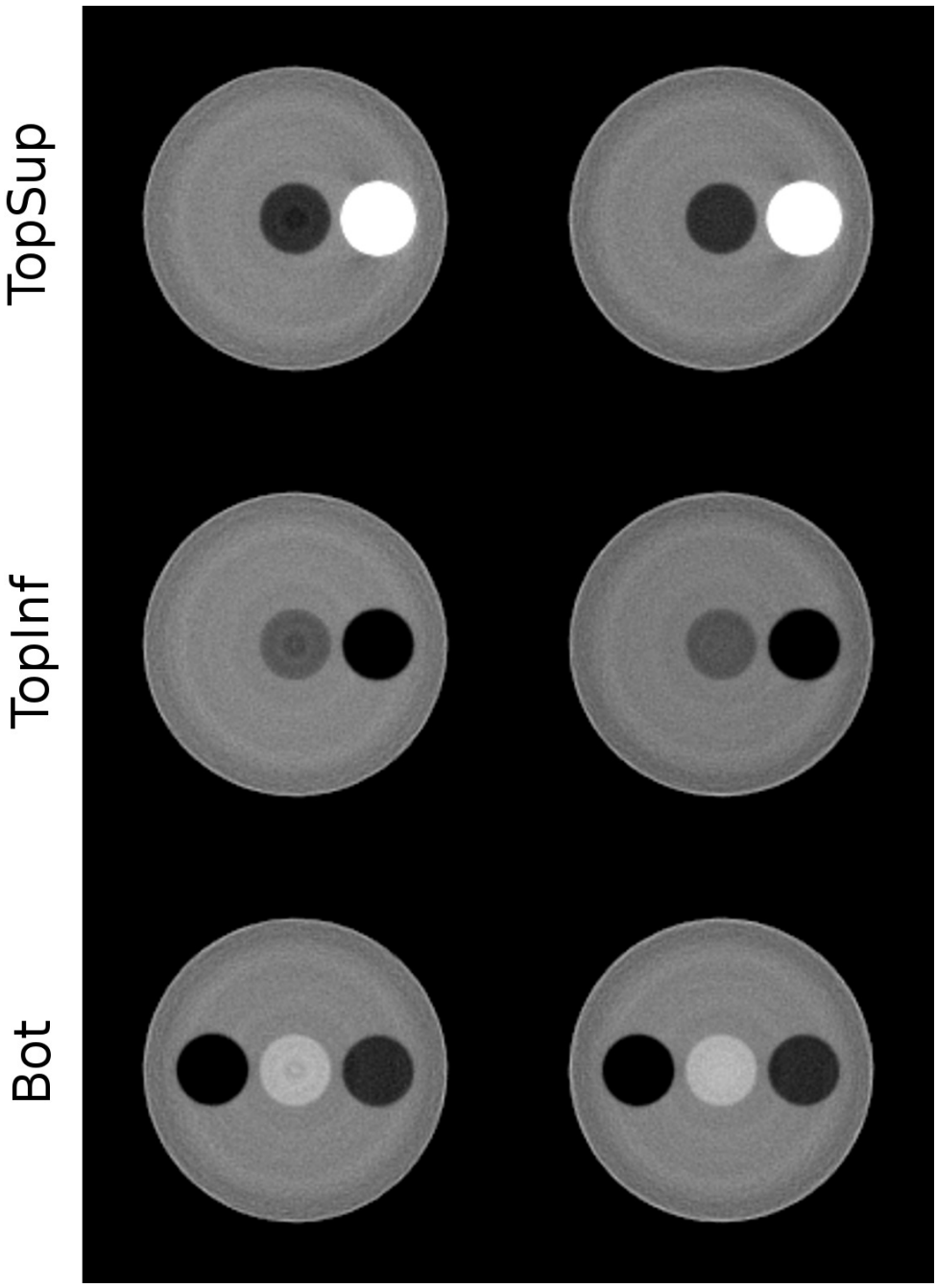

Figure 3.11: Reconstructed pCT images of the LMU phantom. The first column contains images reconstructed from realistically simulated data with the tracker gap material modelled as air. The right column contains images reconstructed from realistically simulated data with the tracker gap material modelled as silicon. An RSP level of 1.15 and window of 0.3 were applied on the images in order to highlight the pCT image artifacts.

Chapter 4

Fluence modulation in proton CT

The investigation presented in this chapter was published in Physics for Medicine and Biology, volume 62, issue 15, pages 6026–6043, in July 2017, with the title *Application of fluence field modulation to proton computed tomography for proton therapy imaging*, by Dedes et al. (2017).

4.1 Introduction

Proton computed tomography (pCT), initially proposed by Cormack (1963), has recently seen considerable research interest as a means of reducing range uncertainties in proton therapy (Paganetti 2012, Yang et al. 2012). By measuring the position and angle of each proton, as suggested by Huesman et al. (1975), along with the energy loss behind the patient, a relative stopping power to water (RSP) map can be directly reconstructed using dedicated reconstruction algorithms (Penfold et al. 2009, Penfold et al. 2010, Rit et al. 2013, Hansen et al. 2014, Hansen et al. 2016). Recent detector developments coupled with improved computing power have permitted the development and operation of two list-mode pCT scanner prototypes based on broad (passively scattered) proton beam delivery (Hurley et al. 2012, Sadrozinski et al. 2016). The RSP accuracy of pCT is expected to improve the current clinical practice, which is based on the calibration of single energy CT scans with a reported uncertainty of 3.5% (Yang et al. 2012). Additionally, pCT has the potential to equal or outperform the accuracy achievable with dual energy CT (Hünemohr et al. 2013, Hünemohr et al. 2013a, Hudobivnik et al. 2016, Möhler et al. 2016), according to the simulation study of Hansen et al. (2015). Initial results from pCT prototypes operating in list–mode suggest at least comparable accuracy to DECT with experimental RSP errors currently below 1.4% for phantoms (Giacometti et al. 2017a).

Besides RSP accuracy, pCT benefits from good dose efficiency, yielding better contrast to noise ratio at equivalent dose levels when compared to x–ray CT (Schulte et al. 2005). Early results from pCT prototypes report doses of about 1 mGy to achieve acceptable image quality (Sadrozinski et al. 2016). This is comparable to the imaging dose used by in–room

cone–beam CT (CBCT) imaging, and lower than typical x–ray CT treatment planning imaging doses (at least 10 mGy and ranging up to 100 mGy, see Table 7 in (Murphy et al. 2007)). A recent DECT publication specific to proton therapy reported doses of 20 mGy (Hudobivnik et al. 2016). This aspect of pCT suggests the technique could be used for daily low–dose in–room imaging and could compete with x–ray CBCT image quality without the need of corrections, as needed for proton dose calculations on CBCT images (Kurz et al. 2015, Landry et al. 2015a, Landry et al. 2015b, Park et al. 2015, Veiga et al. 2015, Veiga et al. 2016). In x–ray CT, several dose reduction techniques have been developed such as the use of bow–tie filters (Graham et al. 2007a, Mail et al. 2009) and automatic exposure control (Mulkens et al. 2005, McCollough et al. 2006, Kalender et al. 2008). An exciting idea consists of prescribing image quality levels prior to CT scanning using modulation of the x–ray fluence within the fan–beam (Graham et al. 2007b, Bartolac et al. 2011, Bartolac & Jaffray 2013, Szczykutowicz & Mistretta 2013a, Szczykutowicz & Mistretta 2013b). This approach is called fluence field modulated CT (FFMCT) and was recently experimentally realized using the imaging system of a TomoTherapy machine (Szczykutowicz & Mistretta 2014, Szczykutowicz et al. 2015). The FFMCT concept is very similar to intensity modulated radiation therapy (IMRT) where the mega–voltage photon fluence is modulated with a multi–leaf collimator. The main difficulty with modulating the x–ray fluence of a conventional CT scanner is the lack of an equivalent modulation device. For this reason experimental realization of FFMCT made use of the TomoTherapy machine's imaging system which is equipped with a 64 leaf binary collimator. Another group has achieved FFMCT by using multiple aperture devices (Stayman et al. 2016).

Given the current rise in the number of proton therapy centers equipped with pencil beam scanning (PBS), one can imagine imitating intensity modulated proton therapy (IMPT) instead of IMRT, to create a pencil beam (PB) pCT system allowing fluence modulated pCT (FMpCT). FMpCT would rely on beam current or delivery time modulation during the acquisition of a pCT projection to achieve the desired proton fluence modulation. Using a treatment planning image as guide, an in–room FMpCT scan could thus be acquired with high image quality in the beam path and a lower image quality in regions receiving negligible therapeutic dose levels, leading to a lower integral imaging dose compared to uniform fluence scans. Such an image could be used for patient positioning, dose recalculation or even re–planning.

The objective of this study was to evaluate the feasibility of FMpCT using a PB pCT Monte Carlo simulation framework and a state–of–the–art pCT reconstruction algorithm. For a simple homogeneous virtual phantom and two patient CT–based virtual phantoms, image quality at different modulation levels was investigated, as well as proton dose calculation accuracy. Furthermore, FMpCT was emulated using a selected subset of experimental data acquired with a state–of–the–art scanner prototype for broad beam proton irradiation.

4.2 Materials and methods

Proton CT simulation

The simulation platform used in this study is based on GEANT4 version 10.01.p02 (Agostinelli et al. 2003) and has been used in previous studies for proton dose calculation in voxelized geometries (Landry et al. 2015b, Schmid et al. 2015, Hudobivnik et al. 2016). The reference physics list `QGSP_BIC_HP` was used for the simulation of interaction of particles with matter. The simulation platform uses CT images which are converted to mass density and tissue composition using the approach of Schneider et al. (2000).

Existing list–mode pCT scanner prototypes rely on two tracker modules located up– and down–stream of the scanned object. The tracker modules are made of pairs of two orthogonal silicon strip detectors allowing position and direction detection. A calorimeter, located after the second tracker module, records the residual energy loss. A more detailed description can be found in (Sadrozinski et al. 2016) (see figures 1 and 2 in that publication). In this study, an ideal pCT scanner in the form of two scoring planes, before and after the scanned object (which is centered at the origin) was simulated. The two ideal scoring planes, which are of the same material as the surrounding world (air), record the exact position, direction and energy of each traversing proton. The planes are positioned perpendicularly to the X axis, which is parallel to the beam, at $-15.88\,\text{cm}$ and $15.88\,\text{cm}$, both covering a surface of $60 \times 60\,\text{cm}^2$. This area was chosen for simplicity and does not represent the field of view of a real scanner. The simulated imaging beam consists of a 2D grid of non–divergent proton PBs arranged at 1 cm intervals in the YZ plane. Each PB has a two–dimensional Gaussian proton distribution with $\sigma = 4\,\text{mm}$ in air. The Z extent (parallel to the superior inferior patient axis) of the PB grid was adjusted according to each virtual phantom, and the Y extent was set to 25 cm. In order to simulate a tomographic scan, the scanned phantom was rotated around the Z axis at 1^o steps, covering an angular span of 360^o.

4.2.1 Virtual phantoms

Three voxelized phantoms were used in this study. The phantoms were derived from x–ray CT scans of IMRT patients.

1. A simple phantom was simulated by overwriting a $1.074 \times 1.074 \times 1\,\text{mm}^3$ voxel size CT image with a 10 cm radius cylinder with 0 HU. Outside the cylinder –1000 HU was used. Note that when using the stoichiometric calibration, 0 HU corresponds to a predefined human tissue composition and not water. For this case a single row of PBs bisecting the cylinder was employed given the Z axis symmetry.

2. The second phantom was derived from a $1.074 \times 1.074 \times 1\,\text{mm}^3$ voxel size CT scan of a patient (Pat1) treated with IMRT for a brain metastasis with a small planning target volume (PTV) located near the base of the skull. The PTV was 5 cm along the Z axis.

3. For the third phantom, a $1.074 \times 1.074 \times 3\,\mathrm{mm}^3$ voxel size CT scan of a paranasal sinus cancer patient (Pat2) was used. The large PTV including lymph nodes was 14.3 cm along the Z axis. For (2) and (3), the PB grid extent in Z was set to cover the PTV plus a margin (see section 4.2.4).

For each phantom, a 360 projections pCT scan with 104 protons per PB per projection was simulated. The corresponding fluence at the first tracker was 9600 protons per cm^2 in a projection. For all simulations the dose to tissue per voxel was scored.

4.2.2 Fluence modulation

The concept of fluence modulation based on proton PB scans presented in this work relies on the definition of regions of interest (ROIs), in which a high image quality is desirable. The term image quality here refers to RSP noise levels and RSP accuracy. The ROIs should ideally cover the beam path including the PTV and could be derived using diagnostic or treatment planning imaging data. Phantom specific ROI generation will be presented below.

A schematic representation of the concept is shown in figure 4.1 for a simplified ROI. In this proof of principle study, the PB modulation pattern was obtained by calculating a binary sinogram (PB index vs. projection angle). The sinogram entries were 1 if the central axis of a PB intersects the ROI in a given projection and 0 otherwise. Using this sinogram, the fluence of PBs assigned 0 is reduced by a given fluence modulation factor (FMF), FMF<1, while those PBs assigned 1 preserve full fluence (FF). The modulation was performed as a post–processing step to allow several FMpCT images from a single simulation. When reducing a PB's fluence, list–mode data were randomly discarded, ensuring that the energy and spatial distribution of the PBs was preserved.

4.2.3 Proton CT reconstruction framework

The reconstruction algorithm chosen for this study was a filtered backprojection (FBP) implementation which accounts for the curved proton paths in the imaged object, mainly due to multiple Coulomb scattering. The main principles of the algorithm are presented in (Rit et al. 2013) and a comparison with different iterative algorithms is presented in (Hansen et al. 2016). The algorithm is based on list–mode data, and the actual path of every proton is approximated by a most likely path (MLP) formulation (Schulte et al. 2008), which uses the position and direction information from the scoring planes.

Protons were selected with a 3 standard deviations cut on the energy and angular distributions around their mean energy and angle per projection pixel, in order to filter out nuclear reactions and large angle scattering events, whose energy loss and path cannot be described by the Bethe formula and the MLP formalism. List–mode data were binned in intermediate projection images with $1 \times 1\,\mathrm{mm}^2$ pixels, which were then filtered and back–projected. Scans simulated with low fluences can suffer from artifacts due to the absence of proton information in some pixels of the intermediate projections. To counter

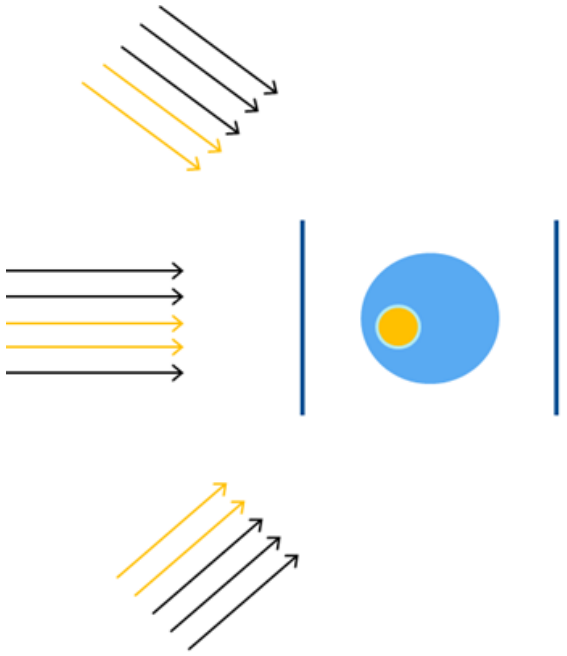

Figure 4.1: Schematic representation of FMpCT using PBS. PBs whose central axis geometrically intersects the ROI preserve full fluence while the others have their fluence reduced. The vertical lines represent the tracker planes for the horizontal PBs. The energy measuring detector is omitted for simplicity.

this, sinogram interpolation (Hansen et al. 2016) was used for all reconstructions. Images were reconstructed on the same grids as the CT scans used to generate the GEANT4 virtual phantoms. For each phantom and each FMF, the following images were reconstructed: (1) a FF image, (2) a FMF·FF uniform fluence image and (3) a FMpCT image with a fluence of FMF·FF outside the ROI.

4.2.4 FMpCT ROIs and FMF

For the cylindrical phantom, simple circular ROIs with 1 cm radius were studied, as shown in figure 4.2. The FMF was set to 0.1, 0.05 and 0.01. ROI1 was at the center of the cylinder, ROI2 was 37.6 mm off–center and ROI3 was 75.2 mm off–center.

For Pat1 and Pat2, proton treatment plans using PBS were generated using a research version of a commercial TPS (Raystation, Raysearch Laboratories, Sweden). For Pat1, a single field uniform dose (SFUD) plan using a 220° gantry angle on the International Electrotechnical Commission (IEC) scale was used to deliver 60 Gy to the PTV (15 cm^3). For Pat2, a 3–field IMPT simultaneous integrated boost plan with beams at 0°, 100° and 260° was used. The 0° field was used only superiorly to the nasal cavities. The high dose PTV (174 cm^3) received 60 Gy and the lymph node PTV (510 cm^3 − 174 cm^3 = 336 cm^3) 50 Gy. For both Pat1 and Pat2, a FMpCT ROI was obtained by using the 10 Gy isodose

line, ensuring inclusion of beam paths, PTVs as well as relevant organs at risk. The ROI volumes were 220 cm^3 for Pat1 and 2021 cm^3 for Pat2, and the ROIs are shown in figure 4.3. For Pat1 and Pat2 FMF was set to 0.1, 0.05 and 0.01.

Figure 4.2: Homogeneous 10 cm radius phantom with three different 1 cm radius ROIs (from top to bottom). The two colors indicate high and low fluence. For each ROI, a uniform FF scan (left), a uniform scan of lower fluence FMF·FF (right), and finally three FMpCT scans of FF inside the ROI and FMF·FF outside (middle) were simulated.

4.2.5 Image quality

The image quality was quantified according to RSP accuracy and noise. For each virtual phantom, a reference RSP distribution was obtained from GEANT4 and accuracy was calculated as the mean of $(RSP–RSP_{ref})/RSP_{ref}$ in the FMpCT ROI. Noise was the standard deviation of $(RSP–RSP_{ref})/RSP_{ref}$ in the same ROI. In this study, for both noise and accuracy, relative differences of less than 0.5% (absolute value) were considered negligible. Additionally, the imaging dose reduction in the FMpCT images with respect to the FF images, as a function of FMF, was also quantified.

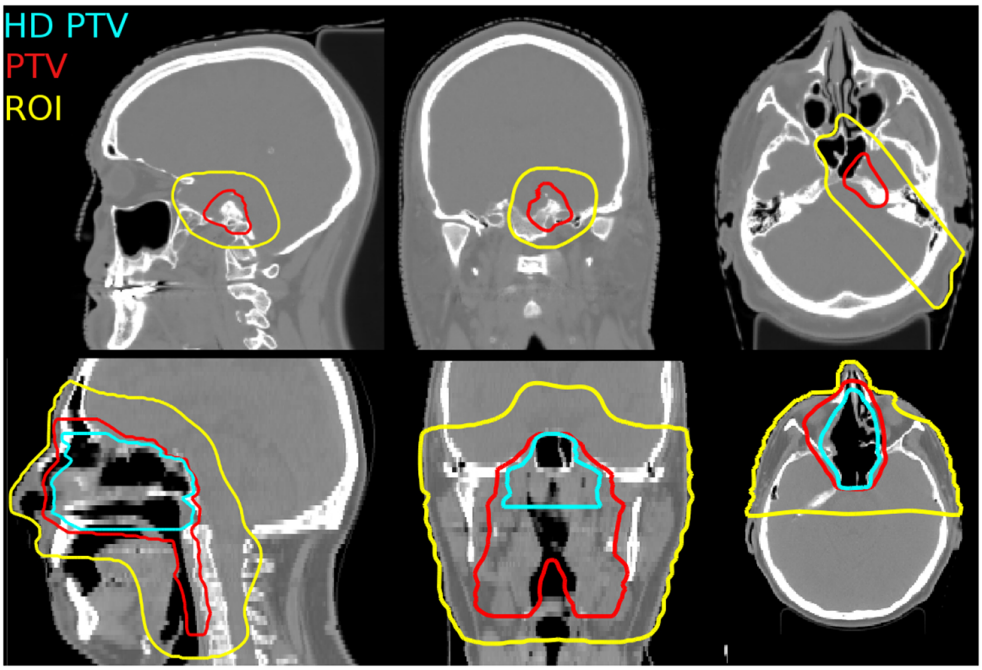

Figure 4.3: (top row) Axial, coronal and sagittal slices of the CT image of Pat1, including the FMpCT ROI and PTV. (bottom row) The same for Pat2, with an additional high dose PTV (HD PTV).

4.2.6 Dose recalculation

To evaluate dose calculation accuracy on FMpCT images, the reference 3D RSP_{ref} distributions were imported in the TPS with a custom lookup table and used to re–optimize the treatment plans of Pat1 and Pat2. Subsequently, the FF, FMF·FF and FMpCT images were also imported in the TPS and used for plan recalculation. For the SFUD beam of Pat1, beam eye view range (80% of dose maximum) differences to the RSP_{ref} image were computed for dose profiles showing maximum doses higher than 50% of the prescription dose. The percentage of profiles within 1 mm and 2 mm of the RSP_{ref} dose distribution was computed. Additionally, dose volume histograms (DVH) and 2%/2 mm gamma pass rates were obtained for the SFUD dose distribution of Pat1 and the IMPT dose distribution of Pat2. For gamma evaluation doses above 50% of the prescription dose were considered.

4.2.7 Experimental data

In addition to the simulated data, FMpCT was also attempted on experimental data obtained with the phase II preclinical prototype pCT scanner (Sadrozinski et al. 2016) of the Loma Linda University and U.C. Santa Cruz at which a pediatric head phantom (715-HN, CIRS, Norfolk, VA) was scanned. The 90 projections scan with $2.5 \cdot 10^6$ protons per projection was performed at the Northwestern Medicine Chicago Proton Center (NMCPC) facility, using a 200 MeV proton wobbled beam from the IBA universal nozzle for another

study (Johnson et al. 2016). The registered proton fluence per projection in the experimental data was $1.2 \cdot 10^4$ cm^{-2}, similar to that from the simulated data. Although the actual scan was not acquired with fluence modulated PBs, fluence modulation was emulated during post–processing of the list–mode data. For every proton, a straight line path was constructed from the entrance and exit coordinates provided by the tracking system of the prototype. The protons whose straight paths did not intersect the defined ROI were removed from the data with a removal probability of 1–FMF. The goal here was to assess the impact of FMpCT on image quality compared to the FF image; the general performance of pCT for this combination of scanner and phantom has been reported elsewhere (Giacometti et al. 2017a). Thus, for this part of the study, the theoretical RSP (from Giacometti et al. (2017a)) in the homogeneous brain section of the phantom was used as reference for numerical analysis. Since no voxelized ground truth was available, the FF pCT image was used as reference in RSP voxel–by–voxel accuracy images. Given the lower number of protons in the experimental data (90 projections instead of 360), higher values of FMF = 0.5, 0.3 and 0.1 were employed. A single cylindrical ROI of 1.5 cm radius was considered (see figure 4.10 in the results section). The ROI was chosen to contain homogeneous phantom brain material for image quality analysis, given the absence of a voxelized reference at the time of writing.

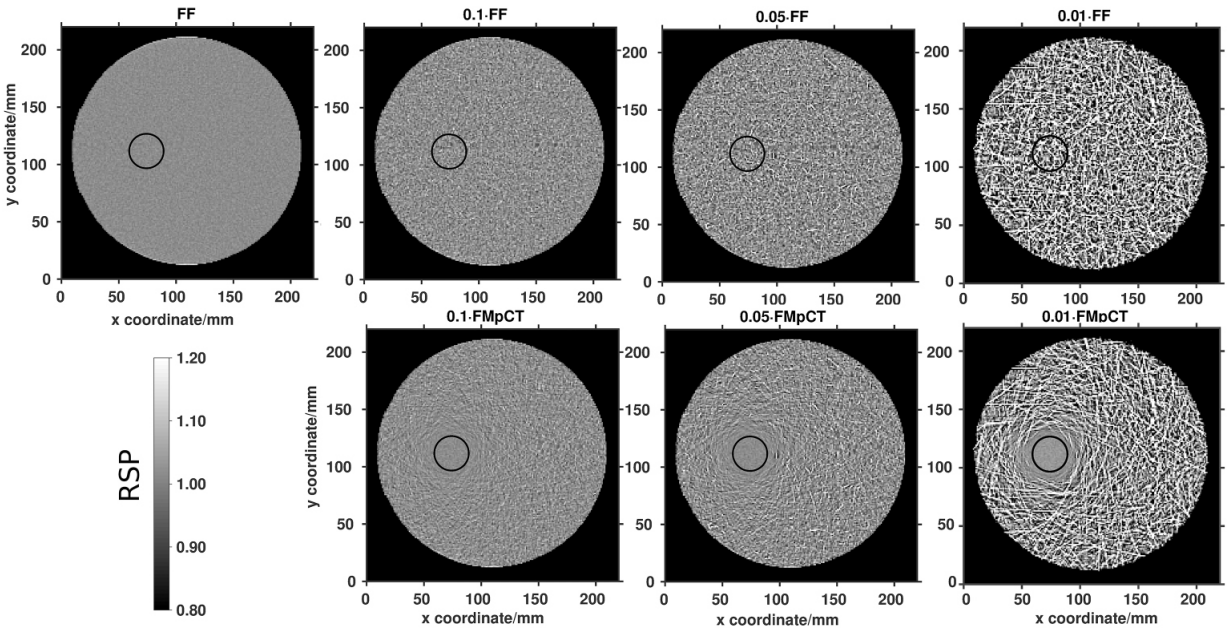

Figure 4.4: Upper row, left to right: FF image (104 protons per PB) and images obtained with homogeneous fluence of 0.1·FF, 0.05·FF and 0.01·FF. Lower row, FMpCT images corresponding to FF in each ROI (black circle) and 0.1·FF, 0.05·FF and 0.01·FF outside. The grayscale corresponds to RSP values.

4.3 Results

4.3.1 Uniform phantom

In figure 4.4, the reconstructed pCT images of the cylindrical phantom are shown. With increasing FMF, the noise in the images increases as expected. In the case of the FMpCT images, the noise and accuracy in the ROIs, indicated by the black circles, showed negligible difference from those of the FF image for FMF = 0.1 and 0.05. At FMF = 0.01, the mean value in the ROI deviated by –1.3% from the reference image (see table 4.4 in supplementary material). The imaging integral physical dose was reduced from 2.3 mGy in the FF image to 0.7 mGy in the FMF = 0.1 FMpCT image (see figure 4.12 in supplementary material). Additionally, the results of the RSP noise and accuracy analysis did not vary as a function of the ROI location (see figure 4.13 in supplementary material).

4.3.2 SFUD and IMPT cases

Figure 4.5 shows the uniform FF as well as FMpCT with FMF = 0.05 and 0.01 for Pat1, where we observed better image quality in the FMpCT ROI used to cover the SFUD beam path than outside. For FMF = 0.05, the RSP/RSP$_{\text{ref}}$ in the FMpCT image is similar in the ROI as the FF image, however for FMF = 0.01 slightly lower RSP values were observed. Figure 4.6 presents noise and accuracy as a function of the FMF for Pat1, where we observed that for FMF < 0.025 the noise increases more than expected from the statistical dependence on (number of protons)$^{-1/2}$ and that the accuracy begins to degrade even with FMpCT.

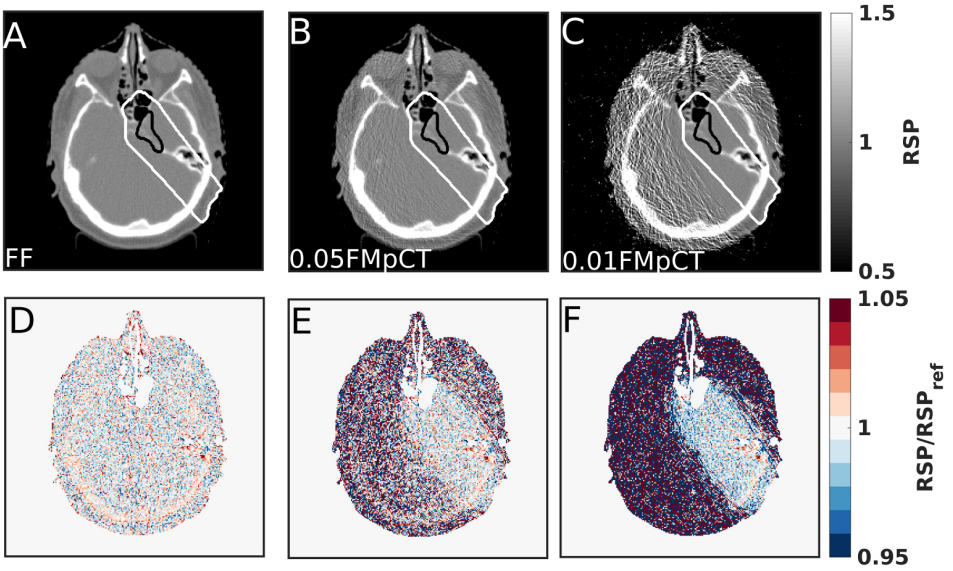

Figure 4.5: (A) FF pCT image for Pat1. FMpCT images with FMF of (B) 0.05 and (C) 0.01. (A)–(C) The PTV (black) and FMpCT ROI (white) are overlaid on the pCT images. (D)–(F) Corresponding relative RSP images.

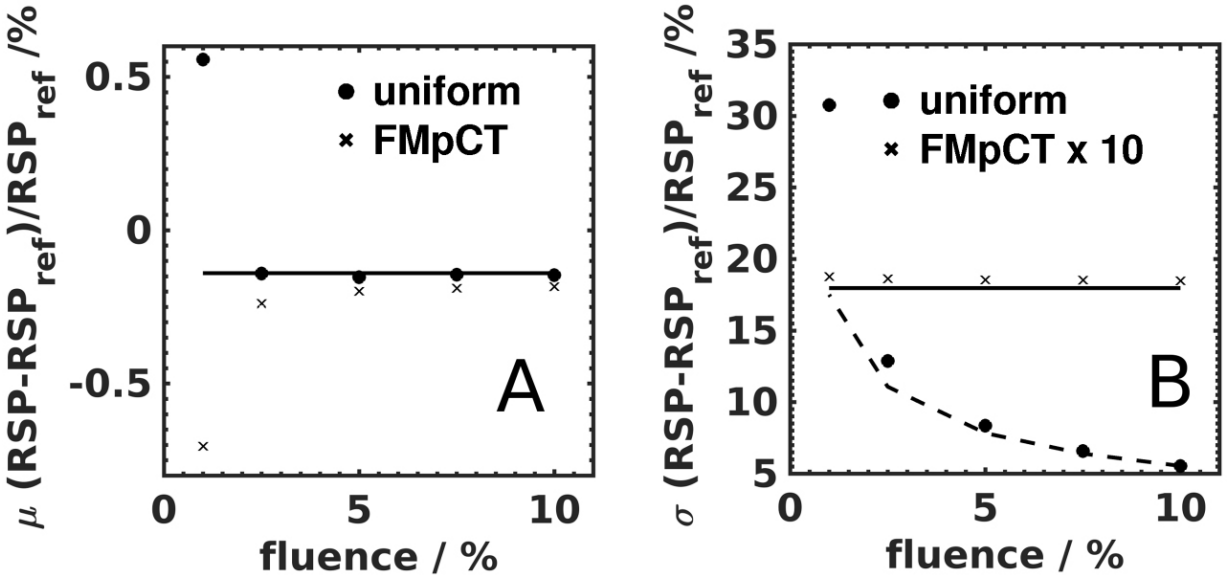

Figure 4.6: (A) Accuracy (mean, μ) of uniform and FMpCT images for various FMF (expressed as percentage of FF) in the ROI of Pat1. The horizontal solid line corresponds to the FF accuracy. (B) Noise (1 standard deviation, σ) of the same distributions. The FMpCT data are multiplied by 10. The horizontal solid line corresponds to the FF noise×10 and the dashed line to $\sigma_{10\%}(\text{fluence}\%/10\%)^{-1/2}$.

Figure 4.7 shows the pCT images for Pat2, where similar image quality preservation as Pat1 was obtained from FMpCT. This can be appreciated in table 4.1 where the results of the RSP noise and accuracy in the ROIs are tabulated for both patients. The noise levels for Pat1 and Pat2 are comparable to those from the uniform cylinder (see table 4.4 in supplementary material) at FMF = 0.1 and FMF = 0.05 . At FMF = 0.01 , the patient images exhibited higher noise. For both Pat1 and Pat2 the FMF = 0.01 caused a slightly increased mean error of 0.6% and 0.5%, respectively, for the uniform fluence, and –0.7% and –0.4%, respectively, for the FMpCT (visible in figures 4.5(F) and 4.7(F)), which is comparable to what was observed with the cylindrical phantom.

In figure 4.8, the imaging dose distributions in the case of uniform FF and FMpCT illustrate the imaging dose reduction achieved for both the SFUD and IMPT cases considered in this study. Table 4.2 presents the integral doses in the whole image as well as outside the ROIs for both patients. For Pat1 with a small PTV, the integral dose reduction was up to 49% for the whole image and 56% outside the ROI with FMF = 0.01. For Pat2, smaller integral dose savings were realized due to the larger PTV and higher number of beams with the same FMF, with a reduction of up to 22% for the whole image and 37% outside the ROI. Interestingly, changing FMF from 0.1 to 0.01 had an effect of less than 5% on integral dose reduction for both cases.

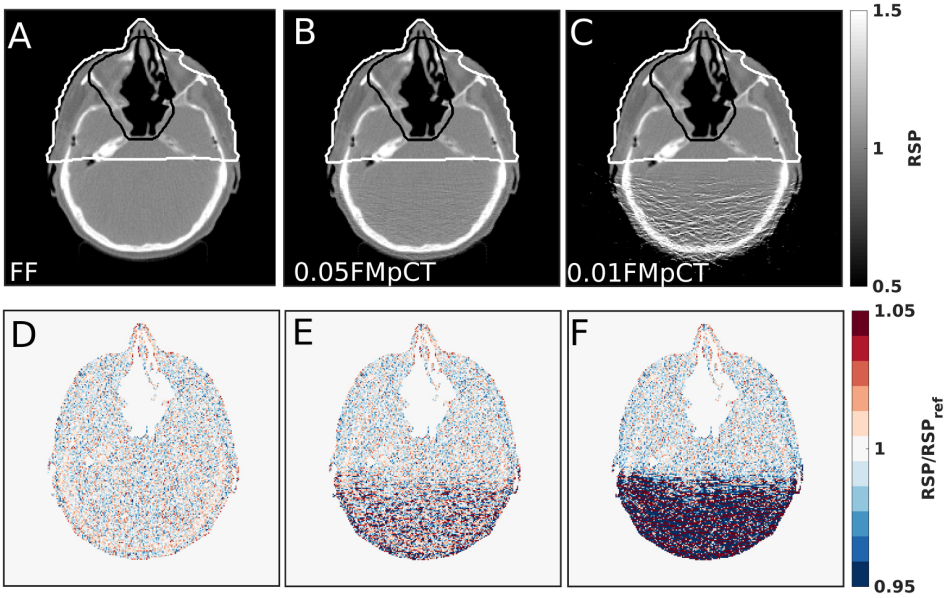

Figure 4.7: (A) FF pCT image for Pat2. FMpCT images with FMF of (B) 0.05 and (C) 0.01. (A)–(C) The lower dose PTV (black) and FMpCT ROI (white) are overlaid on the pCT images. (D)–(E) Corresponding relative RSP images.

Table 4.1: RSP noise (1 standard deviation) and accuracy (mean) results in ROIs for the SFUD (Pat1) and IMPT (Pat2) cases. The second column shows RSP noise values for uniform scans with different FMF. The third column reports noise levels with FMpCT. The fourth and fifth columns refer to the RSP accuracy in the ROIs for the same imaging configuration.

	$(RSP–RSP_{ref})/RSP_{ref}$ (%) Noise		$(RSP–RSP_{ref})/RSP_{ref}$ (%) Mean	
Pat1	Uniform	FMpCT	Uniform	FMpCT
FF	1.8	–	-0.1	–
$0.1 \cdot FF$	5.5	1.8	-0.1	-0.2
$0.05 \cdot FF$	8.3	1.8	-0.2	-0.2
$0.01 \cdot FF$	30.1	1.9	0.6	-0.7
Pat2				
FF	1.4	–	-0.2	–
$0.1 \cdot FF$	4.2	1.5	-0.2	-0.2
$0.05 \cdot FF$	6.3	1.6	-0.2	-0.2
$0.01 \cdot FF$	32.5	1.6	0.5	-0.4

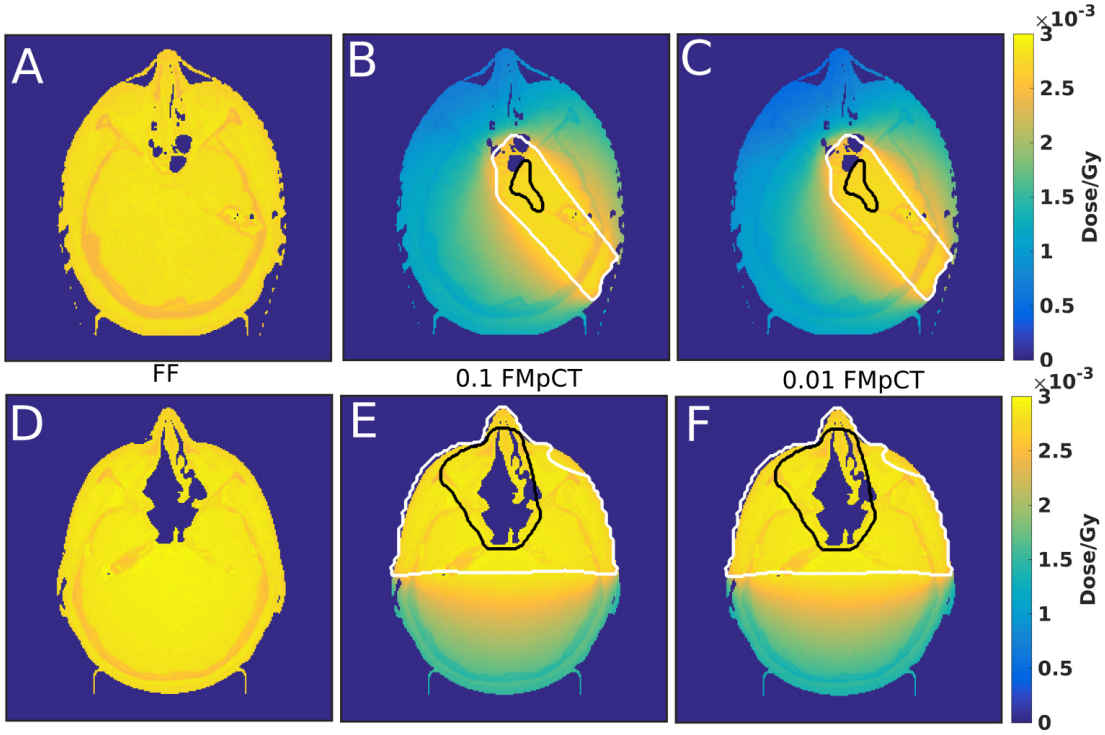

Figure 4.8: (top row) Pat1 imaging dose distributions on the (A) FF image, (B) 0.1 FMpCT and (C) the 0.01 FMpCT image. (bottom row) (D)–(F) Same for Pat2.

Table 4.2: Imaging dose in the complete volume and outside the ROI for both patients.

Pat1	Integral dose (mGy)		Integral dose reduction $1 - \mathrm{FMF} \cdot \mathrm{FMpCT/FF}$	
	Whole image	Outside ROI	Whole image	Outside ROI
FF	2.57	2.56	–	–
$0.1 \cdot$ FF	1.42	1.25	0.45	0.51
$0.05 \cdot$ FF	1.35	1.18	0.47	0.54
$0.01 \cdot$ FF	1.30	1.12	0.49	0.56
Pat2				
FF	2.67	2.65	–	–
$0.1 \cdot$ FF	2.13	1.76	0.20	0.34
$0.05 \cdot$ FF	2.10	1.71	0.21	0.35
$0.01 \cdot$ FF	2.08	1.67	0.22	0.37

The dose calculation accuracy for the SFUD and IMPT treatment plans is presented in figure 4.9 for FMF = 0.01. We observed that the dose distributions calculated on the FMpCT images agreed with those calculated on the reference RSP_{ref} image in terms of isodose levels and DVH curves. This was not the case for the uniform FMF = 0.01 image where dose calculation accuracy was degraded as shown on the DVH of figure 4.9. This was confirmed by the (2%, 2 mm) gamma index analysis presented in table 4.3. For FMF = 0.05, the DVH curves of uniform fluence and FMpCT were identical to the reference.

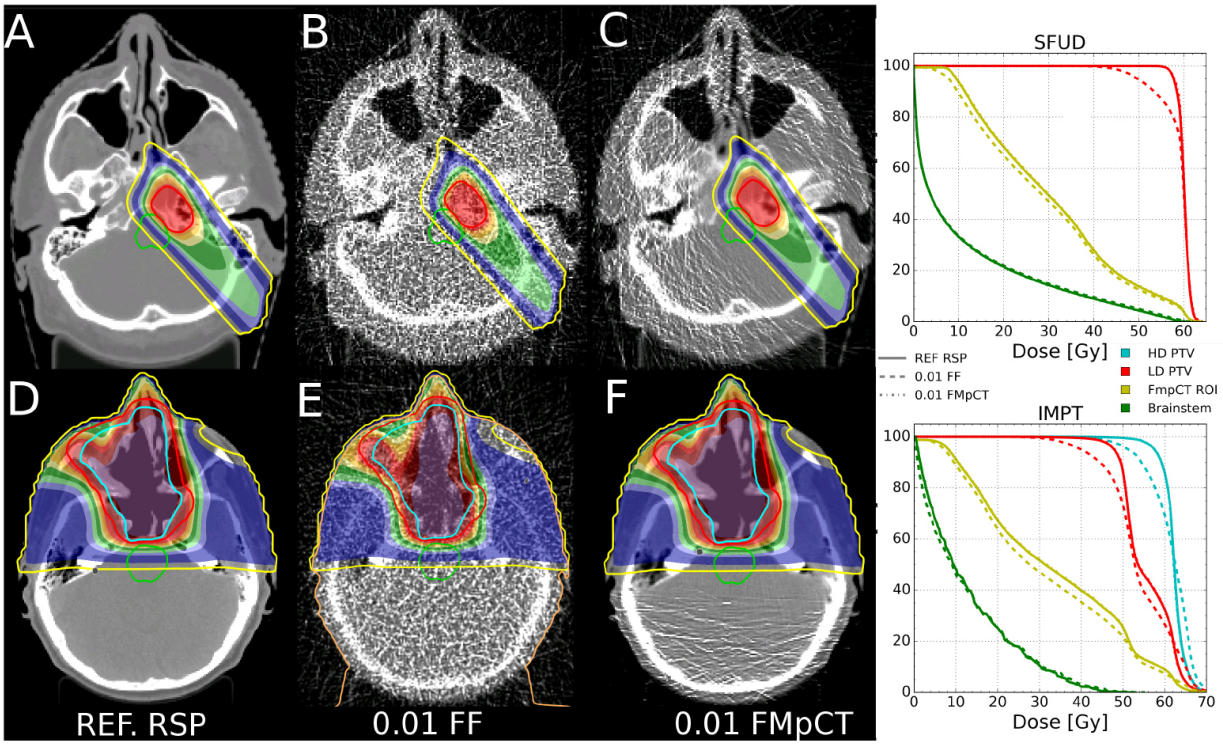

Figure 4.9: (top row) Pat 1 SFUD dose calculation on the (A) RSP_{ref} image, (B) 0.01·FF image and (C) the 0.01·FMpCT image. (bottom row) Pat 2 IMPT dose calculation on the (D) RSP_{ref} image, (E) 0.01·FF image and (F) the 0.01·FMpCT image. The corresponding DVHs are shown on the right.

For the SFUD beam of Pat1, 96% of BEV dose profiles had range differences lower than 2 mm versus RSP_{ref} for FMF = 0.01 FMpCT, while with the uniform FMF = 0.01 pCT only 1% of profiles passed this test. The FMF = 0.05 FMpCT results compared to RSP_{ref} show limited range agreement improvement compared to reduced uniform fluence (90% versus 87% respectively). When comparing the FMpCT and reduced uniform fluence to the FF case a clear improvement was seen (99% versus 88% respectively). This result indicates that at this modulation level (FMF = 0.05), FMpCT retains better agreement with the FF image than the uniform FMF = 0.05 image. The worst pass rate for FMpCT was 66% of profiles with range differences of less than 1 mm when comparing FMF = 0.01 and RSP_{ref}. However when comparing the same FMpCT image to the FF image, the

Table 4.3: SFUD BEV range difference for Pat1. The percentage of profiles with range differences (RDs) below 1 mm and 2 mm are shown. RD are reported using the dose distribution calculated on the RSP_{ref} image and the FF image as reference. For Pat1 and Pat2, the percentage of voxels passing the (2%, 2 mm) gamma evaluation is also reported.

	Pat1				Pat1		Pat2	
	RD < 1 mm(%)		RD < 2 mm(%)		Gamma (2%/2 mm)			
	Uniform	FMpCT	Uniform	FMpCT	Uniform	FMpCT	Uniform	FMpCT
FF - RSP_{ref}	93	–	99	–	99	–	99	–
0.05·FF - RSP_{ref}	87	90	99	99	99	99	99	99
0.01·FF - RSP_{ref}	0	66	1	96	90	98	59	99
0.05·FF - FF	88	99	99	99	–	–	–	–
0.01·FF - FF	0	97	0	99	–	–	–	–

pass rate rose to 97%, indicating that the lower pass rate was caused by a combination of sub-mm FMpCT errors compared to FF, and sub–mm systematic FF errors compared to RSP_{ref}. The range analysis was found to be sensitive to the air cavity abutting the PTV (see figure 4.5(A)) which caused small range shifts in tissue to be expanded in air. We verified that the 93% pass rate of FF pCT versus RSP_{ref} was caused by beam profiles ending in the air cavity.

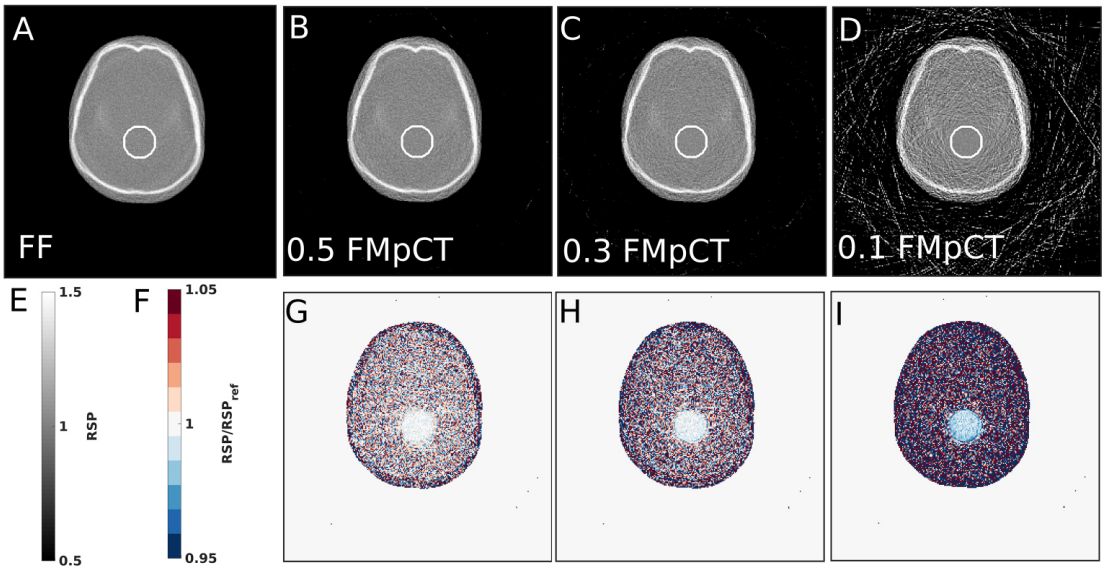

Figure 4.10: (A) FF pCT and (B) 0.5 (C) 0.3 and (D) 0.1 FMpCT images for the pediatric head phantom and (E) their colorbar. (A)–(D) The FMpCT ROI (white) used for modulation and analysis is overlaid on the pCT images. (G)–(I) Corresponding relative RSP images and (F) their colorbar. In this case RSP_{ref} is the FF image.

The results of applying FMpCT to broad beam pCT scans of the pediatric head phantom are shown in figure 4.10 for the three FMF investigated. We observed a similar trend

of reduced accuracy at lower FMF, however this happened at higher FMF for the experimental data. The noise and accuracy are reported in figure 4.11, where we observed that FMpCT with FMF below 0.3 showed accuracy degradation larger than 1% as well as increasing noise in the ROI. In Giacometti et al. (2017a), the theoretical RSP for the brain material of the pediatric phantom is reported as 1.047; the mean value in the ROI of the FF pCT image was 1.04 ± 0.03, i.e. –0.7% lower than the reference. Please note that the theoretical RSP may differ from that measured in a proton beam due to uncertainties on the material composition.

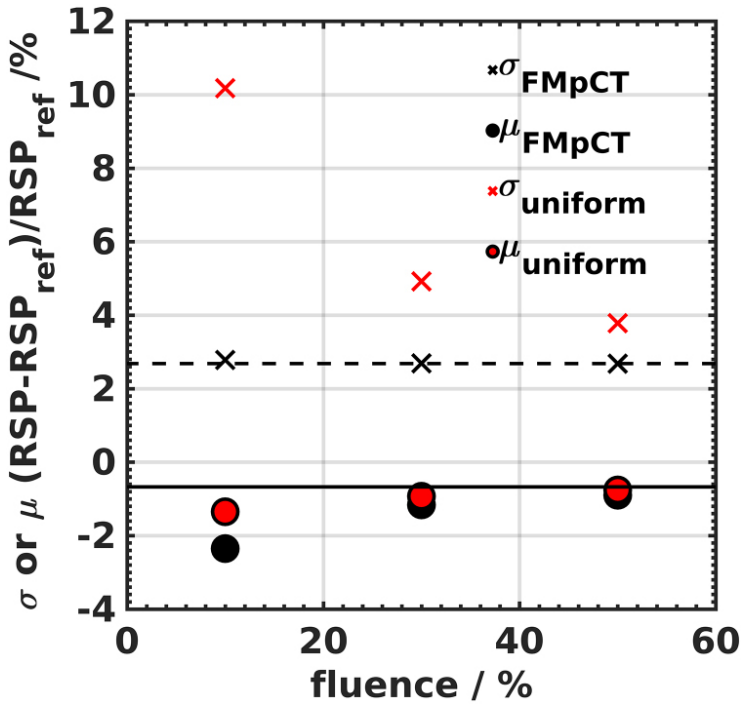

Figure 4.11: Accuracy (mean, μ) and noise (1 standard deviation,σ) of FMpCT images (black data points) and uniform images (red data points) for various FMF (expressed as percentage of FF) in the ROI of the pediatric head phantom. The horizontal dashed (solid) line corresponds to the FF noise (accuracy).

4.4 Discussion

The results shown in figure 4.4 qualitatively support our hypothesis that FMpCT image reconstruction is feasible when employing the binary fluence modulation scheme presented in figure 4.1. For the three levels of fluence reduction outside ROI2, two distinct image quality levels were seen in each image. The image quality metrics reported in table 4.4 of the supplementary material show that in the case of homogeneous fluence, the noise was approximately inversely proportional to the square root of the proton fluence, while FMpCT reconstruction preserved image quality (both standard deviation and mean

value) down to 5% of the FF. However for FMF = 0.01, we observed a break from the (number of protons)$^{-1/2}$ dependence of the noise as well as a –1.3% shift of the mean values in the ROI, indicating a limit to the modulation level feasible with FMpCT. A similar effect was observed for the patient and experimental data. Our initial investigations on the root of this error indicate that one cause of the problem is the inaccurate estimate of the 3 standard deviations cuts from a small number of protons which then fail to eliminate some protons that have encountered nuclear interactions.

The higher threshold where this happens for the experimental data compared to the simulated data may be explained by the use of ideal detectors in the simulations. A follow up study making use of the simulation framework presented in Giacometti et al. (2017a) would help clarify this.

Given the nature of the FMpCT approach, the dose reduction was not uniform throughout the images (see figures 4.8 and 4.12) but was the highest at the largest distance to the ROI. The imaging dose reduction outside the ROI might be particularly important for proton therapy imaging since the non–target treatment dose is usually lower than for conventional x–ray based radiotherapy, potentially reducing the induction of secondary cancers (Miralbell et al. 2002, Paganetti 2012, Fuji et al. 2013). Avoiding additional non–target dose from image guidance is thus warranted.

For the homogeneous fluence imaging case, the imaging dose was uniform partly due to the fact that the Bragg peak falls outside the object, which means that the protons traversing it produce a relatively flat dose distribution along their path. This situation is very different from kV x–ray CT where higher dose is observed at the object's edge due to the shape of the depth dose profile of photon beams (Bartolac et al. 2011, Arbor et al. 2015).

The dose levels reported in this study for the virtual phantoms are based on MC simulation of ideal detectors. Effect such as pileup would contribute to a higher dose compared to our ideal simulation at equivalent image quality. However, the magnitude of the relative dose reduction should not be greatly influenced by the efficiency of real detectors. The lowest possible FMF where image quality is maintained would however probably be higher for realistic detectors.

The experimental list–mode data made use of 90 projections with $1.2 \cdot 10^4$ cm^{-2} at the center of a projection while the simulated list–mode data used 360 projections with $9.6 \cdot 10^3$ cm^{-2} (3.2 times more protons). For both experimental and simulated data, one should keep in mind that the list–mode data used for image reconstruction are not equivalent to the proton fluence incident on the first tracker module since they consist of protons with entrance and exit coordinates. In the simulated data this excludes mostly protons which underwent nuclear interactions. For the experimental data, various additional detector effects exclude protons from the list–mode data, making estimation of the dose for experimental scans based on simulated scan doses a rough approximation (about 1 mGy). At the time of writing, no direct dose estimation was available for the experimental data. A separate project involving a dosimetric phantom and appropriate detectors is underway by some of the co–authors.

In this first attempt at FMpCT, we made use of a simple binary fluence modulation scheme; a more sophisticated fluence optimization approach may yield different spatial dose distributions for equivalent image quality and deserves further investigation. The conformity of the imaging dose to the ROI may be improved by employing continuously varying fluence modulation profiles and an optimization procedure instead of the binary scheme employed in this proof of principle study. The reason our simple modulation scheme yielded rather acceptable results is attributed to the low attenuation of protons. Indeed, the attenuation of protons traversing 20 cm of water is around 19% for 200 MeV protons and 18% for 300 MeV protons (Quiñones et al. 2016). Additionally, the inelastic proton cross section is almost independent of the proton energy in the energy range $100-300$ MeV. Compared to x–rays, for which 97% of 80 keV photons would be attenuated after 20 cm of water, the calculation of the modulation pattern necessary to obtain the desired image quality is thus more forgiving for FMpCT than FFMCT. Proton dose deposition in the $200-300$ MeV range also has a lower material dependency (mass stopping power ratio) than photons in the diagnostic energy range (mass energy absorption coefficient ratio), yielding more homogeneous imaging dose distributions.

We have found very little published literature on the impact of CT noise on proton therapy dose calculation, with one report stating that discontinuities in the CT number to RSP lookup table might introduce range shifts in the presence of stochastic noise (Brousmiche et al. 2015). This is not the case for proton CT which has a linear lookup table (Arbor et al. 2015). High levels of noise on CT images can additionally cause a spread of the range at which protons stop, the width of which (4 standard deviations) can reach a few percent of the range at 5.0% CT number noise (Chvetsov & Paige 2010).

This study was based on the use of 1 cm (FWHM) proton PBs. This value is a realistic estimation of the PB size available at clinical facilities. The current prototype scanner has been used with PB of minimum size of 4 cm (FWHM). Due to pileup in the tracker, the maximum counting rate for a PB size of 1 cm is approximately 200 kHz. For Pat1, this would mean about 9 s per projection at full fluence, resulting into a total scan time of 13.5 min for 90 projections or 54 min for 360 projections. For realistic scan duration, the electronics of the phase II prototype would have to be modified. The current electronics were designed according to a conservative approach and assuming only scattered proton beams. Therefore, an upgrade to faster electronics is not considered a technical obstacle and significantly faster electronics have been already developed and used in other pCT prototypes (Taylor et al. 2016). To achieve a pCT scan with 1 cm PBs, without any hardware modification to the scanner, a reduction of the beam intensity will be required. It remains to be investigated whether that can allow for a dynamic range necessary for FMpCT studies. Alternatively, a more sophisticated fluence modulation technique than the binary intersection pattern could compensate for the larger PB size.

Although initially explored in the context of pCT with advanced detectors tracking individual protons, the method could also be applicable to integrating detectors such as range telescopes (Krah et al. 2015, Farace et al. 2016) or using 2D detectors (Lee et al. 2015, Tanaka et al. 2016) being also under development. Moreover, the same concept could be extended to imaging with other ions such as ^{4}He and ^{12}C (Shinoda et al. 2006, Hansen

et al. 2014, Rinaldi et al. 2014).

The concept of fluence modulation can be supplemented by energy modulation as well, as pCT imaging dose and noise levels are energy dependent, and we intend to investigate it as a continuation of the current study.

4.5 Conclusion

In this work, we have applied the concept of fluence field modulation to proton CT, inspired by earlier work applied to x–ray CT. Using Monte Carlo simulations of an ideal proton CT scanner, we have confirmed that image quality could be varied across the proton CT image by modulating the proton fluence in a binary fashion. Our approach was successful for both homogeneous and anthropomorphic virtual phantoms, potentially allowing clinical imaging dose reductions ranging from 37% to 56% outside the treatment area, while preserving full fluence image quality inside regions of interest. We additionally virtually implemented the method on broad beam proton CT experimental data and showed that fluence modulated proton CT should be realizable if proton PB intensity and detector count rates can be adjusted to achieve the desired modulation levels.

Acknowledgments

This work was supported by the German Research Foundation (DFG) Cluster of Excellence Munich–Centre for Advanced Photonics (MAP), by the Bavaria–California Technology Center (BaCaTeC) project Nr. 28 [2015-2] and by the Bavaria–France Cooperation Centre (BFHZ). Dr Ute Ganswindt from the Department of Radiation Oncology of the LMU is acknowledged for sharing and delineating the patient CT scan used in this study. The authors thank Erik Traneus from RaySearch Laboratories for his support on the RayStation TPS. Advice on PBS treatment planning from Martin Hillbrand is gratefully acknowledged.

4.6 Supplementary material

Figure 4.12 shows an example of the 2D dose distribution in a slice of the homogeneous phantom, for the case of FF and FMpCT with 0.1·FF for ROI2.

Figure 4.13 shows FMpCT images with 0.1·FF for all three ROIs. The results of the RSP noise and accuracy analysis as a function of the ROI location were as in table 4.4 and did not vary with ROI location.

Table 4.4: RSP noise (1 standard deviation) and accuracy (mean) results in ROI2. The second column reports RSP noise values for uniform scans with fluence modulation as indicated in the first column. The third column depicts the noise of FMpCT scans. Similarly, the reconstructed RSP mean values in the fourth and fifth columns are given for uniform and FMpCT simulated scans. The reference RSP of the simulated material was 1.02.

Modulation	$(RSP–RSP_{ref})/RSP_{ref}$ (%) Noise		$(RSP–RSP_{ref})/RSP_{ref}$ (%) Mean	
	Uniform	FMpCT	Uniform	FMpCT
FF	1.2	–	-0.2	–
$0.1 \cdot FF$	4.1	1.2	-0.2	-0.2
$0.05 \cdot FF$	7.0	1.2	-0.3	-0.3
$0.01 \cdot FF$	19.0	1.2	-0.2	-1.3

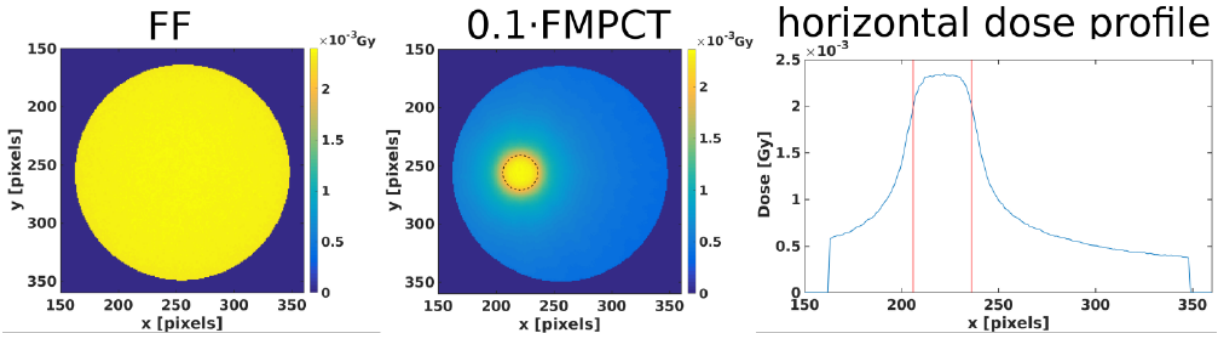

Figure 4.12: Imaging physical dose distribution in the homogeneous phantom for FF (left). Same quantity shown for an FMpCT scan (0.1·FF) for ROI2 (center). The imaging integral physical dose was reduced from 2.3 mGy in the FF image to 0.7 mGy in the FMpCT image. (right)A dose profile drawn horizontally through ROI2 is shown.

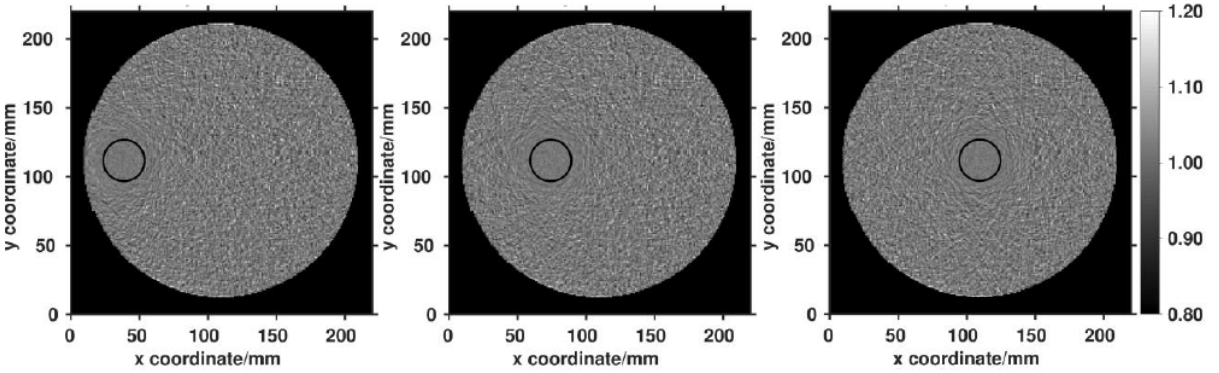

Figure 4.13: Three different ROIs defined in the homogeneous phantom, indicated by the black circles, are overlaid on FMpCT images of 0.1·FF.

Chapter 5

Experimental fluence–modulated proton CT

The material contained in this chapter was published in Medical Physics, in volume 45, issue 7, pages 3287–3296, in July 2018, with the title *Experimental fluence-modulated proton computed tomography by pencil beam scanning*, by Dedes et al. (2018).

5.1 Introduction

The use of x–ray computed tomography (CT) scans for relative (to water) proton stopping power (RSP) estimation, a quantity necessary for dose calculation, contributes considerably to range uncertainties in proton therapy (Paganetti 2012, Yang et al. 2012). The potential of reducing these uncertainties by direct RSP measurements at the treatment position has motivated the recent revival of proton computed tomography (pCT), which was first proposed by Cormack (1963) in the early 1960s. By measuring the positions and residual energies of the protons behind (and in some designs also in front of) the patient in a series of projections, a RSP image can be reconstructed (Hansen et al. 2016, Penfold et al. 2009, Penfold et al. 2010, Poludniowski et al. 2014, Rit et al. 2013). Currently, several groups are known to be designing, building, or operating pCT (or heavier ion CT) prototypes (Rinaldi et al. 2013, Sadrozinski et al. 2016, Taylor et al. 2016, Meyer et al. 2017, Tanaka et al. 2018), and initial reports of RSP accuracy are encouraging (Giacometti et al. 2017a).

The eventual use of pCT for frequent imaging in treatment position is supported by the fact that pCT dose efficiency, evaluated by metrics such as contrast to noise ratio, is superior to x–ray CT (Schulte et al. 2005) and may allow lower imaging doses. Recently, pCT scans using imaging doses as low as 1 mGy have been achieved (Sadrozinski et al. 2016). The imaging dose from pCT may be further reduced by employing the concept of fluence field modulation (Bartolac et al. 2011, Bartolac & Jaffray 2013, Graham et al. 2007b, Stayman et al. 2016, Szczykutowicz et al. 2015, Szczykutowicz & Mistretta 2013a, Szczykutow-

icz & Mistretta 2013b), where the fluence of particles used for imaging is adjusted within a projection to yield spatially varying image quality. For proton therapy guided by pCT, we could thus envision imaging the therapeutic beam paths with high image quality (low noise) and sacrificing image quality (high noise) where therapy–related dose computation is not required. This is particularly attractive for proton therapy given the low–integral dose nature of the modality (Fuji et al. 2013). A recent simulation study suggested that fluence–modulated pCT (FMpCT) may yield imaging dose reduction of 30% and 50% in head–and–neck and brain cancer patients, respectively, while preserving the dose calculation accuracy of a full fluence image (Dedes et al. 2017).

FMpCT may be achieved by acquiring pCT scans using the pencil beam scanning (PBS) functionality of modern proton therapy facilities instead of the broad beams (cone or wobbled beams typically) currently employed in most pCT prototypes (Sadrozinski et al. 2016). The use of PBS would greatly simplify the fluence modulation task by allowing the prescription of proton fluence on a pencil beam by pencil beam basis. PBS scans have been previously acquired for carbon ion CT with a flat panel detector (Telsemeyer et al. 2012) as well as with an ionization chamber stack detector (Meyer et al. 2016, Magallanes 2017) which functions in PBS mode by design (Rinaldi et al. 2013). Helium CT results obtained from PBS have also been presented for a particle tracking scanner, and the same study reported the acquisition of PBS pCT as well, but no results were reported (Volz et al. 2017).

In this study, we performed PBS pCT scans and first experimental realization of FMpCT by making use of the phase II preclinical prototype pCT scanner of the Loma Linda University and U.C. Santa Cruz, together with the PBS capability of the Northwestern Medicine Chicago Proton Center. Our aim was to demonstrate the equivalence, in terms of image quality, of PBS and broad beam acquisitions, and the feasibility of FMpCT using a simple water cylinder phantom with a centrally located FMpCT region of interest (FMpCT-ROI) for high image quality. Specifically, this work aimed at demonstrating that (a) PBS pCT can achieve the same image quality as cone beam or wobbled beam pCT in uniform fluence scenarios and (b) that FMpCT scans preserving image quality in the FMpCT-ROI are feasible with PBS pCT.

5.2 Materials and methods

5.2.1 Experimental setup

The experimental part of this study was performed at the Northwestern Medicine Chicago Proton Center using the phase II preclinical pCT prototype scanner of the Loma Linda University and U.C. Santa Cruz, described in detail in Johnson et al. (2016). The scanner consists of two tracking modules and a five–stage water equivalent path length (WEPL) detector coupled to a data acquisition system capable of acquiring broad beam proton events at a sustained rate in excess of 1 MHz.

The front and rear tracking modules together contain 32 single-sided silicon strip detectors (SSD) with a strip pitch of 228 and 400 μm thickness. Four SSDs are assembled

on a printed circuit board (PCB) in a pattern that forms a continuous sensitive area of $8.95 \times 35.6\,\mathrm{cm}^2$. Two back–to–back PCBs with orthogonal orientation form a 2D coordinate detector, and each tracking module consists of two such 2D coordinate detectors. The tracking system therefore provides four sets of 3D coordinates, allowing the estimation of the curved proton path using the measured position and direction vector of each proton before and after the object.

The five–stage plastic scintillating detector has a dynamic range of 0–260 mm WEPL. Each 52 mm thick stage is made of polystyrene of RSP=1.038 and has a lateral area of $10 \times 40\,\mathrm{cm}^2$. The WEPL information for every proton is given by the signal from the stage in which the proton stopped (Bashkirov et al. 2016).

5.2.2 PBS beam line

The PBS beam was delivered in a clinically commissioned fixed beam line using a universal nozzle manufactured by IBA (IBA, Belgium). The accelerator source for this room is an IBA C230 cyclotron that has been in clinical use since 2010 at the Northwestern Medicine Chicago Proton Center. For proton imaging, extremely low proton flux is needed. The standard method of regulating dose using signal from the nozzle's monitor chambers was not feasible due to the very low signal to noise ratio of the imaging beam. Instead, spot dwells were based on a fixed dwell time, which is the beam–on time at a specific position, along with setting the cyclotron output current once to provide the desired total fluence and flux.

5.2.3 Phantom

For this experimental proof–of–principle work, we chose a simple cylindrical water phantom. The phantom was made of a PMMA cylindrical container with inner diameter of 137.8 mm and outer diameter of 150.5 mm. The height of the cylinder was 28 mm and the two PMMA endcaps had a thickness of 6 mm each. The container was filled with water and placed on a remotely controlled rotating stage.

5.2.4 PBS scan patterns

Proton CT images were acquired using active pencil beam (PB) scanning. A 200 MeV PB with a spot size of 1.37 cm FWHM in air at isocenter was magnetically deflected in order to cover an area of $30 \times 10\,\mathrm{cm}^2$ in the case of the calibration runs (300 PBs) and $20 \times 40\,\mathrm{cm}^2$ in the case of the imaging scans (200 PBs). The beam was turned off for the duration of the deflection to the next point. Beam spot locations were spaced 1 cm apart. The pencil beam grid was offset by a quarter of the PB spacing in the direction normal to scan's axis of rotation to avoid overlapping PB from opposing projections and thereby improve the measurement uniformity.

5.2.5 Pileup study

The pCT phase II scanner had been so far mostly used with broad proton beams: either a passively scattered cone beam at the Loma Linda University Medical Center (Bashkirov et al. 2016a), covering the whole field of view of the scanner, or at the Northwestern Medicine Chicago Proton Center with a wobbled proton beam of 4–7 cm FWHM spot size (Sadrozinski et al. 2016). The single report of scanner operation in PBS mode makes no mention of pileup (Volz et al. 2017). Before performing the calibration and imaging runs, a study was made on the detector performance for a 1.37 cm FWHM PB, concentrating especially on the resulting pileup in the silicon strips of the tracker modules. The main concern was that the amplifier time–over–threshold for the silicon strips could be such that if two protons impinge on the same strip within a time window of less than about one microsecond, the second one will be missed.

For the purpose of the pileup study, a $30 \times 10\,\mathrm{cm}^2$ field (300 PBs) was employed, with the beam running at cyclotron currents of 2, 4, 6, 8, 10, and 12 nA in six different runs and impinging upon the detector without any object in the field of view. The PB dwell time was 8 ms. This dataset was employed to select a suitable cyclotron current, taking into account the trade–off between pileup and acquisition speed. No formal optimization was done, however, as a decision for the subsequent runs had to be made within a few minutes.

We have subsequently further analyzed the data from these six runs in order to understand the impact of pileup on the efficiency of individual proton detection in the same PB. The efficiency to detect a proton "hit" in the silicon–strip detectors can be directly measured in a given layer for either of the two views (vertical strips, coordinate V and horizontal strips, coordinate T) by fitting a straight line to the hits in the other three layers to predict the expected location of a hit in the layer of interest and verifying whether it was recorded. With zero pileup that efficiency is generally found to be above 99%, with most of the loss coming from the gaps between individual silicon wafers (Johnson et al. 2016).

5.2.6 Calibration run

A calibration procedure that transforms the detector single–proton measurement to WEPL is necessary and is discussed in Bashkirov et al. (2016). The pCT collaboration has developed an improved version in which a polystyrene wedge phantom is used in ref. (Johnson et al. 2017) instead of the stepped pyramid phantom. In addition to the wedge, four polystyrene blocks are required to bring the proton Bragg peak into each stage of the five–stage plastic scintillator detector (Bashkirov et al. 2016).

The calibration was performed at the beginning of a scanning session in five consecutive runs, in addition to a run made without an object. The wedge provides a continuous WEPL range that is traversed by protons, whose position and hence path length in the phantom is provided by the tracking system. The first calibration run was acquired with only the wedge, resulting in protons stopping at different depths in the last stage. Four additional runs were acquired by adding every time one block behind the wedge, so that protons stop in each of the four remaining stages. A run with no phantom was also acquired and

was used to map the spatial dependence of the scintillator signals as well as to provide a conversion from analog–to–digital converter (ADC) counts to MeV (Bashkirov et al. 2016). The result of the calibration run was a look–up table transforming an energy measurement to the traversed WEPL for a proton stopping in a particular stage. All calibration runs were acquired using a $30 \times 10\,\mathrm{cm}^2$ (300 PBs) field with a cyclotron current of 4 nA with 8 ms dwell time per PB position, with six repetitions of the scan pattern.

5.2.7 pCT scan acquisitions

Each scan consisted of 45 projections, acquired at eight–degree steps, with the phantom rotating between the acquisitions of two consecutive projections (step–and–shoot mode). Following the pileup study and calibration, a $20 \times 10\,\mathrm{cm}^2$ scanning pattern with 27 ms dwell time delivered in a single pass with a cyclotron current of 4 nA was chosen. This resulted in a 400 kHz counting rate (averaged over a projection).

The fluence of each PB was controlled by changing its dwell time. For the full fluence (FF) scan, the dwell time was kept at 27 ms. In that case, the acquisition of a single projection required 7 s, and 2.2 million protons were recorded per projection. For the FMpCT scans, a FMpCT–ROI was imaged with the FF, while for the remaining PBs in the $20 \times 10\,\mathrm{cm}^2$ imaging field, the PB fluence was reduced by 50% (FMpCT50) and 20% (FMpCT20) by reducing the dwell time to 13.5 and 5.4 ms, respectively. The central FMpCT–ROI was defined as the region irradiated by the 5·10 central PBs. An unintentional misalignment caused a shift of 1 cm in the PB scan pattern, resulting in a smaller FMpCT–ROI covered by FF PBs in all projections. The final effective size of the FMpCT–ROI was 3 cm in diameter and 10 cm in height.

5.2.8 Image reconstruction

Images were reconstructed for this study with a filtered backprojection (FBP) implementation that accounts for the curved proton paths in the imaged object, mainly due to multiple Coulomb scattering. The main principles of the algorithm are presented in (Rit et al. 2013). The path of every proton is approximated by a most likely path (MLP) formulation introduced by Schulte et al. (2008), which uses the position and direction information from the tracking modules (see figure 3 in that publication for simulated proton paths and their MLP).

Protons were selected with a 3 standard deviations cut on the energy and angular distributions around their mean energy and angle per projection pixel. List–mode data were binned in intermediate projection images with $0.8 \times 0.8\,\mathrm{mm}^2$ pixels, which were then filtered and backprojected. Sinogram interpolation (Hansen et al. 2016) was used for all reconstructions.

To better understand the noise behavior outside the FMpCT–ROI, images were additionally reconstructed from homogeneous scans with fluence virtually reduced to 50% and 20% of the FF scan. Those scans were not explicitly acquired during the experiment but

were obtained from the FF dataset by assigning a uniform selection probability to every proton.

Finally, the results of the PB pCT scans acquired in this experiment were compared to an older scan of the same phantom imaged with a wobbled proton beam with a size of 4 cm FWHM (Sadrozinski et al. 2016). To ensure comparability, the wobbled beam scan data were processed so as to contain the same number of protons per projection and the same number of projections as the FF PB scan.

5.2.9 RSP noise and accuracy quantification

The image quality of the acquired pCT scans was quantified in terms of RSP accuracy and RSP noise. Utilizing the cylindrical symmetry of the scanned object, the image quality metrics were assessed in annular ROIs with increasing radius, covering different radial extents of the phantom. For comparability, all ROIs contained approximately 1000 pixels, resulting in variable annulus thicknesses. The RSP accuracy in an annulus was defined as the mean value of the RSP distribution of all pixels contained in it, expressed as a percentage difference from the theoretical value for water (RSP = 1). The RSP noise in an annulus was defined in a similar manner, using the standard deviation of the RSP distribution. Finally, the inner radius of each annulus denotes the radial distance from the center of the cylindrical phantom. The ROIs are shown in figure 5.11 (supplementary material).

5.2.10 Imaging dose

During the experiment, there was no possibility of dose estimation for each scan. In order to quantify the dose gains of FMpCT with respect to the FF scans, we simulated the dose in the water phantom using Monte Carlo simulations. A full simulation was employed that modeled in high detail all the active and passive scanner elements. The simulation application based on GEANT4, version 10.02.p1, was presented in Giacometti et al. (2017a). The dose in the water phantom was estimated by a GEANT4 primitive dose scorer in a $1 \times 1 \times 1\,\mathrm{mm}^3$ voxel grid using a proton phase space distribution estimated from the front tracking module. The resulting dose distributions were normalized to the FF scan to estimate dose savings from FMpCT. Absolute doses were however not reported.

5.3 Results

Evaluation of pileup with pencil beam scanning

Figure 5.1 shows the efficiency measurements resulting from the pencil beam runs. The highest trigger rate of 910 kHz is only about 30% below the maximum event rate that the data acquisition can sustain. At the data rate (400 kHz) employed in the experiments described below, the hit inefficiency caused by pileup is only about 1% and is thus negligible.

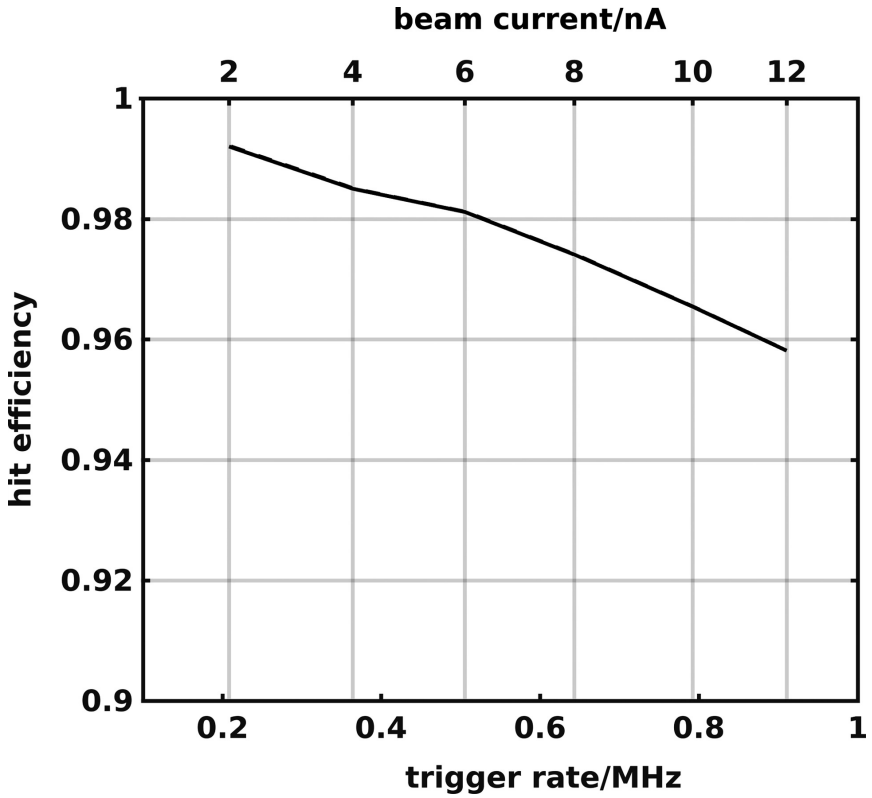

Figure 5.1: The measured hit efficiency averaged over all eight tracker layers vs the trigger rate. The falloff in efficiency with increasing rate is due to pileup in the amplifiers of the silicon–strip detectors.

Figure 5.2 shows the fraction of events for which a proton track could be reconstructed along the vertical and transverse directions as a function of the trigger rate corresponding to the different cyclotron currents considered. We observed an increase in pileup with increasing current, with the tracking efficiency dropping from 90% at 2 nA and 200 kHz to 85% at 12 nA and 900 kHz. The remainder of the experiments and the results presented in this section were obtained with 4 nA and a trigger rate of 400 kHz.

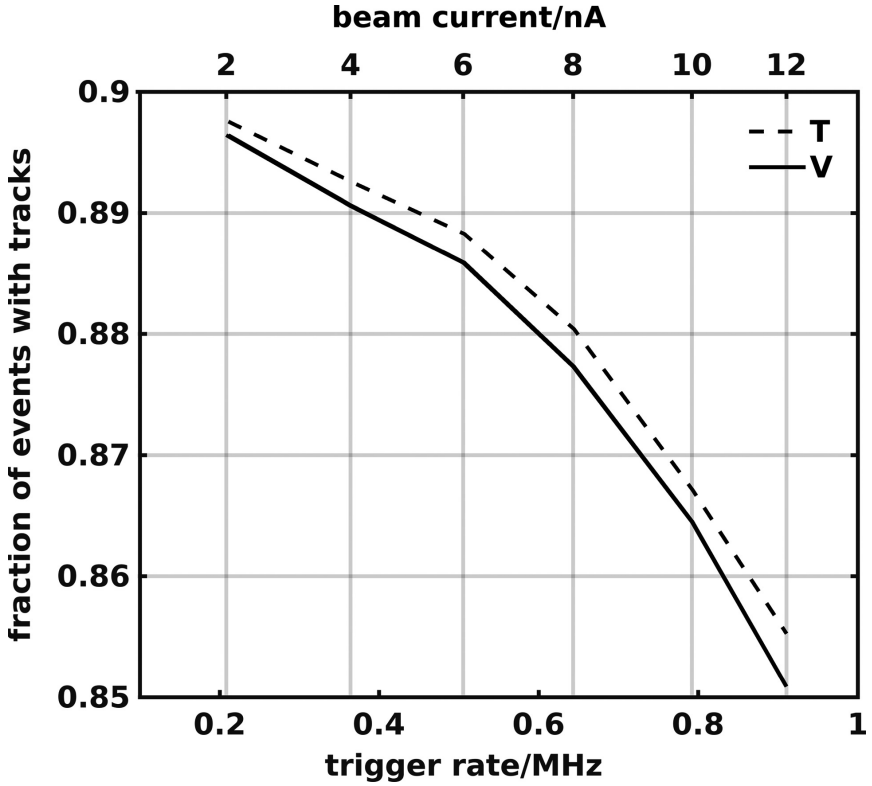

Figure 5.2: The fraction of triggered events for which proton tracks in the vertical (V) and transverse (T) directions could be obtained as a function of the trigger rate or cyclotron current.

5.3.1 Comparison to established imaging with wobbled delivery

Figure 5.3 shows the detector's stage–wise energy–to–WEPL calibration curves obtained with an open field using the wobbled beam and PBS. The curves nearly overlap, indicating that the PBS and wobbled beam calibrations are equivalent.

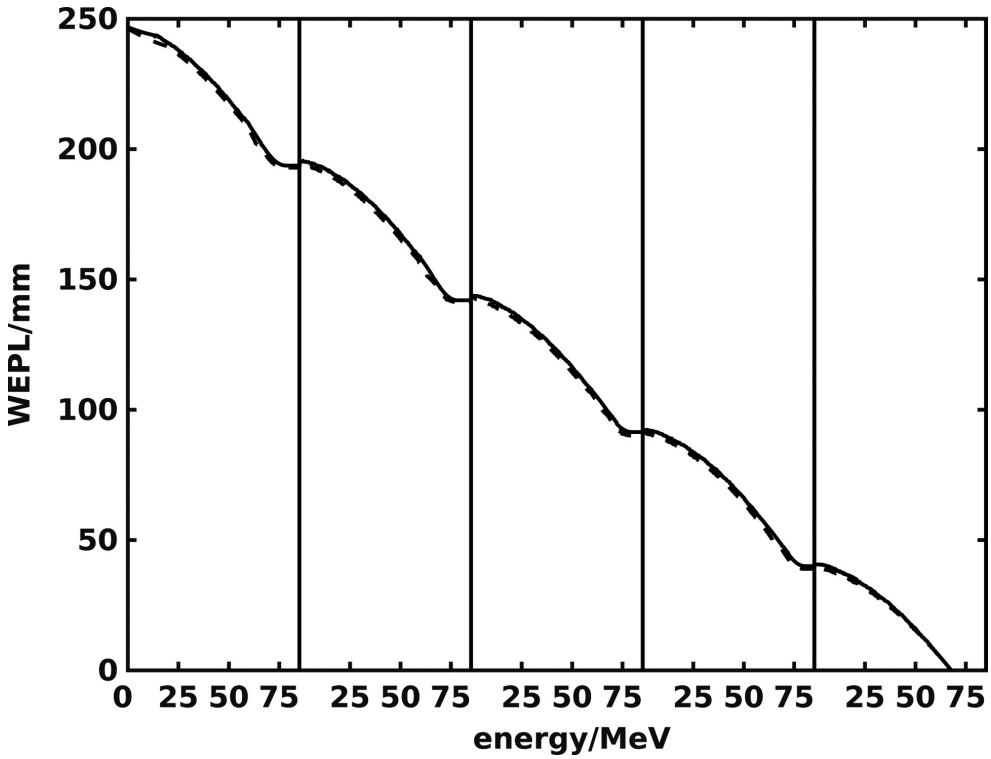

Figure 5.3: Calibration plots for the five consecutive energy detector stages based on the wobbled (dashed line) and pencil beam scanning (solid line) beams. The horizontal scale refers to the energy deposition in the stage in which the proton apparently stopped.

Figure 5.4 shows reconstructed pCT images from full fluence scans of the water cylinder phantom acquired with a wobbled beam and with pencil beam scanning. The images are generally similar, with the wobbled beam exhibiting a slightly darker artifact at the center of the water cylinder. Using a ROI covering 90% of the water cylinder's radius, the mean (standard deviation) of the water RSP were 1.011 (0.053) and 0.993 (0.049) for the wobbled and pencil beams, respectively. Figure 5.5 makes use of annular ROIs (see figure 5.11 – supplementary material) to present the RSP accuracy and noise. We observed that the pencil beam scanning images had better accuracy (results closer to 0% error), and that the difference between the wobbled and PBS images was limited to about 1%. Both images showed increasing noise with radius. There was good agreement between the noise levels in the center of the cylinder, but slightly higher values for the wobbled beam were observed at the object's edge.

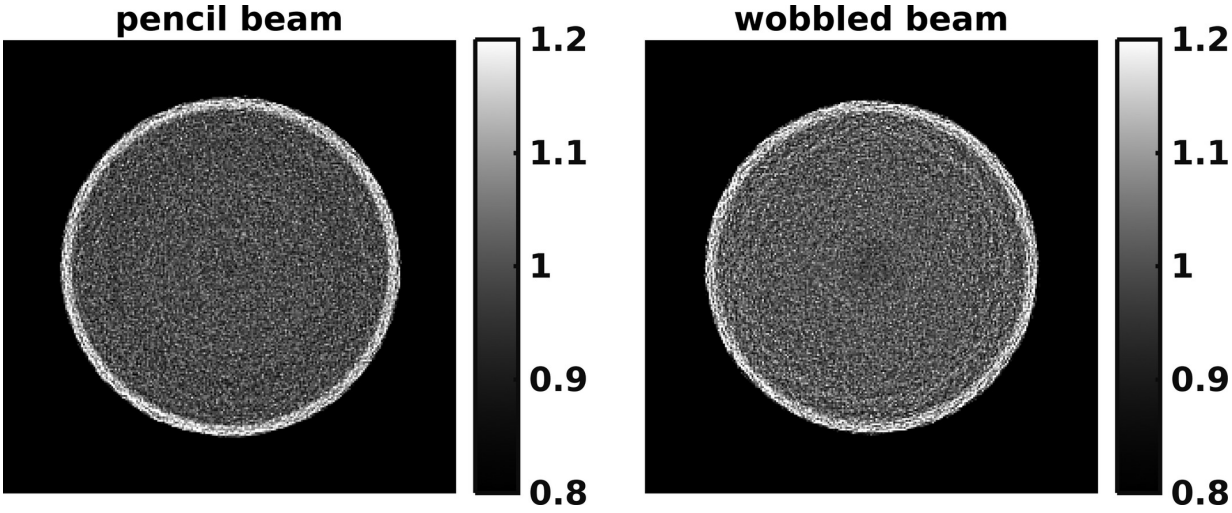

Figure 5.4: pCT images (image values are RSP) acquired with pencil beam scanning and a wobbled beam.

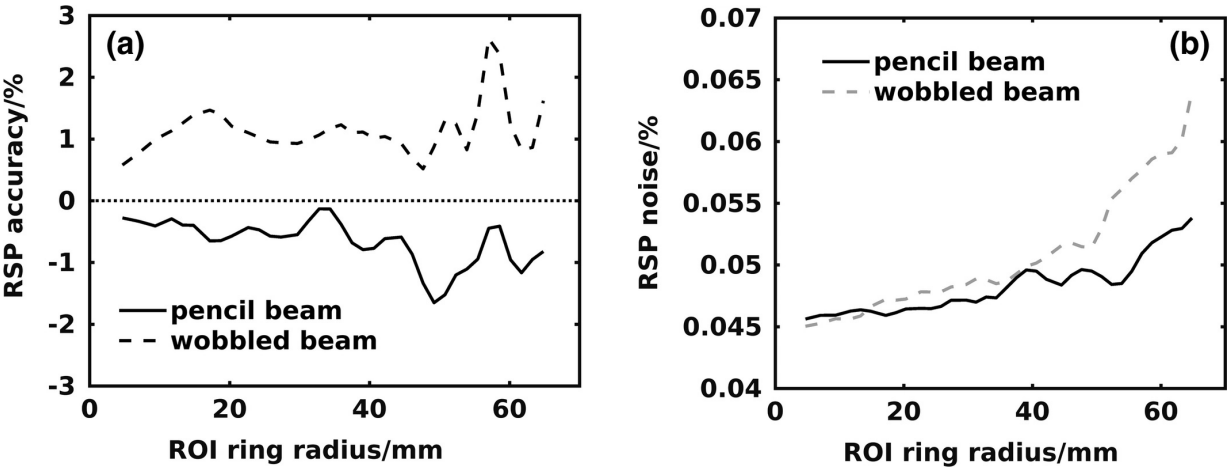

Figure 5.5: (a) Error on mean RSP of water estimated using annular ROIs in pCT images acquired with pencil beam scanning and a wobbled beam. (b) Corresponding standard deviation (noise).

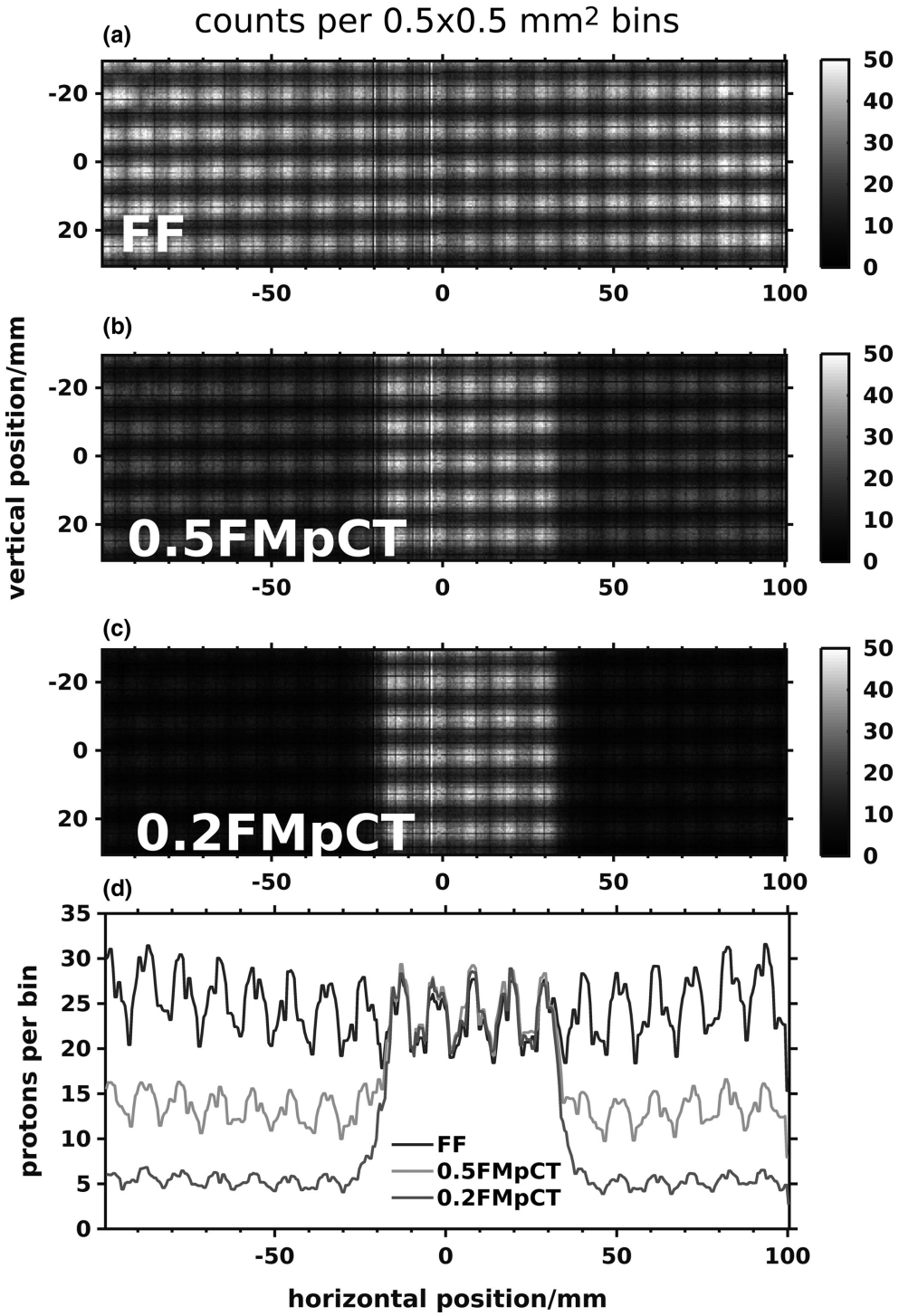

Figure 5.6: Front tracker proton counts for (a) full fluence (FF), (b) FMpCT with FMF 0.5, and (c) 0.2. The horizontal axis is parallel to the cylindrical phantom's diameter and the vertical axis is parallel to the phantom's rotation axis. (d) Horizontal profiles of proton counts per bin averaged along the vertical axis of (a–c). (d) A 1–dimensional median filter was applied for display purposes.

5.3.2 FMpCT

Figure 5.6 presents the front tracker proton counts showing the pencil beam fluence used for each projection of the FF and FMpCT scans. Individual pencil beams are resolvable as well as the quarter detector shift employed to increase the in–slice dose uniformity. We observed that due to an experimental misalignment of the pencil beam scanning system and the phantom's rotation stage, the high fluence region (5 PBs) was shifted by one pencil beam. We also noticed a slight horizontal tilt attributed to the scanner not being perfectly leveled. The pattern of darker lines is attributed to gaps between active tracker elements.

Figure 5.7 presents the reconstructed PBS pCT images for varying levels of uniform fluence as well as FMpCT images with FMF of 0.5 and 0.2. The expected increase of noise with lower proton fluence was observed for uniform fluence images, while with FMpCT, the image quality in the center of the phantom appears qualitatively preserved and the noise at the object's edge tends to approach the uniform fluence scenario.

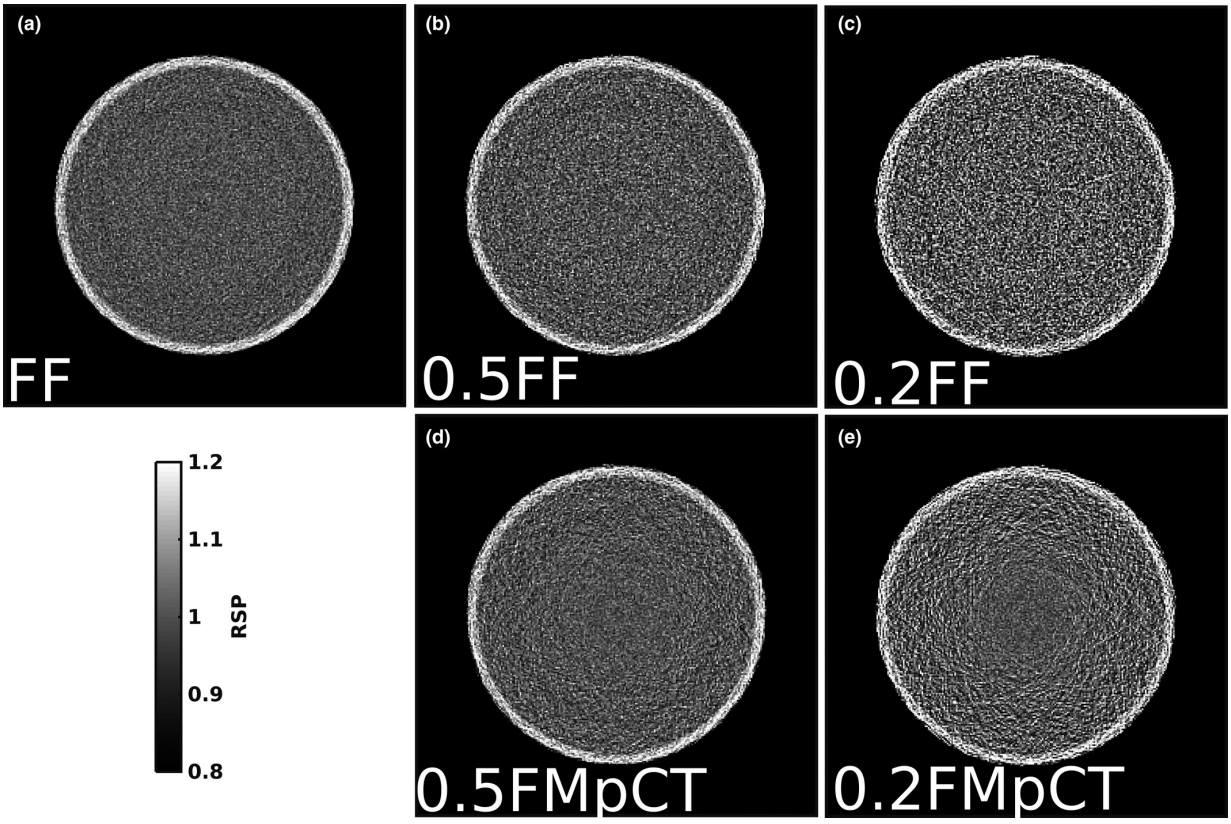

Figure 5.7: PBS pCT RSP images with (a) homogeneous full fluence (FF) and reduced uniform fluence by factors (b) 0.5 and (c) 0.2. (d,e) PBS FMpCT RSP images with FMF of (d) 0.5 and (e) 0.2.

This finding was quantified in figure 5.8, where the standard deviation in the annular ROIs is reported for the images presented in figure 5.7. We can observe that in the low noise ROI, the standard deviation of the FMpCT images matches the one obtained from the FF

image, while the uniform low fluence images have overall higher noise. For FMpCT images, the noise increases rapidly with the distance from the ROI, approaching the uniform fluence cases at the edge of the object.

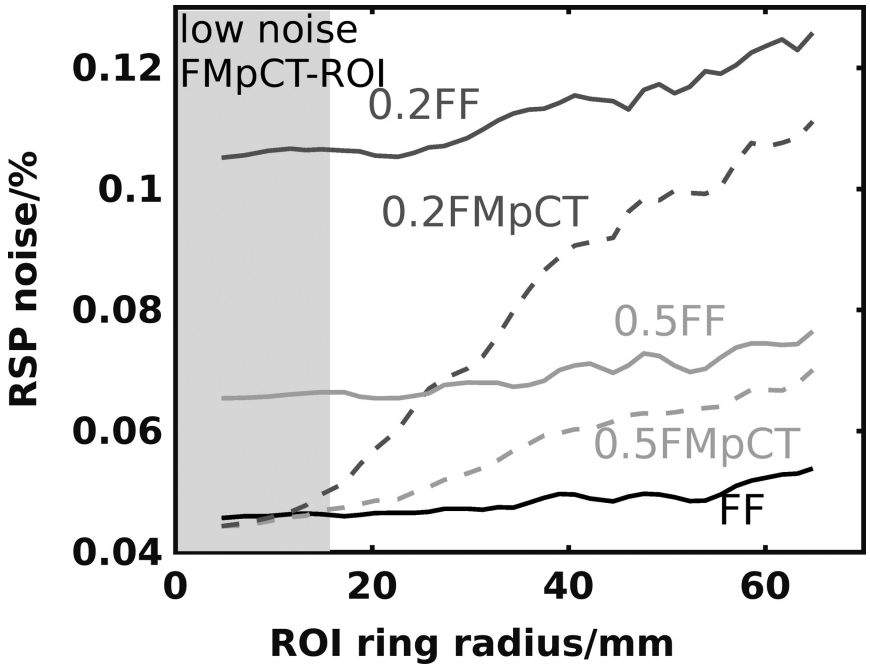

Figure 5.8: Standard deviation (noise) of RSP of water estimated from pCT images using annular ROIs. The extent of the FMpCT–ROI is indicated in gray.

Figure 5.9 reports the RSP error (for water) of the mean values for the uniform FF and FMpCT images in the annular ROIs. We see that the use of FMpCT with FMF 0.5 does not alter the RSP accuracy inside and outside the low noise ROI. A slight variation of accuracy for FMF 0.2 was observed in the low noise FMpCT–ROI. This can be observed in table 5.1 where we see that the mean value in the case of FMpCT with FMF 0.2 is 0.2% lower than in the corresponding uniform fluence case.

Table 5.1: Mean RSP noise and standard deviation in the FMpCT–ROI. The second column shows RSP mean values for uniform scans with different FMF. The third column reports mean values with FMpCT. The fourth and fifth columns refer to the corresponding RSP standard deviations in the same ROI.

	Mean		Standard deviation	
Pat1	Uniform	FMpCT	Uniform	FMpCT
FF	0.992	–	0.047	–
$0.5 \cdot$ FF	0.992	0.992	0.066	0.046
$0.2 \cdot$ FF	0.992	0.990	0.109	0.044

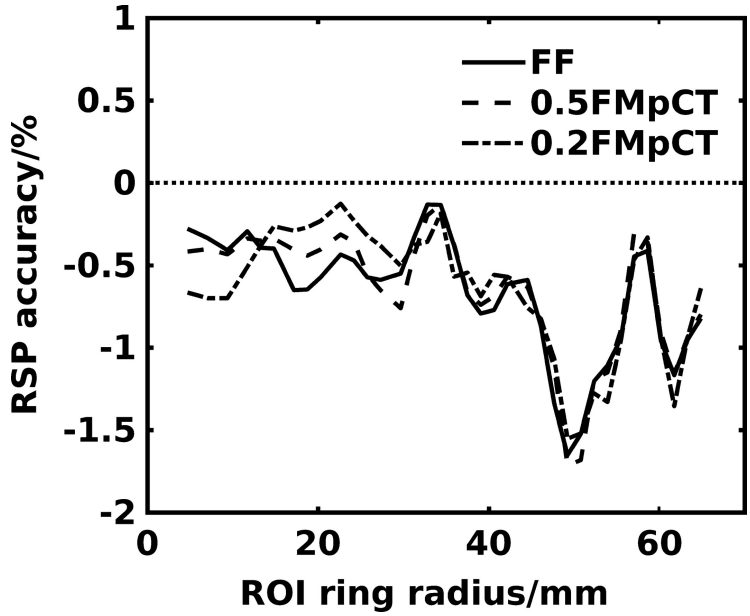

Figure 5.9: Error on mean RSP of water estimated from pCT images using annular ROIs.

Figure 5.10 presents the relative dose distributions, normalized to their maximum, of the FF and FMpCT scans. We observed the expected dose reduction outside the FMpCT–ROI when compared to the mostly uniform FF dose. At the object's edge, the imaging dose was 60% and 40% of the FF dose for FMpCT with FMF of 0.5 and 0.2, respectively, while in the center, the same dose was obtained for all cases. Due to the misaligned high fluence pattern, an intermediate 80% of FF dose halo was observed around the maximum dose region.

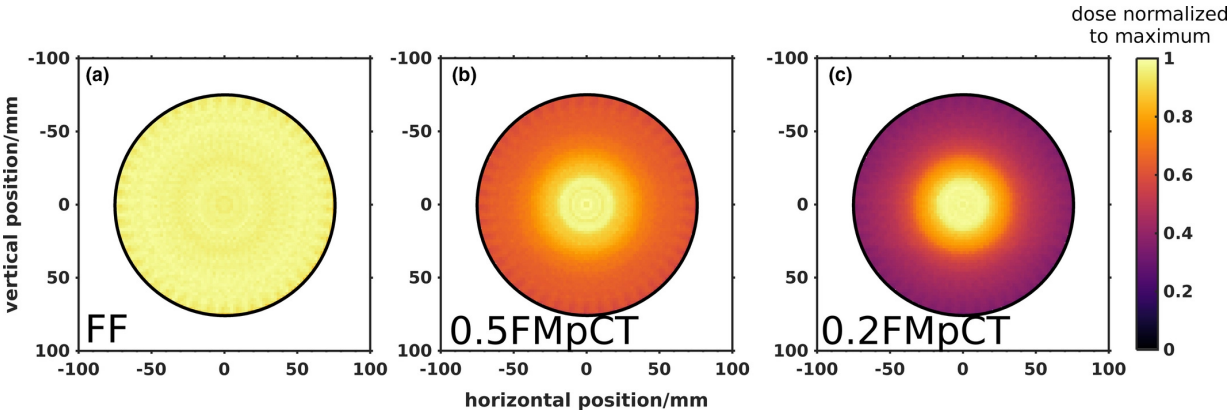

Figure 5.10: Normalized pCT scan doses in a central slice of the phantom (circumference indicated by dashed line). (a) Full fluence (FF), FMpCT with (b) FMF 0.5 and (c) FMF 0.2.

5.4 Discussion

In this experimental study, we have demonstrated that the phase II pCT scanner prototype could be operated in PBS mode at acceptable pileup levels. Previous mentions of scanner operation in proton PBS mode did not present pileup analyses (Volz et al. 2017). The chosen cyclotron current of 4 nA led to a scanner trigger rate of approximately 400 kHz. This was lower than the 1 MHz rate achievable with wobbled or cone beams (Sadrozinski et al. 2016). In retrospect, even at the 1 MHz rate typically used with this scanner, the pileup inefficiency would be acceptable. Although 96% hit efficiency would result in only 70% of protons being measured with all eight hits, our analysis generally allows a single missing hit in both views, considerably reducing the impact of the inefficiency. Furthermore, when a phantom is placed in the beam, multiple scattering broadens the beam in the rear tracker, significantly reducing pileup. Note that loss of single hits in the front tracker is relatively unimportant given that the incoming beam direction is well known.

The experiment was performed using slow step–and–shoot acquisition mode. The suggested increase in cyclotron current and pencil beam size might facilitate scan acquisition using continuous rotation. The step–and–shoot mode we employed would not be compatible with the time constraints associated with clinical operation. Realizing FMpCT with PBS based on continuous rotation in a reasonable time frame (about 1 min) is one the main challenges to this approach, and subsequent studies will aim at evaluating whether this goal is achievable with current beamlines and acquisition rates.

We made use of a quarter PB spacing shift to interleave PBs from opposing projections, thus rendering the dose in a slice homogeneous. However, in the vertical dimension, interleaving is not possible and the dose distribution shows peaks and valleys corresponding to the PB pattern shown in figure 5.6. It might be necessary to reduce the PB spacing in that direction to ensure homogeneous image quality.

The accuracy of the RSP for water was mostly within 1% for the PBS pCT scan and comparable or better than what had been achieved with the wobbled beam. For both scan acquisition methods, we observed increasing noise with object radius. This is currently the subject of a separate investigation and requires careful modeling of noise reconstruction in pCT as well as understanding of the various sources of projection noise, including detector effects. These might be modeled using the simulation platform of Giacometti et al. (2017a). The slightly higher noise levels at the object's edge observed with the wobbled beam may be caused by a falloff of the fluence at the edge of the beam; however, this has yet to be verified.

We were able to confirm the results of a previous simulation study suggesting that FMpCT should be feasible without loss of accuracy in the FMpCT-ROI (Dedes et al. 2017). We saw a very slight change (0.2%) in the water RSP when employing the 20% FMF. This appears consistent with the findings of Dedes et al. (2017), where accuracy was degraded when using FMF lower than 30% (see figure 11 in that reference). The FF noise level in the FMpCT–ROI was preserved for both FMF values, and we observed that the noise level outside the FMpCT–ROI approached the uniform low fluence results with increasing distance from the FMpCT–ROI. At a given FMF, the FMpCT scan noise was always lower

than the uniform fluence scan noise since the full fluence PBs cross any given voxel at least once.

The relative maximum dose reduction achieved in this study was 40% and 60% for FMF of 50 and 20, and is comparable to what was achieved in the simulation study of Dedes et al. (2017). Of course, these dose savings highly depend on the shape of the FMpCT–ROI. In this study, we opted for a simple central FMpCT–ROI, thus allowing us to neglect the synchronization of the PB fluence pattern and rotation angle. This was deemed sufficient for a proof–of–principle study, and next experiments will aim at performing FMpCT for arbitrarily shaped FMpCT–ROI. The unfortunate misalignment of the PB high fluence pattern meant that the dose fall–off with distance from the FMpCT–ROI's edge was not as sharp as it could have been.

5.5 Conclusion

In this study, we have demonstrated that FMpCT scans are feasible. Furthermore, we have shown that PBS pCT scans can achieve equivalent accuracy as those obtained from broad beams. Image accuracy and noise were successfully preserved in the central FMpCT–ROI chosen for this study, and dose reduction of up to 60% at the object's edge was realized.

Acknowledgments

This work was supported by the German Research Foundation (DFG) project #388731804 "Fluence modulated proton computed tomography: a new approach for low-dose image guidance in particle therapy" and the DFG's Cluster of Excellence Munich–Centre for Advanced Photonics (MAP), by the Bavaria–California Technology Center (BaCaTeC) project Nr. A1 [2018-2] and by the Bavaria–France Cooperation Centre (BFHZ). Dr. Valentina Giacometti is gratefully acknowledged for developing and sharing the Geant4 simulation platform used for dose calculation.

5.6 Supplementary material

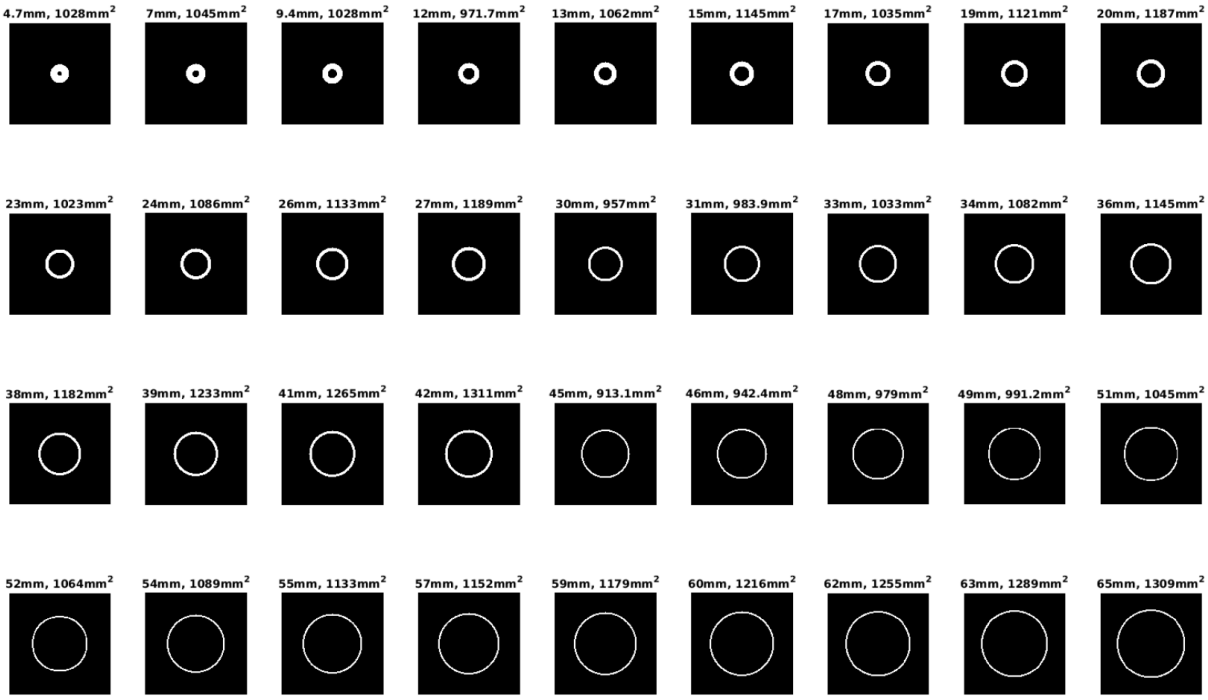

Figure 5.11: The annular regions of interest (ROI) used to analyse the relative stopping power (RSP) images reconstructed in this work. The inner radius and area are reported above each ROI.

Chapter 6

Noise reconstruction in proton CT

This chapter contains the work published in Physics for Medicine and Biology, volume 63, issue 21, page 215009, in October 2018, with the title *Two dimensional noise reconstruction in proton computed tomography using distance–driven filtered back-projection of simulated projections*, by Rädler et al. (2018) – (senior author: Dedes).

6.1 Introduction

Relative proton stopping power (RSP), the ratio of the proton stopping power of a given material to that of water at a given energy, is necessary for most clinical dose calculation methods used in proton therapy. RSP is currently estimated by x–ray computed tomography (CT) scans in clinical practice (Taasti et al. 2018). The conversion from photon attenuation coefficients to RSP contributes considerably to range uncertainties (Paganetti 2012, Yang et al. 2012). The potential of reducing these uncertainties by direct RSP measurements at the treatment position has motivated the recent revival of proton computed tomography (pCT), which was first proposed by Cormack in the early 1960s (Cormack 1963). Modern pre–clinical pCT scanners measure the positions and residual energies of the protons behind (and in some designs also in front of) the patient in a series of projections, from which an RSP image can be reconstructed (Hansen et al. 2016, Penfold et al. 2009, Penfold et al. 2010, Poludniowski et al. 2014, Rit et al. 2013). Many groups are known to be designing, building or operating pCT (or heavier ion CT) prototypes (Rinaldi et al. 2013, Sadrozinski et al. 2016, Taylor et al. 2016, Meyer et al. 2017, Tanaka et al. 2018) and initial reports of RSP accuracy support these endeavours (Giacometti et al. 2017a).

The concept of fluence field modulation computed tomography (FFMCT), initially suggested for x–ray CT by Graham et al. (2007b) and pioneered by the Toronto (Bartolac et al. 2011, Bartolac & Jaffray 2013) and Madison groups (Szczykutowicz & Mistretta 2013a, Szczykutowicz & Mistretta 2013b), allows the tailoring of the spatial distribution of image noise and dose by modulating the x–ray fluence within a given CT projection. Fluence modulation has been realized by employing a digital beam attenua-

tor (Szczykutowicz & Mistretta 2014), the binary collimator of a Tomotherapy machine (Szczykutowicz et al. 2015), multiple aperture devices (Stayman et al. 2016) or piecewise-linear dynamic attenuators (Shunhavanich et al. 2018). While fluence modulation capability is crucial in achieving FFMCT, a mathematical model relating x–ray fluence and image noise and/or radiation dose is required to optimize the FFMCT fluence pattern (Bartolac et al. 2011). Several publications cover the theory of noise reconstruction for x–ray CT for parallel (Gore & Tofts 1978, Huesman et al. 1977, Huesman 1984, Kak & Slaney 1988, Buzug 2008), fan (Wunderlich & Noo 2008) and cone beam (Zhang & Ning 2008, Shäfer et al. 2015) acquisitions.

Clinical implementation of FFMCT would thus rely on prior imaging data to generate a patient model, which would be used as input to algorithms predicting noise projections required for noise reconstruction (Bartolac et al. 2011). The patient model could thus be established on the basis of prior diagnostic imaging studies or even using an atlas.

Dedes et al. (2017) proposed adapting FFMCT to proton computed tomography (pCT) scans acquired with pencil beam scanning (PBS) beamlines found in modern proton therapy facilities. While they could show the feasibility of fluence modulated pCT (FMpCT) in a simulation study, they relied on a "forward planning" approach where simple geometric considerations guided a binary fluence modulation on a pencil beam by pencil beam basis. The same approach was employed for the recent experimental realization of FMpCT using the proton tracking phase II pCT prototype of the Loma Linda University and University of California Santa Cruz (Dedes et al. 2018). Further developments in FMpCT thus require the modeling of the relation between proton fluence and pCT image noise to allow using an optimization strategy where pencil beam fluence could be continuously adjusted to achieve image noise prescriptions.

Preliminary work by Schulte et al. (2005) for the noise of the central pixel in a pCT image of a water cylinder, using proton projections binned at the rear tracker, laid the groundwork for noise reconstruction in pCT. However, Schulte et al. (2005) did not account for the impact of multiple Coulomb scattering (MCS) near object edges, and was published prior to the development of state of the art filtered backprojection (FBP) along most likely paths (MLP) (Rit et al. 2013), which makes use of distance driven binning (DDB) to create depth dependent projections for which rear tracker binning is a special case. As we will present in this paper, these effects have a non–negligible, non–trivial impact on two dimensional (2D) image noise in pCT.

The goal of this paper was thus to realize 2D noise reconstruction for simulated pCT scans of a water cylinder, assuming an ideal version of proton tracking pCT scanners, and accounting for the impact of MCS and the distance driven binning (DDB) which underpins FBP along MLPs. To do so, we extended the FBP along most likely paths to allow noise reconstruction, and made use of projection noise calculated on the basis of Monte Carlo (MC) simulations of ideal pCT scans, as well as from a dedicated analytical model.

6.2 Material and methods

6.2.1 MC simulation and geometry

In order to validate the noise reconstruction methods presented in the following sections, a MC simulation of a pCT scan of a water cylinder with a diameter of 25 cm was carried out, assuming ideal detectors (see figure 6.1). We chose a 260 mm × 50 mm rectangular proton field covering the whole diameter of the cylinder and 50 mm along the cylinder's axis. The fluence of the beam was chosen to be 200 protons/mm^2, all protons were launched perfectly parallel with random starting positions from the source plane, and the initial proton energy ($E_{\text{in}} = 250\,\text{MeV}$) was monoenergetic. The proton path was tracked on two parallel planes on the front and rear side of the water cylinder (see figure 6.1 for the details of the geometry), perpendicular to the incident beam, returning the initial and final position and momentum direction of each proton along with their exit energies.

The simulation platform was based on GEANT4 version 10.01.p02 (Agostinelli et al. 2003). The reference physics list `QGSP_BIC_HP` was used for the simulation of the interaction of particles with matter, which relies on `G4EmStandardPhysics` for electromagnetic interactions. MCS is modeled via the `G4WentzelIVIModel` (Ivanchenko et al. 2010). The tabulation of energy loss, range and inverse range, which are calculated during initialization, are done with 84 bins. More details on the energy loss are described in (GEANT-Collaboration et al. 2016).

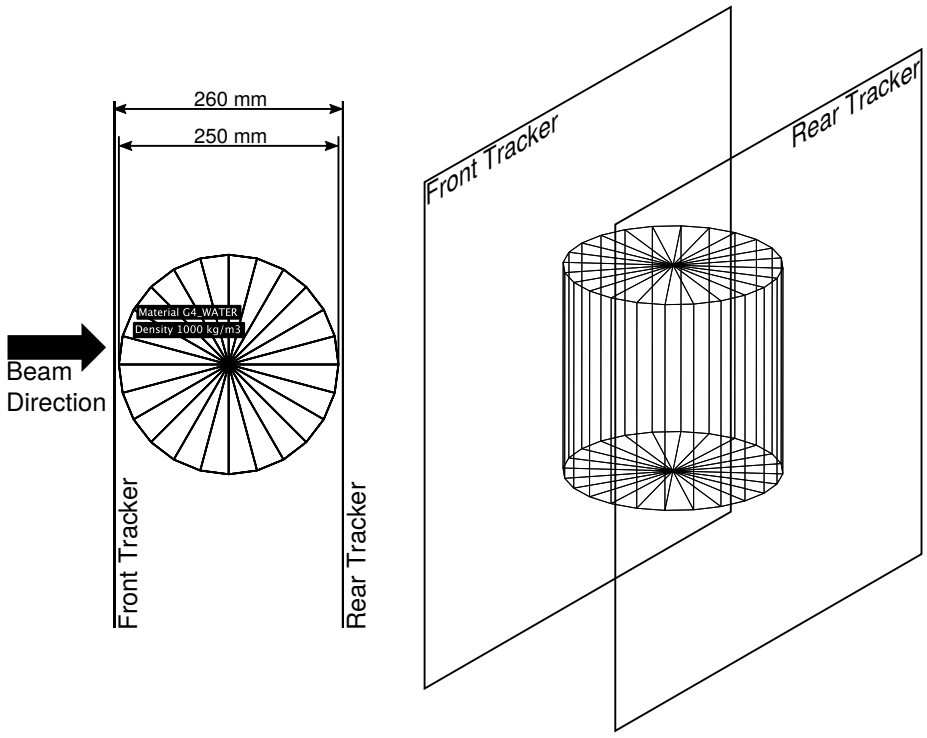

Figure 6.1: Geometry of the MC simulation.

6.2.2 Noise reconstruction formalism for pCT

Literature refers alternatively to noise images (one standard deviation, σ) or variance images (σ^2), with similar naming at the projection level. In this paper, we have opted to systematically employ the term noise reconstruction, which implies the trivial step of taking the square root of variance reconstructions. The noise reconstruction formalism presented below applies for pCT images reconstructed through distance driven binning (DDB), which was introduced by Rit et al. (2013). By doing so, one is able to include the influence of MCS, as protons traversing curved paths will be binned into different detector pixels at different binning depths. The fluence modulation approach, as proposed by Dedes et al. (2017), is based on parallel pencil beam irradiation. Therefore, we will solely discuss the parallel beam case in 2D slices.

In section 6.2.2, after a brief summary of the image reconstruction, we will review the quantification of noise in the pCT projections followed by the noise reconstruction of pCT binned at the rear tracker. The noise reconstruction including DDB is shown thereafter, given the noise projections binned at variable depth. We discuss the calculation of noise projections binned at the rear tracker and with DDB in section 6.2.3.

Image reconstruction

The coordinate system used in this paper is illustrated in figure 6.2. The FBP of an image slice $f(x, y)$, given the discrete projection values $p_{\gamma_n}(m\Delta\xi)$ acquired at discrete angles γ_n with a $\Delta\xi$ spacing on the one–dimensional projection grid using a discrete number of projections N_p, is given by

$$f(x, y) = \frac{\pi}{N_p} \sum_{n=1}^{N_p} h_{\gamma_n}(x \cos(\gamma_n) + y \sin(\gamma_n)), \tag{6.1}$$

where $h_\gamma(j\Delta\xi)$ are the convolved projections

$$h_{\gamma_n}(j\Delta\xi) = \Delta\xi \sum_{m=-D/2}^{D/2-1} p_{\gamma_n}(m\Delta\xi) g((j-m)\Delta\xi). \tag{6.2}$$

We chose the simplest convolution kernel from Ramachandran and Lakshminarayanan (Ramachandran & Lakshminarayanan 1971) (Ram–Lak), which results from band limiting the ramp kernel

$$g(j\Delta\xi) = \begin{cases} 1/(2\Delta\xi)^2 & \text{for } j = 0, \\ 0 & \text{for } j \text{ even } (j \neq 0), \\ -1 & \text{for } j \text{ odd.} \end{cases} \tag{6.3}$$

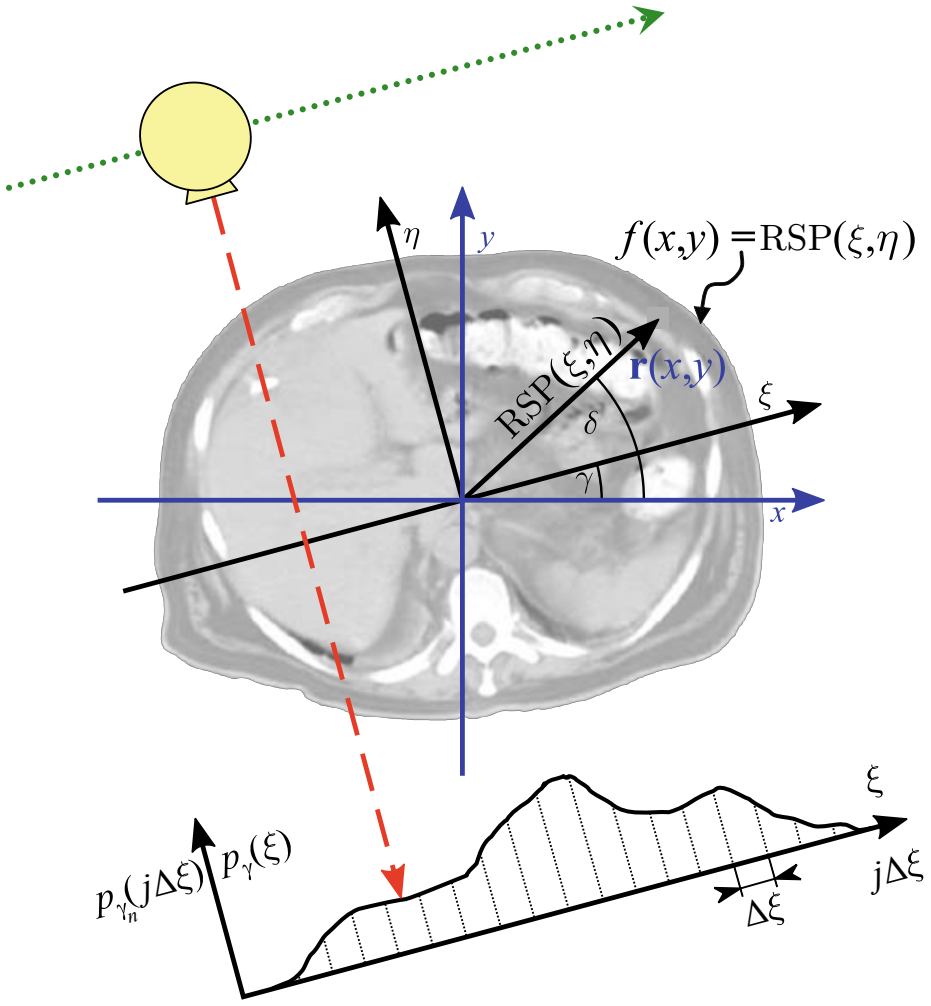

Figure 6.2: Setup of the acquisition and the coordinate systems. The source (yellow circle) moves along the dotted green line in order to generate one projection.

For a reconstruction using a given image pixel grid, with the pixel centers located at (x_p, y_p), the convolved projections of equation 6.1 require interpolation, as the sampled projection values do not necessarily coincide with the sample points $\xi_n(x_p, y_p) = x_p \cos(\gamma_n) + y_p \sin(\gamma_n)$. Interpolation reduces the noise and should be taken into account, when estimating noise in reconstructed images from noisy projection images (Huesman et al. 1977, Kak & Slaney 1988). For a linear interpolation between the two adjacent pixels j and $j+1$, the complete reconstruction from the FBP becomes

$$f(x_p, y_p) = \frac{\pi}{N_p}\Delta\xi \cdot \sum_{n=1}^{N_p} \sum_{m=-D/2}^{D/2-1} p_{\gamma_n}(m\Delta\xi) \left\{ g((j-m)\Delta\xi) \cdot [1-u] + g((j+1-m)\Delta\xi) \cdot u \right\}, \quad (6.4)$$

where both $j = j(\xi_n)$ and the weights $u = u(\xi_n)$ are determined by the location of the query point relative to the two adjacent projection sample values

$$u(\xi_n) = \frac{\xi_n - j\Delta\xi}{\Delta\xi}. \quad (6.5)$$

Statistical limitations of the acquisition

Proton tracking pCT reconstruction with FBP relies on binning individual protons into projection pixels. For regular FBP (i.e. non–DDB), this can be done by using the data from the rear or front trackers. For FBP along most likely paths based on DDB, the paths of individual protons are reconstructed and protons are binned into projections with variable η (see figure 6.2) (Rit et al. 2013).

After binning the protons into their respective pixels, one calculates their water equivalent path length (WEPL) through

$$\mathrm{WEPL}_i = \int_{E_{\mathrm{in}}}^{E^i_{\mathrm{out}}} \frac{\mathrm{d}E}{S_{\mathrm{W}}(E)}, \quad (6.6)$$

where $S_{\mathrm{W}}(E)$ is the stopping power of water and i refers to individual measured protons with energy E_{in} before the object and measured energy E^i_{out} beyond. Then one estimates the mean to obtain the projection value

$$p_{\gamma_n}(j\Delta\xi) = \frac{1}{N_{\gamma_n}(j\Delta\xi)} \sum_{i=1}^{N_{\gamma_n}(j\Delta\xi)} \mathrm{WEPL}_i, \quad (6.7)$$

where $N_{\gamma_n}(j\Delta\xi)$ is the number of protons in pixel $j\Delta\xi$ at the projection angle γ_n. At each pixel, the mean carries an intrinsic uncertainty in itself, typically expressed as the standard deviation of the mean. The variance of equation 6.7 is then

$$\sigma^2_{\gamma_n}(j\Delta\xi) = \frac{\sigma^2_{\mathrm{WEPL},\gamma_n}(j\Delta\xi)}{N_{\gamma_n}(j\Delta\xi)}. \tag{6.8}$$

The variance of the WEPL (without index i since we refer to average WEPL in a projection pixel), σ^2_{WEPL}, in turn depends on the uncertainty of the proton energies, which is generally attributed to energy straggling (Schulte et al. 2005) (additional detector uncertainties will not be taken into account in this study). Therefore, the error of the exit energy propagates into the WEPL values, which is described by the error propagation formula. The first order approximation is sufficient as the second order contribution is already four orders of magnitude below the first order term. With the mean energy of the detected protons $\overline{E}_{\mathrm{out}}$, one obtains

$$\sigma^2_{\mathrm{WEPL}} = \left(\frac{\partial \mathrm{WEPL}\left(\overline{E}_{\mathrm{out}}\right)}{\partial E}\right)^2 \cdot \sigma^2_{E_{\mathrm{out}}} = \frac{\sigma^2_{E_{\mathrm{out}}}}{S^2_{\mathrm{W}}(\overline{E}_{\mathrm{out}})}. \tag{6.9}$$

Together with equation 6.8, the variance of the projection value is given by Schulte et al. (2005)

$$\sigma^2_{\gamma_n}(j\Delta\xi) = \frac{\sigma^2_{E_{\mathrm{out}},\gamma_n}(j\Delta\xi)}{N_{\gamma_n}(j\Delta\xi) \cdot S^2_{\mathrm{W}}(\overline{E}_{\mathrm{out},\gamma_n}(j\Delta\xi))}. \tag{6.10}$$

2D noise reconstruction without DDB

The basics of the noise reconstruction from the FBP for pCT were outlined by Schulte et al. (2005) for the central pixel of pCT images, and are analogous to the x–ray CT noise reconstruction techniques shown by Huesman et al. (1977) or Gore & Tofts (1978). Since the projection values $p_{\gamma_n}(m\Delta\xi)$ carry an error, we will treat them as random variables, with their mean and variance given by equations 6.7 and 6.10 respectively. In general, the variance of a weighted sum of random variables X_i with the weights a_i is

$$\mathrm{Var}\left[\sum_{i=1}^{M} a_i X_i\right] = \sum_{i,j=1}^{M} a_i a_j \mathrm{Cov}\left[X_i, X_j\right] = \sum_{i=1}^{M} a_i^2 \mathrm{Var}[X_i] + 2\sum_{i,j|i<j}^{M} a_i a_j \mathrm{Cov}\left[X_i, X_j\right]. \tag{6.11}$$

The summation over $p_{\gamma_n}(m\Delta\xi)$ in equation 6.4 is threefold: the sum over the angles, the projection values (convolution) and the interpolation. We use the approximation that there is no covariance among the projection values $p_{\gamma_n}(m\Delta\xi)$ since individual protons are tracked and pileup is assumed negligible.

But, due to the convolution, the filtered projections carry a mutual dependency. Each filtered projection $h_{\gamma_n}(j\Delta\xi)$ value is the linear combination of the surrounding projection values $p_{\gamma_n}(m\Delta\xi)$. As the projection values $p_{\gamma_n}(m\Delta\xi)$ are independent, we have

$$\mathrm{Cov}\left[p_{\gamma_n}(m\Delta\xi), p_{\gamma_{n'}}(m'\Delta\xi)\right] = \delta_{n,n'}\delta_{m,m'}\mathrm{Var}\left[p_{\gamma_n}(m\Delta\xi)\right] = \delta_{n,n'}\delta_{m,m'}\sigma^2_{\gamma_n}(m\Delta\xi), \quad (6.12)$$

since $\mathrm{Cov}[X,X] = \mathrm{Var}[X]$. δ_{ij} is the Kronecker delta, which is defined as

$$\delta_{ij} = \begin{cases} 0 & \text{if } i \neq j, \\ 1 & \text{if } i = j. \end{cases} \quad (6.13)$$

The weights from equation 6.11 become

$$\begin{aligned} a_i &\to \frac{\pi}{N_p}\Delta\xi \cdot g((j(\xi_n) - m)\Delta\xi)\left[1 - u(\xi_n)\right] \\ a_j &\to \frac{\pi}{N_p}\Delta\xi \cdot g((j(\xi_n) + 1 - m)\Delta\xi)u(\xi_n). \end{aligned} \quad (6.14)$$

Finally, the noise of the image can be reconstructed through

$$\begin{aligned} \mathrm{Var}\left[f(x_p, y_p)\right] = &\left(\frac{\pi}{N_p}\Delta\xi\right)^2 \cdot \\ &\sum_{n=1}^{N_p}\left\{\left[1-u\right]^2 V_{\gamma_n}(j\Delta\xi) + 2\left[1-u\right]uC_{\gamma_n}(j\Delta\xi, (j+1)\Delta\xi) + u^2V_{\gamma_n}((j+1)\Delta\xi)\right\} \end{aligned} \quad (6.15)$$

where $j = j(\xi_n)$ and $u = u(\xi_n)$, just as in equation 6.4. Following (Wunderlich & Noo 2008), we introduced the variance and covariance terms

$$V_{\gamma_n}(j\Delta\xi) = \sum_{m=-D/2}^{D/2-1} g^2((j-m)\Delta\xi)\sigma^2_{\gamma_n}(m\Delta\xi), \quad (6.16)$$

$$C_{\gamma_n}(j\Delta\xi, j'\Delta\xi) = \sum_{m=-D/2}^{D/2-1} g((j-m)\Delta\xi)g((j'-m)\Delta\xi)\sigma^2_{\gamma_n}(m\Delta\xi). \quad (6.17)$$

When summing $p_{\gamma_n}(m\Delta\xi)g((j(\xi_n)-m)\Delta\xi)\left[1-u(\xi_n)\right]$ and $p_{\gamma_n}(m\Delta\xi)g((j(\xi_n)+1-m)\Delta\xi)u(\xi_n)$ for interpolation, we produce two variance and two covariance terms through $\mathrm{Var}[a_1X + a_2X] = a_1^2\mathrm{Var}[X] + a_2^2\mathrm{Var}[X] + 2a_1a_2\mathrm{Cov}[X,X]$. The C_{γ_n} term is the covariance of the filtered projections and only the inner two sums of equation 6.4 bear this covariance, as we do not convolve in the angular dimension.

Wunderlich & Noo (2008) noticed that by defining

$$g_C(j\Delta\xi) = g(j\Delta\xi)g((j+1)\Delta\xi), \quad (6.18)$$

equation 6.17 can be written as a convolution

$$C_{\gamma_n}(j\Delta\xi,(j+1)\Delta\xi) = \sum_{m=-D/2}^{D/2-1} g_{\mathrm{C}}((j-m)\Delta\xi)\sigma^2_{\gamma_n}(m\Delta\xi). \tag{6.19}$$

In general, the noise reconstruction algorithm is similar to a FBP. We merely use a different prefactor, interpolation and different convolution kernels.

Furthermore, one is able to approximate the effect of the interpolation and reduce it to a single factor. A simplified variance reconstruction is then given by

$$\mathrm{Var}\left[f(x_p,y_p)\right] = f_{\mathrm{interp},\mu}\left(\frac{\pi}{N_p}\Delta\xi\right)^2 \sum_{n=1}^{N_p} V_{\gamma_n}(j\Delta\xi), \tag{6.20}$$

where $f_{\mathrm{interp},\mu} = 2/3 - 2/\pi^2$. In quantitative terms, the linear interpolation in combination with the Ram–Lak filter reduces the standard deviation by about 32% (more precisely: $\sqrt{2/3-2/\pi^2} \approx 0.681193$). See appendix A.0.1 for the detailed derivation of this approximation. For our 2D noise reconstruction, including the 2D noise reconstruction using DDB, we utilize this simplification.

2D noise reconstruction including DDB

Given the projections from a single binning depth (e.g. the rear tracker), we had to use one–dimensional interpolation between the sampled (and convolved) data points (see equation 6.4). In order to take the projections from different depths into account, a two–dimensional interpolation is necessary. However, as the projections from two neighboring depths are hardly any different, the interpolation along η has a negligible contribution to the variance reconstruction, if the spacing $\Delta\eta$ is sufficiently small. For the reconstruction of a 2D slice, the Radon space becomes now three–dimensional (γ,ξ,η) through the additional dimension in the η–direction: $p_{\gamma_n}(j\Delta\xi) \to p_{\gamma_n}(j\Delta\xi, k\Delta\eta)$. The DDB noise reconstruction becomes then

$$\begin{aligned}\mathrm{Var}\left[f\left(x_p,y_p\right)\right] = &\left(\frac{\pi}{N_p}\Delta\xi\right)^2 \cdot \\ &\sum_{n=1}^{N_p}\Big\{\left[1-u\right]^2 V_{\gamma_n}(j\Delta\xi,k\Delta\eta) + 2\cdot\left[1-u\right] u C_{\gamma_n}(j\Delta\xi,(j+1)\Delta\xi,k\Delta\eta) \\ &\qquad + u^2 V_{\gamma_n}((j+1)\Delta\xi,k\Delta\eta)\Big\},\end{aligned} \tag{6.21}$$

where $k\Delta\eta$ is closest to the corresponding binning depth (nearest neighbor interpolation). A more detailed discussion can be found in appendix A.0.2. The additional simplification involving the interpolation factor $f_{\mathrm{interp},\mu}$ (see equation 6.20) can also be applied here.

6.2.3 Noise of the projections

Recall that for 2D noise reconstruction, we need to know the variance of all the projections binned at different depths, which consists of the variance of the energy, the number of protons within the pixels and the stopping power, evaluated at $\overline{E}_{\text{out}}$

$$\sigma^2_{p_{\gamma_n}}(j\Delta\xi, k\Delta\eta) = \frac{\sigma^2_{E_{\text{out}},\gamma_n}(j\Delta\xi, k\Delta\eta)}{N_{\gamma_n}(j\Delta\xi, k\Delta\eta) \cdot S^2_{\text{W}}(\overline{E}_{\text{out},\gamma_n}(j\Delta\xi, k\Delta\eta))}. \tag{6.22}$$

The latter is certainly the easiest to calculate, as we can use the $\overline{E}_{\text{out}}$–values straight from the scan and evaluate the stopping power of water at $\overline{E}_{\text{out}}$. The remaining two components of the variance require more detailed discussions.

In this section, we first show an analytical approach to calculate and explain the energy straggling and proton counts for *rear tracker projections* using theoretical proton energy straggling and scattering models, commonly used in the pCT reconstruction. Results from this model will be compared to the results of the MC simulation.

Since the extension of the analytical model to arbitrary distances for DDB is non-trivial, we subsequently report how the noise of *DDB projections* was calculated from the MC simulation data.

These calculations of the noise in pCT projections are an extension of Schulte's (Schulte et al. 2005) work detailing the noise at the center of a cylindrical object, which will be referred to as *central pixel model*.

Noise of rear tracker projections

Proton counts

For the 2D noise reconstruction, the proton counts N could be taken directly from the MC simulation or scan data. However, for the proton fluence used in this work (200 protons/mm^2), the statistical fluctuation of the proton counts at the rear detector is large. This fluence corresponds to an imaging dose of about 3 mGy (Schulte et al. 2005), which is already relatively high in the context of daily image guidance with pCT.

Accurate and smooth proton count data can be calculated through the transport theory, i.e. Fermi–Eyges theory–(Gottschalk 2012, Fermi 1940, Eyges 1948). It is a bivariate Gaussian theory, which is able to predict proton MCS with sufficient accuracy. More complete models, e.g. Molière's theory (Molière 1947, Molière 1948) are not necessary, as the additional tails of the distributions, as predicted by Molière's theory, will be subject to the 3 standard deviations data cuts, i.e. the rejection of protons which have undergone large angle scattering or nuclear interactions (Schulte et al. 2005). $\mathcal{F}(\xi, \theta, \eta)\mathrm{d}\xi\mathrm{d}\theta$ is the probability to find a proton within the lateral displacement $[\xi, \xi + \mathrm{d}\xi]$ and traveling along the angle $[\theta, \theta + \mathrm{d}\theta]$ at depth η, which was initially at $\xi_0 = 0$ and $\theta_0 = 0$ at depth $\eta_0 = 0$

$$\mathcal{F}(\xi, \theta, \eta)\mathrm{d}\xi\mathrm{d}\theta = \frac{1}{2\pi\sqrt{B(\eta)}} \exp\left[-\frac{1}{2}\frac{A_0(\eta)\xi^2 - 2A_1(\eta)\xi\theta + A_2(\eta)\theta^2}{B(\eta)}\right] \mathrm{d}\xi\mathrm{d}\theta, \tag{6.23}$$

where

$$B(\eta) = A_0(\eta)A_2(\eta) - A_1^2(\eta). \tag{6.24}$$

For the scattering integrals

$$A_n(\eta) = \int_0^\eta (\eta - x)^2 \, T(x)\mathrm{d}x \tag{6.25}$$

we chose (as it is chosen in other pCT related work e.g. (Quiñones et al. 2016) or (Bopp 2014)) the scattering power proposed by Gottschalk (2010), which reads

$$T_{\mathrm{dM}} = f_{\mathrm{dM}}(p, v, p_1, v_1) \cdot \frac{1}{X_s} \left(\frac{E_s}{pv}\right)^2, \tag{6.26}$$

where $E_s = 15.0\,\mathrm{MeV}$, X_s is the material dependent scattering length and

$$\begin{aligned} f_{\mathrm{dM}} \equiv 0.5244 + 0.1975 \log_{10}\left[1 - \left(\frac{pv}{p_1 v_1}\right)^2\right] + 0.2320 \log_{10}\left[\frac{pv}{\mathrm{MeV}}\right] \\ -0.0098 \log_{10}\left[\frac{pv}{\mathrm{MeV}}\right] \log_{10}\left[1 - \left(\frac{pv}{p_1 v_1}\right)^2\right]. \end{aligned} \tag{6.27}$$

In order to carry out these integrals, we used the analytical expression of the cylindrical hull, but a prior reconstruction could also be used in the case of patient imaging. Our primary interest is the spatial distribution of the protons, thus we calculate $A_2(\eta)$, since

$$\langle \xi^2(\eta) \rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \xi^2 \mathcal{F}(\xi, \theta, \eta) \mathrm{d}\xi \mathrm{d}\theta = A_2(\eta). \tag{6.28}$$

The additional proton drift from the object edge (hull) to the tracker can be calculated using a quadratic law under the assumption that the scattering power of air is negligible (Gottschalk 2012)

$$A_2(D(\xi) + d(\xi)) = A_0(D(\xi))d^2(\xi) + 2A_1(D(\xi))d(\xi) + A_2(D(\xi)). \tag{6.29}$$

See figure 6.3 for the definition of $d(\xi)$ and $D(\xi)$.

We can apply this theory to determine the rear tracker fluence by calculating the width of the proton beam for any ξ. Then we superimpose the distributions, weighted by their corresponding attenuation caused by nuclear reactions, as already described by Schulte et al. (2005). The exponential attenuation of the initial fluence Φ_0 is given by

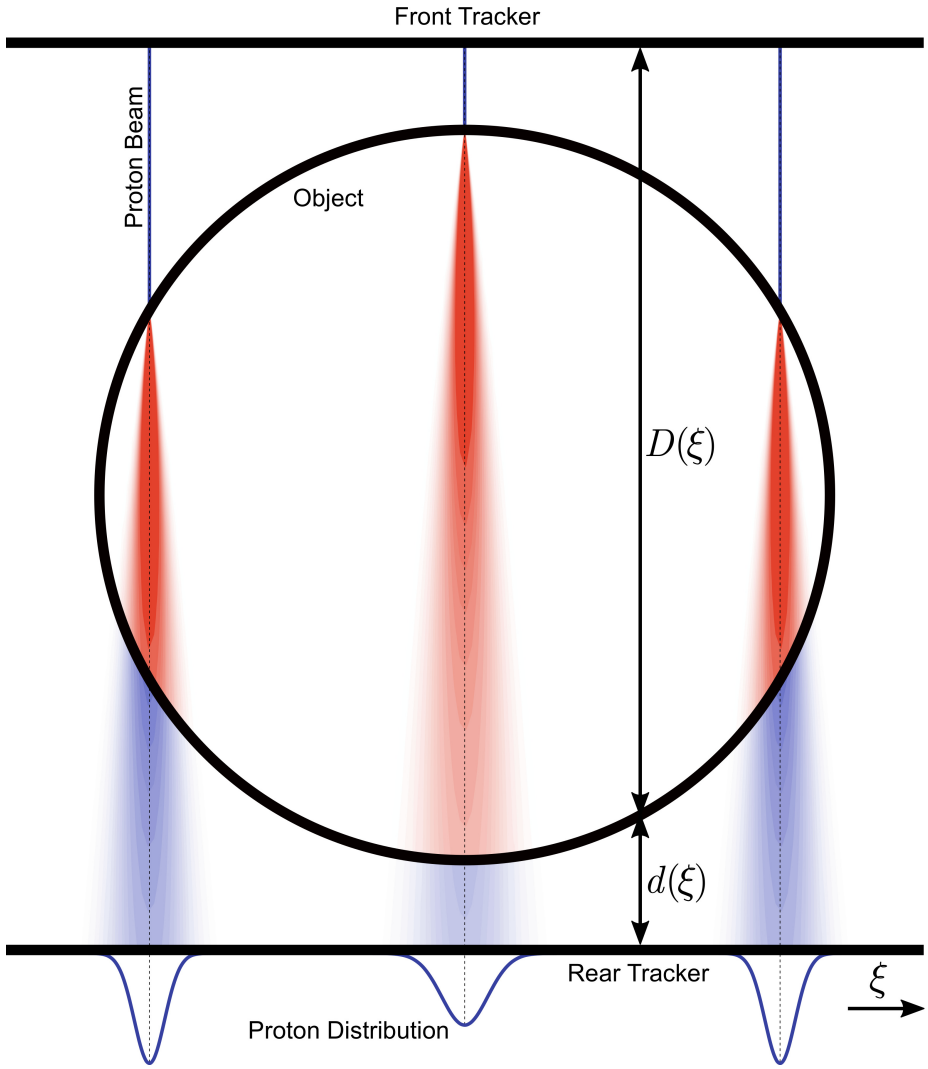

Figure 6.3: Transport theory applied to pCT. For the red colored proton beam, we have to evaluate the scattering integrals of equation 6.25, while for the subsequent blue colored drift through air, the quadratic law of equation 6.29 is sufficient. Prior to entering the object, the width of proton beam is considered infinitely small, i.e. we consider a parallel beam. Note that the scattering is exaggerated.

$$\Phi(\mathrm{WEPL}) = \Phi_0 \cdot e^{-N\sigma_{\mathrm{nuc}}\cdot\mathrm{WEPL}} = \Phi_0 \cdot e^{-\kappa\cdot\mathrm{WEPL}}, \tag{6.30}$$

where N is the target (nuclei) density, σ_{nuc} is the nuclear cross section, and $\kappa = N\sigma_{\mathrm{nuc}}$ is the linear attenuation coefficient.

The attenuation coefficient can be determined by taking elastic ($\approx 80\,\mathrm{mb}$) and inelastic ($\approx 270\,\mathrm{mb}$) cross sections into account (values taken from (Quiñones et al. 2016) → figure 3.9 and → figure 3.12 for "G4_O" above approximately 150 MeV; Schulte et al. (2005)

determined the attenuation coefficient, neglecting the elastic contribution, in a similar fashion) which results in an attenuation coefficient $\kappa = 0.0131\,\mathrm{cm}^{-1}$. The normalized fluence at each detector pixel is the sum of all beams that scatter into a given pixel.

Standard deviation of the exit energy

The determination of the standard deviation of the exit energy $\sigma_{E_{\mathrm{out}}}$ at each detector pixel is a somewhat more challenging task. Schulte et al. (2005) suggested to calculate it from the exit energy $\overline{E}_{\mathrm{out}}$, or the WEPL value, in combination with an evaluation of Payne's (Payne 1969) or Tschalär's (Tschalär 1968a, Tschalär 1968b) theories, which establish a connection between the exit energy or the WEPL to the energy straggling. However, if we want to perform 2D noise reconstruction, then Schulte's approach is not valid away from the central pixel due to the interplay of MCS and the high gradient of the object's hull along ξ. In the following, we will present an analytical approach, much like (Schulte et al. 2005), which includes Tschalär's/Payne's theoretical energy straggling and also accounts for the effect of MCS. Given the proton transport and thus $A_2(\xi)$ for every exit detector pixel, that we used to determine the proton counts earlier in section 6.2.3, we can answer the inverse question as well: given some exit detector pixel $j\Delta\xi$, what is the distribution of initial proton positions (or initial position distribution, short IPD) on the front tracker, that scatter into $j\Delta\xi$. This process is demonstrated in figure 6.4. We take the distributions of the surrounding entrance pixels of $j\Delta\xi$ and calculate how much they contribute to the exit pixel sited at $j\Delta\xi$. Additionally, we weight the result with the attenuation.

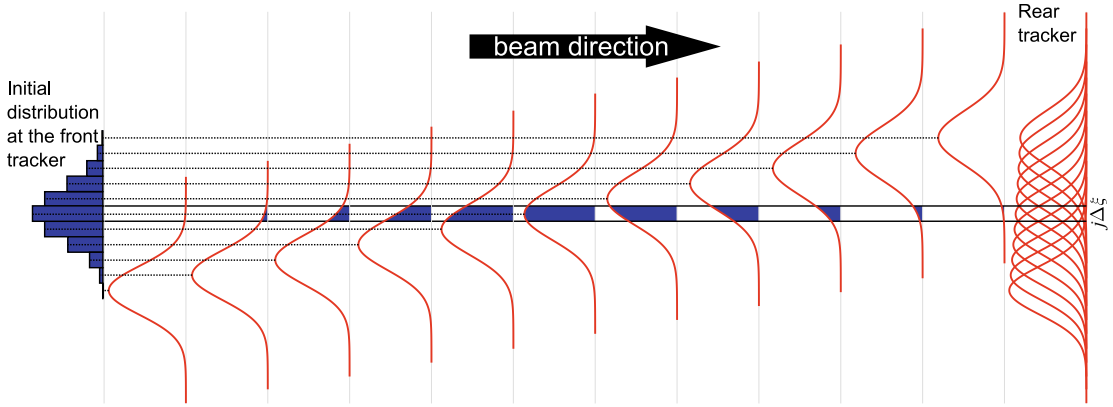

Figure 6.4: Visualization of the calculation of the initial distribution from the given proton distributions at the rear tracker. The proton distributions at the detector were spread horizontally for visual clarity. This has no geometrical meaning. We calculate the initial distribution of the pixel $j\Delta\xi$ delimited by the solid lines. The probabilities for the surrounding positions to scatter into $j\Delta\xi$ is given by the areal overlap of their distributions within $j\Delta\xi$. The corresponding probabilities are then arranged in the initial distribution along the dotted lines.

We include the effect of the MCS through the IPD. Since protons from different initial positions (IP) scatter into the same pixel, they must traverse different path lengths, i.e.

different parts of the objects. Therefore they lose different amounts of energy, which eventually broadens the energy spectrum. We used a straight line approximation between the entry point and the detector pixel coordinate, which may seem a poor approximation for a 25 cm diameter object. However, notice that the broadest IPD of the detector pixel at the center covers approximately only two centimeters.

Our goal is to calculate the distribution of *mean* exit energies that are collected within each detector pixel. This can be done by mapping the IPD with some function $F(j\Delta\xi, \mathrm{IP})$ to the corresponding distribution of energy losses. This function in turn can be calculated through sinogram interpolation, taken from a prior scan in combination with the straight line approximation. See appendix A.0.3 for details.

Now we transform the IPDs into distributions of $\overline{E}_{\mathrm{out}}$. The transformation is given by

$$p_\mu(j\Delta\xi, \mu_n) = \sum_{x \in F^{-1}(j\Delta\xi, \mu_n)} f_{\mathrm{IP}}(x). \tag{6.31}$$

In general, the IPDs are closely distributed around their corresponding exit detector pixel. Despite the fact that the IPDs are the broadest at the center, the transformed distributions of mean energy losses will more closely resemble a delta distribution. Put simply, no matter where the protons that scatter into the central pixel enter the object, on average they have lost approximately the same amount of energy. On the other hand, at the object edges the energy transformation varies more rapidly. Even though the IPDs become increasingly narrow at the edges, the corresponding $\overline{E}_{\mathrm{out}}$ distribution might be broader, if the traversed thickness decreases rapidly, which is the case with the 25 cm cylinder we used. In other words, only small changes of the IPD cause large changes in the average energy loss. This is due to the more rapidly changing hull and therefore more rapidly changing path lengths.

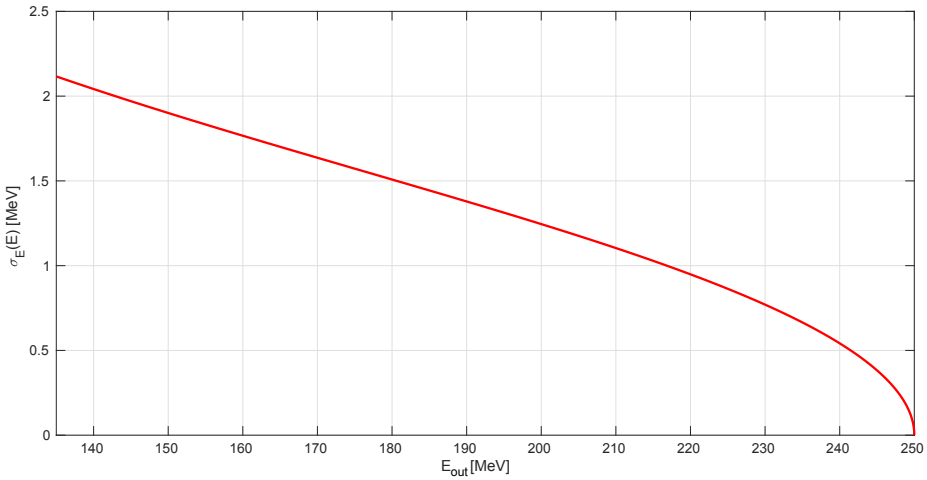

Figure 6.5: Energy straggling determined via solving the integral given in equation 6.34 for protons with an initial energy of $E_{\mathrm{in}} = 250$ MeV.

Finally, at this stage we will apply the theoretical energy straggling (see figure 6.5), which is governed by the differential equation of Tschalär (Tschalär 1968a, Tschalär 1968b), here expanded up to the first order

$$\frac{\mathrm{d}\sigma_E^2(x)}{\mathrm{d}x} = \chi_2(E(x)) - 2\frac{\partial\chi_1(E(x))}{\partial E}\sigma_E^2(x), \tag{6.32}$$

where χ_1 is given by the stopping power and χ_2 is the straggling parameter

$$\begin{aligned}\chi_1(E) &= K_1\frac{1}{\beta^2(E)}\left[\ln\left(\frac{2mc^2\beta^2(E)}{I(1-\beta^2(E))}\right) - \beta^2(E)\right]\\ \chi_2(E) &= K_2\frac{1-\beta^2(E)/2}{1-\beta^2(E)}.\end{aligned} \tag{6.33}$$

For protons stopping in water, we have $K_1 \approx 170\,\mathrm{keV/cm}$ and $K_2 \approx 0.087\,\mathrm{MeV^2/cm}$. Equation 6.32 can be solved analytically, as outlined by Payne (1969):

$$\sigma_E^2(E) = \chi_1^2(E)\int_E^{E_{\mathrm{in}}}\frac{\chi_2(E')}{\chi_1^3(E')}\mathrm{d}E'. \tag{6.34}$$

The last step involves the calculation of the standard deviation of the energy σ_E from the distribution of mean energy losses $\overline{E}_{\mathrm{out}}$. Each detector pixel collects protons that traversed different material thicknesses with different intensities (or normalized probabilities p_μ). Thus the energy distributions at the detector pixels consist of a superposition of the individual energy spectra with their respective means μ_n and standard deviations σ_n. We can retrieve the detector distribution from the distribution of means by convolving the means distributions with the normal distribution from the energy straggling theory. Note that this is not shift–invariant convolution, as the Gaussian convolution kernel is energy dependent.

If $\mu_{\mathrm{Det}}(j\Delta\xi)$ is the mean energy loss at each exit detector pixel $(j\Delta\xi)$, i.e. the projection value prior to the transformation into WEPL, then the variance of the energy is given by

$$\sigma_E^2(j\Delta\xi) = \sum_n p_\mu(j\Delta\xi,\mu_n)\left(\sigma_n^2 + (\mu_n - \mu_{\mathrm{Det}}(j\Delta\xi))^2\right). \tag{6.35}$$

As the distribution is particularly narrow at the center of the water cylinder (all protons traversed a similar length, the diameter), the only contribution to the standard deviation comes from a single mean energy loss, which is equivalent to calculating the standard deviation solely from the straight line from the front to the rear tracker. In general, there are three opposing effects at each detector pixel, which influence each other. Firstly, the IPD, which is broad at the center and narrow at the edges. Secondly, the mean energy loss distribution, which is broad at the edges and narrow at the center. Finally, the convolution of p_μ with Gaussians according to Tschalär's theory, which has a big influence at the center (since it predicts broader Gaussian distributions at large energy losses) and a small effect at the edges, as the additional effect of energy straggling is small at small energy losses.

Noise of DDB projections

For the calculation of the noise of DDB projections at depth η we made use of the path of every proton for which entrance and exit coordinates were recorded by the MC simulation.

The paths were reconstructed by cubic splines with adjusted velocity boundary conditions, similar to (Fekete et al. 2015), using the position and momentum direction information from the tracker planes. Li et al. (2006) showed that the use of regular cubic splines has little effect on the spatial resolution in the pCT reconstruction, and the improved formulation by Fekete et al. (2015) provides paths which are nearly congruent with the path determined by the original MLP formulation (Schulte et al. 2008).

In section 6.2.3, we refer to issues with low count statistics when generating projections from the proton fluence used in our MC simulations. To circumvent this issue, a high–statistics MC dataset was generated by combining all simulated projections. This smoother dataset was used for the DDB 2D noise reconstruction by exploiting the rotational symmetry of our water cylinder, after scaling back the counts N to the original fluence.

Prior to binning the data into projections, protons were selected with a 3 standard deviations cut on the energy and angular distributions around their mean energy and angle per projection pixel. This was done based on front tracker binning, as in the implementation of Rit et al. (2013)

Thus for a given η, the proton tracks crossing 1 mm bins were used to calculate $\sigma^2_{p_{\gamma n}}$ using equation 6.22 where $\sigma^2_{E_{\text{out}},\gamma_n}$ was obtained from Gaussian fitting of the E_{out} distribution of the binned protons. N_{γ_n} was simply the number of proton paths crossing the bin and S^2_{W} was evaluated at $\overline{E}_{\text{out}}$.

6.2.4 RSP image reconstruction and noise quantification

RSP images were reconstructed for this study with an implementation of DDB FBP, using the formalism of section 6.2.2. The main principles of the algorithm are presented in (Rit et al. 2013). The path of every proton was obtained from the splines described in the previous section 6.2.3. The data cuts of section 6.2.3 were used.

The validation of the 2D noise reconstructions was performed against the noise calculated from RSP images reconstructed from the MC simulation data. Utilizing the radial symmetry of the water cylindrical phantom, annular regions of interest (ROIs) with varying radii were defined. The number of pixels in each ROI was fixed to 1000 to ensure statistical accuracy, with the radial thickness varying accordingly. The noise from the MC RSP image at a given radius was defined as the standard deviation of the distribution of RSP values within a ROI. The standard deviation was calculated from a Gaussian fit in each RSP distribution. For the central pixel and improved models, the noise determination as a function of the distance from the center of the object was calculated by means of a line profile across a diameter on the 2D noise reconstruction. The pixel grid used for all image reconstructions shown was 280 mm × 280 mm with 1 mm × 1 mm pixels.

Ideal WEPL projections for parallel rays, calculated analytically for the water cylinder,

were discretized on the same grid as the rear–tracker or DDB binned projections. These were used to reconstruct RSP images as described above. We used these images to evaluate the impact of partial volume effects (for example at the object's edge) and reconstruction from discretized projections on the standard deviation calculated with the annular ROIs. This was done by calculating the standard deviation analytically instead of using Gaussian fits.

6.3 Results

Equation 6.22 gives σ^2_{WEPL} of a projection as a function of S_{W}, $\sigma_{E_{\mathrm{out}}}$ and N within the pixels. In figure 6.6, each of the aforementioned components is shown along the lateral coordinate, for the MC data, the central pixel model and the improved model taking into account the effect of MCS. The three curves for S_{W} and N were nearly indistinguishable, while for $\sigma_{E_{\mathrm{out}}}$ and σ^2_{WEPL} good agreement between the MC data and the improved rear tracker model accounting for MCS was observed. The largest σ_{WEPL} error between the MC data and the improved model was about 8% at $\xi = 100\,\mathrm{mm}$. The central pixel model, which ignores MCS, failed to correctly predict $\sigma_{E_{\mathrm{out}}}$ and σ^2_{WEPL} away from the object's center.

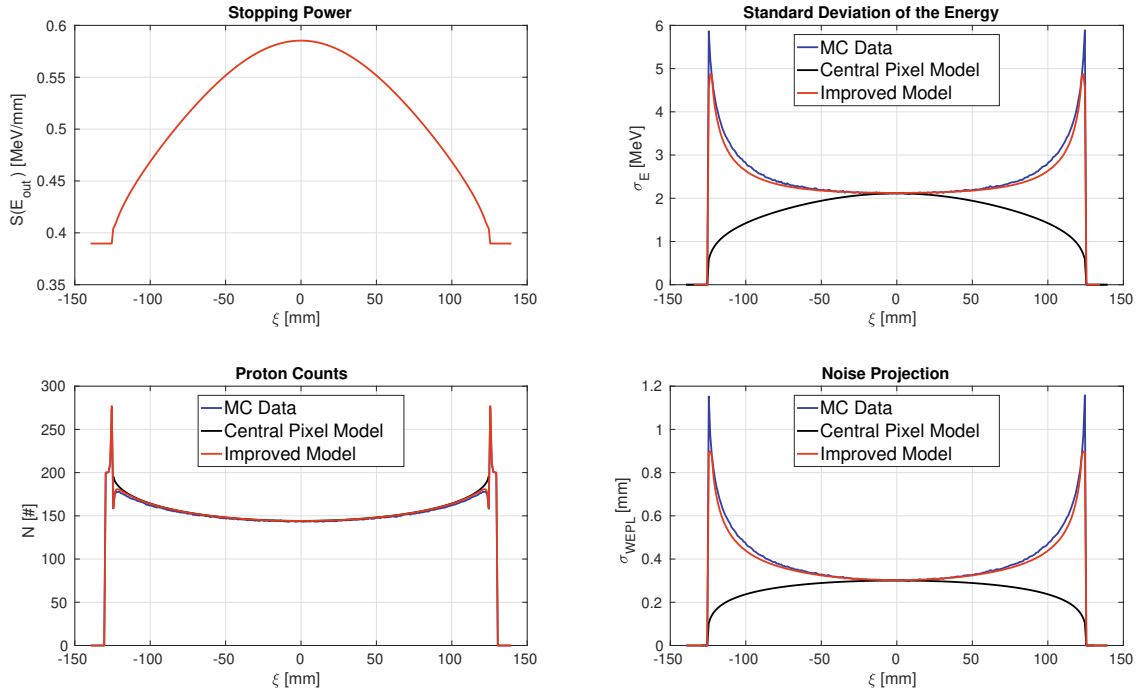

Figure 6.6: Results from rear tracker binning ($\eta = 260\,\mathrm{mm}$) for a single projection as profiles along the lateral coordinate ξ for the components of equation 6.22 for the central pixel model, the improved model and MC data. The stopping power evaluated at the mean exit energy (upper left) is shown only for the MC data, as the three curves overlapped. Data from high statistics MC simulations were used.

Figure 6.7 shows the effect of the binning location on σ_{WEPL}. The distance is measured from the front tracker ($\eta = 0$). Binning at the rear tracker ($\eta = 260\,\mathrm{mm}$) results in high noise at the edges of the object (equivalent to the MC data σ_{WEPL} of figure 6.6). We observed that the increase of σ_{WEPL} with ξ approaching the object's edge was most pronounced at the rear tracker, and that this effect gradually disappeared as η approached 0 near the front tracker.

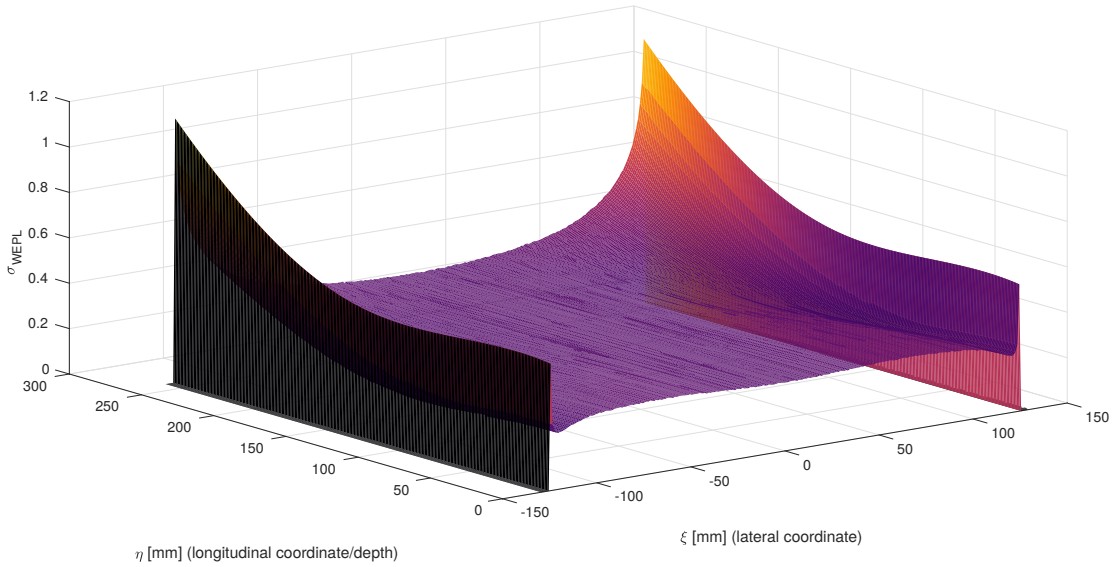

Figure 6.7: Noise projection profiles along the lateral coordinate ξ, shown as a function of the longitudinal coordinate of the binning position η. $\eta = 0$ corresponds to a binning at the front tracker, while $\eta = 260\,\mathrm{mm}$ at the rear tracker. Data from high statistics MC simulations were used.

Figure 6.8 presents a 2D noise reconstruction obtained using either noise projections obtained by binning the protons at the rear tracker (equation 6.15) or with DDB (equation 6.21). The effect on the noise image of the "interference" between the 2D image pixel grid and the 1D projection grid, as well as that of using a constant term for the linear interpolation as explained in 6.2.2, are shown. Generally, with rear tracker binning, the noise increased towards the object's edge, while for DDB it appeared constant with a slight decrease at the edge. High noise was observed at the object's boundary.

Finally, in figure 6.9, profiles through the 2D noise reconstructions based on the central pixel model, the improved model for rear tracker binning and DDB (in this case direct use of MC data, see section 6.2.3) are compared to that obtained from the MC reconstructed RSP image (using annular ROIs), as a function of the radius from the center of the object. For indicative purposes, the standard deviation for the RSP image from discretized ideal projections is also shown. We observed that the improved model and MC data–based DDB accurately reproduced the behavior of the noise observed in the reconstructed RSP images. When using rear tracker binning, an increase of 60% in image noise was observed at the edge of the 12.5 cm radius object when compared to its center. This effect was poorly

captured by the central pixel model, which underestimated noise by up to 40% in this case. Interestingly, DDB negated the radial noise increase observed with rear tracker binning, producing generally lower noise values which decreased less than 5% with radius. The ideal projections yielded large standard deviations at the object's edge which corresponded to the spikes observed in the noise from the annular ROIs.

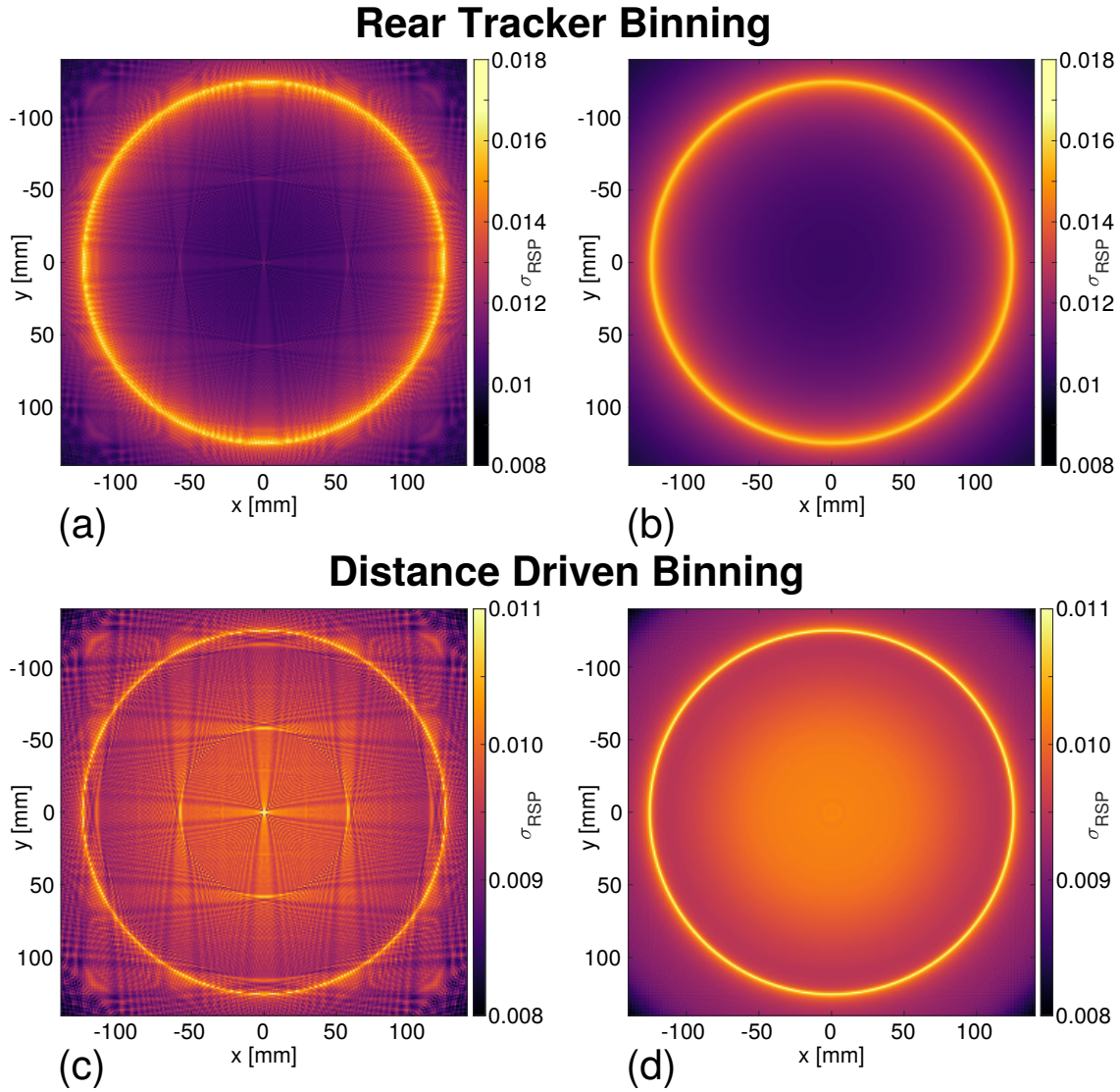

Figure 6.8: Noise images reconstructed without DDB (a,b) and with DDB (c,d). The left column (a,c) shows the effect of 'interference' between the 2D image pixel grid and the 1D projection grid. The right column (b,d) shows the smoothed versions reconstructed by using a constant term for the linear interpolation. The results from rear tracker binning were obtained from the improved model while for the DDB directly from high statistics MC data. Notice the different scales.

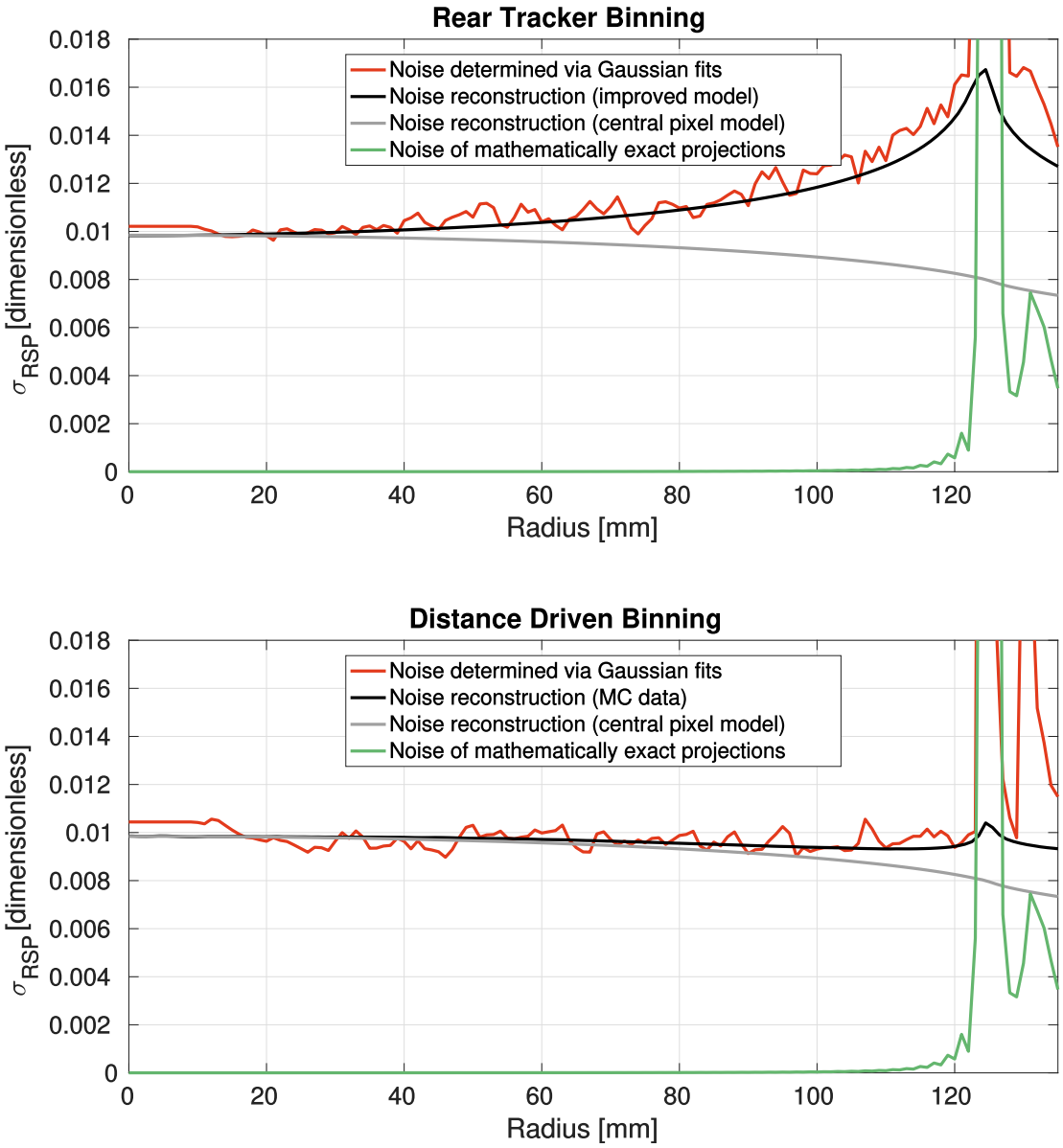

Figure 6.9: Noise profile comparison, as a function of the radius, between MC, the central pixel model, the improved model and exact mathematical projections, for rear tracker binning (upper) and distance driven binning (lower) reconstructions. The noise from the MC and from the exact mathematical projections RSP image is obtained from the annular ROIs. The noise of the central pixel model and improved model are obtained from a line profile along the diameter of the reconstructed noise map (with the simplified interpolation effect, i.e. figure 6.8 (b) & (d)).

6.4 Discussion

In this study, we presented a formalism for 2D noise reconstruction in pCT based on a beam of incident protons traveling in parallel. 2D noise reconstructions of a cylindrical water phantom were obtained from variance projections (equation 6.22) using an FBP algorithm, either for rear tracker binning or DDB.

For projection noise, results from MC simulation data, a central pixel model and an improved model were shown for rear tracker binning (see figure 6.6). Good agreement was obtained for rear tracker binning when using the improved model and the MC data. We made use of the improved analytical model, in addition to MC simulations, to better isolate the contributions to projection noise and help explain the shortcomings of the central pixel model, which were not readily deduced from MC data. As a side note, the central pixel and improved models yielded smoother variance projections and noise images as they do not suffer from statistical fluctuations as MC does, while being less demanding on computing resources. This may increase convergence speeds when optimizing fluence patterns for FMpCT.

An important finding of this work is the influence of MCS on the calculation of the variance projections. As shown in figure 6.6, the result of the calculation of the variance projection using the central pixel model, which neglects the effect of MCS, deviates considerably from the results obtained with MC. The deviation mainly stems from the estimation of the variance of the energy in a pixel. Even for a mono–energetic proton beam, there are two main contributions to the variance of the energy. The first is the proton energy straggling. It is the dominant contribution at the center of the object and as it is accounted for in the central pixel model, both analytical models (central pixel and improved) and MC yield very similar results for the variance projections in this region of the phantom. The second contribution to the variance of the energy comes from MCS. Protons can scatter in a pixel having traversed very different paths. This leads to an increase of the variance of the energy beyond the level expected from energy straggling of protons which follow very similar paths and cannot be described by the central pixel model. The improved model takes this effect into account and therefore reaches very good agreement with the MC. Nevertheless, the results shown in this study refer to a homogeneous cylindrical phantom. Homogeneous water is yet a typical assumption for the MLP, especially if no prior image is available. The impact of heterogeneities as well as more complex phantom surfaces will have to be incorporated to the improved model.

As shown in figure 6.7, the shape of the noise projection profiles along the lateral coordinate ξ changes when binning protons at different depths along the longitudinal coordinate η. The noise projection is described by higher noise at the edges of the object. For rear tracker binning ($\eta = 260\,\text{mm}$), this effect is very pronounced due to the importance of drift along the increasing air gap between the object and the rear tracker which causes protons with widely different paths to reach the same projection pixel. DDB mitigates this for projections with η in the object by using the MLP. This can also be appreciated in figure 6.9 where we observed that the image noise from DDB is about 80% lower at 100 mm radius than for rear tracker binning, and more interestingly, relatively flat vs. the

object's radius.

Good agreement between the results from our noise reconstruction formalism and noise analysis using annular ROIs on the reconstructed image was observed. Slight deviations at the object's edge could be attributed to effects present in the reconstruction of discretized ideal projections (see figure 6.9). In order to reduce the computational time needed, we exploited the radial symmetry of the phantom and the resulting radial symmetry of the reconstructed noise map and therefore used annular ROIs for the quantification of the noise from the MC. This assumes that there is no correlation between the different pixels, which is not entirely true (Wunderlich & Noo 2008). A noise quantification from MC that would bypass this assumption would be done pixel–wise on a large set of different RSP images. In that case, the noise in every pixel would be the standard deviation of the RSP values of that pixel from all RSP image realizations.

Note that any analytical noise reconstruction depends on the choice of the convolution kernel. It has a significant impact on the noise, via the corresponding frequency windowing. Alternatives include the Shepp–Logan or the Hanning convolution kernels, which show greater noise suppression and reduction of ringing artifact (Buzug 2008). In the present work, we chose the Ram–Lak kernel, given in equation 6.3 or shown in figure A.1 of the appendix. Its alternating side–lobes cancel one another in the expression of the convolution kernel, which results in minimal correlation (only nearest neighbors contribute to the covariance) as discussed in appendix A.0.1. The general pCT noise reconstruction with an arbitrary filter is given by equation 6.21 , where different expressions for $g\left(j\Delta\xi\right)$ can be implemented in equations 6.16 and 6.17.

How different convolution kernels affect the noise in the reconstructed image has been investigated by e.g. Zhang & Ning (2008) for x–ray cone–beam CT.

In addition to the intrinsic physical effects mentioned above, real detector performance will also affect the noise in a pCT image. In reality, the energy of every proton can be measured with finite accuracy. This will be manifested as increased variance of the energy. Bashkirov et al. (2016) reported that for their pCT setup, the energy detector uncertainty was 3 mm water equivalent path length for any given object's water equivalent thickness. Similarly, the tracking system position resolution will impact the estimation of the proton trajectory, which in its turn will magnify the MSC effect on the final noise image. Finally, other detector limitations such as pileup and non–uniform detector performance could alter the noise image with respect to what is reconstructed assuming that every proton that exits the object can be detected with the same accuracy. Further work will aim at carefully investigating the impact of detector uncertainties by making use of the validated simulation platform of Giacometti et al. (2017a) as well as experimentally acquired pencil beam scanning data from Dedes et al. (2018).

The formalism for 2D noise reconstruction we presented was developed for FMpCT; Figure 6.10 illustrates how a clinical implementation would rely on prior imaging data to generate a patient model used for calculating $\sigma^2_{p_{\gamma_n}}$ as input to noise reconstruction. $\sigma^2_{p_{\gamma_n}}$ would be calculated on the basis of an extension of the improved model we presented, or using MC simulation to fully account for heterogeneities in the patient. By comparing the noise reconstruction for a given fluence to a prescribed $\mathrm{Var}\left[f(x_p, y_p)\right]$, the fluence may

be optimized in an iterative procedure, as in (Bartolac et al. 2011). The patient model may be generated on the basis of an existing diagnostic CT scan, a previous full fluence pCT scan or even a pseudo–CT generated from a magnetic resonance imaging scan (Rank et al. 2013, Koivula et al. 2016, Maspero et al. 2017). In addition to incorporating a realistic detector model, future work will also establish the FMpCT fluence optimization strategy, the development of the corresponding PBS modulation as well as the validation of noise reconstruction using experimental data.

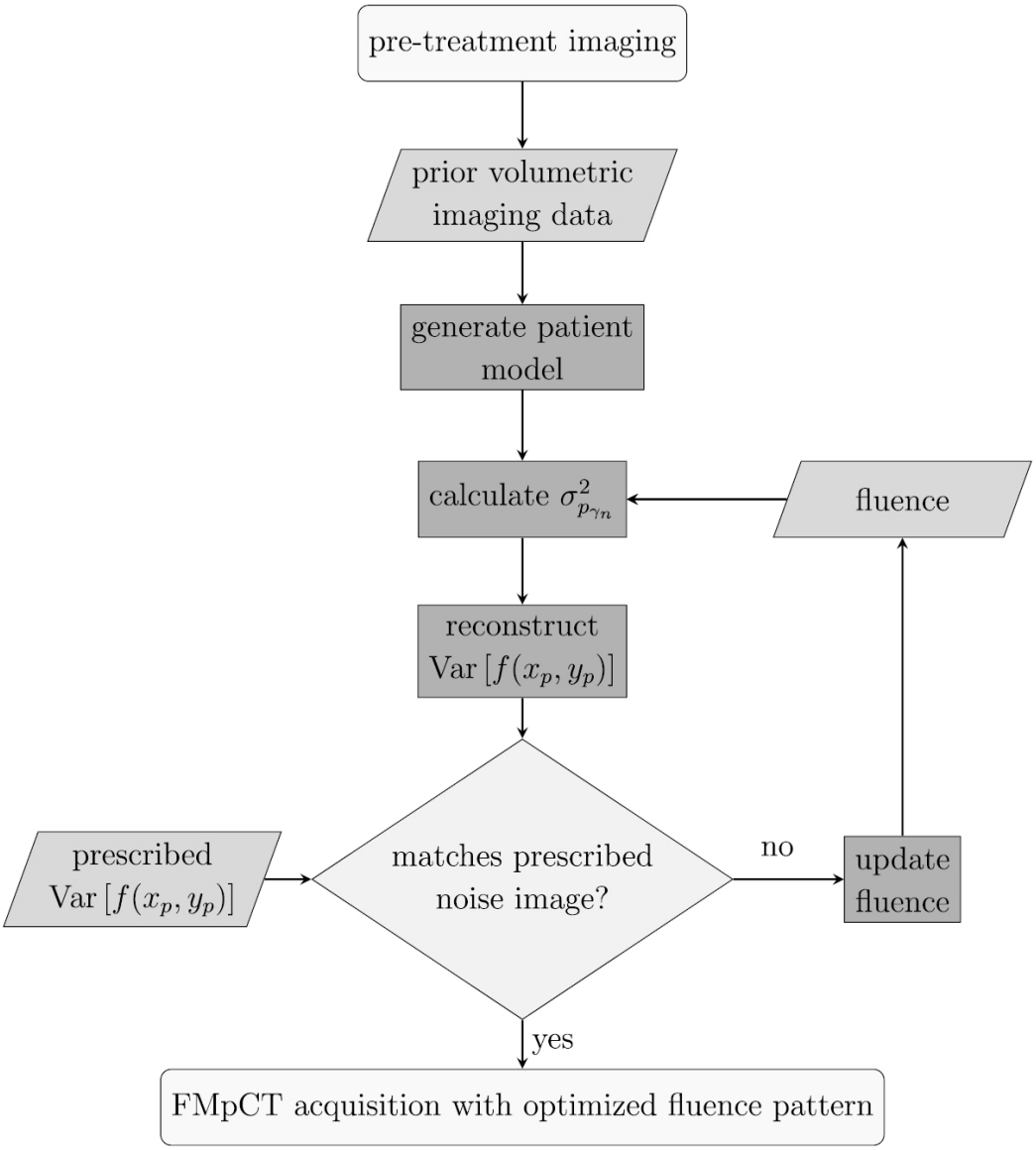

Figure 6.10: Flowchart illustrating how a fluence pattern would be optimized on the basis of the noise reconstruction formalism presented in this work.

6.5 Conclusion

In this paper, we developed a 2D image noise reconstruction formalism to account for both rear tracker binning and DDB in pCT in homogeneous media, assuming parallel proton beams for eventual use in FMpCT fluence optimization. We obtained good agreement between our formalism and with noise estimated from reconstructed images using annular ROIs. The use of DDB slightly decreased the image noise when compared to rear tracker binning and yielded more uniform noise throughout the image. MCS should not be neglected when predicting image noise for pixels away from the center of an object in a pCT scan.

Acknowledgements

This work was supported by the German Research Foundation (DFG) project #388731804 "Fluence modulated proton computed tomography: a new approach for low-dose image guidance in particle therapy" and the DFG's Cluster of Excellence Munich–Centre for Advanced Photonics (MAP), by the Bavaria–California Technology Center (BaCaTeC) project Nr. 28 [2015-2] and by the Bavaria–France Cooperation Centre (BFHZ). David Hansen and Chiara Gianoli are acknowledged for fruitful discussions. We thank Jannis Dickmann for help at the proofreading stage as well as Marco Riboldi for his support of this work.

Chapter 7

Image noise contributions in proton CT

The material contained in this chapter was published in Physics for Medicine and Biology, volume 64, issue 14, page 145016, in July 2019, with the title *Prediction of image noise contributions in proton computed tomography and comparison to measurements*, by Dickmann et al. (2019) – (senior author: Dedes).

7.1 Introduction

The use of proton computed tomography (pCT), initially proposed by Cormack (1963), promises superior accuracy of relative stopping power (RSP) images required for proton therapy dose calculations employed in treatment planning or for various adaptive proton therapy strategies. Instead of using a stoichiometric calibration to convert x–ray linear attenuation coefficients to proton stopping power (Paganetti 2012, Yang et al. 2012), pCT scanners directly measure RSP line integrals, which can be used to reconstruct RSP images using dedicated reconstruction algorithms (Li et al. 2006, Penfold et al. 2009, Penfold et al. 2010, Rit et al. 2013, Poludniowski et al. 2014, Hansen et al. 2016). First prototype proton (and heavier ion) CT scanners (Rinaldi et al. 2013, Coutrakon et al. 2013, Sadrozinski et al. 2016, Johnson et al. 2016, Meyer et al. 2017, Pettersen et al. 2016, Esposito et al. 2018) obtain an RSP line integral — the water equivalent path–length (WEPL) — by measuring the proton's residual energy behind the patient and converting it to WEPL using a calibration obtained prior to the measurement. Reports on RSP accuracy (Giacometti et al. 2017a, Esposito et al. 2018, Volz et al. 2018) suggest that pCT (and heavier ion CT) could equal or outperform accuracy currently achievable with dual–energy CT (Hudobivnik et al. 2016, Wohlfahrt et al. 2017b). Dosimetric accuracy was studied in simulations (Arbor et al. 2015, Meyer et al. 2019) and suggests good performance of using ion CT images for treatment planning.

Apart from RSP accuracy, pCT benefits from its dose–efficiency (Schulte et al. 2005). Reported doses for experimental operation of certain designs and for central pixel noise levels comparable to those of x–ray CT are at only 1 mGy (Sadrozinski et al. 2016).

This is comparable or lower to an in–room cone beam CT (Alaei & Spezi 2015) and can pave the way for daily in–room imaging prior to each treatment session to prevent inaccurate dose delivery during treatment due to positioning errors or anatomical changes. An emerging modality to further reduce imaging dose is fluence modulation, which was originally proposed for x–ray CT (Graham et al. 2007a, Bartolac et al. 2011). Fluence–modulated imaging aims at a dose reduction in parts of the patient by delivering an inhomogeneous imaging dose, and, therefore achieving different noise levels within the image. Fluence modulation has recently gained a strong research interest, particularly due to technical improvements allowing implementation of such systems in x–ray CT (Bartolac & Jaffray 2013, Szczykutowicz & Mistretta 2013a, Szczykutowicz & Mistretta 2013b, Szczykutowicz et al. 2015, Stayman et al. 2016, Gang et al. 2017, Mao et al. 2018, Shunhavanich et al. 2019, Huck, Parodi, Stierstorfer & Fung 2019).

Fluence modulation is particularly meaningful for dose recalculation during particle therapy, where good image quality is only required in the proximity of the beam path and imaging dose can be reduced in healthy tissue where an increased noise level is acceptable (Dedes et al. 2017). This could allow for frequent imaging prior to treatment while maintaining the low integral dose to normal tissue achievable with particle therapy. For imaging with pCT, imaging dose can be accurately delivered in prescribed patterns using the pencil beam scanning functionality of current treatment systems, and the feasibility of FMpCT has been demonstrated both in simulations and proof–of–concept experiments (Dedes et al. 2017, Dedes et al. 2018). Investigations so far focused on static non–optimized fluence maps. To optimize FMpCT plans such that they yield a prescribed image noise map, a prior treatment planning CT may be used as a guide from which pCT noise levels at any fluence setting can be deduced. This requires a method to reconstruct image noise maps from raw pCT data as well as a Monte Carlo simulation that can accurately predict such raw data. A noise reconstruction has been developed in Rädler et al. (2018) for use with filtered backprojection along curved proton paths (Rit et al. 2013). It has so far only been investigated for idealized pCT data of a homogeneous water cylinder. Monte Carlo simulations based on GEANT4 have been used to study a pCT scanner in Giacometti et al. (2017a).

The objective of this study was to demonstrate the feasibility of using a realistic Monte Carlo simulation to accurately predict three–dimensional image noise maps for a given fluence setting. This extends previous studies on pCT noise that either relied on an idealized detector geometry (Rädler et al. 2018) or on a central pixel noise model (Schulte et al. 2005). Data for this study were acquired at a specific pCT prototype scanner described in Sadrozinski et al. (2016). We compared simulated noise predictions to experimental data for three different phantoms, including an anthropomorphic head phantom. We aimed at explaining the shape of resulting image noise maps by disentangling noise contributions in the simulation. Therefore, we hypothesized noise contributions due to multiple Coulomb scattering, energy straggling within the object, uncertainty of the tracking information, energy detection process, and beam energy spread. Finally, we exploited simulated noise maps to apply a fluence profile that yields homogeneous image noise and thus has the same effect as a bow–tie filter for x–ray CT.

7.2 Materials and methods

7.2.1 Experimental setup

Experimental data for this study were acquired using the phase II preclinical pCT prototype scanner developed at Loma Linda University and the University of California, Santa Cruz with details published in Johnson et al. (2016). The single proton tracking scanner consists of two tracking modules, one prior to and one after the imaging object, and a five–stage scintillating detector, which is described in Bashkirov et al. (2016). The scintillators are built from the polystyrene–based material UPS-923A (Artikov et al. 2005) with an RSP of about 1.038, according to Bashkirov et al. (2016). For each proton, two positions before and two after the object, as well as five energy deposits are recorded. For each pair of position information a direction vector can be calculated. A schematic drawing of the scanner geometry is shown in figure 7.1.

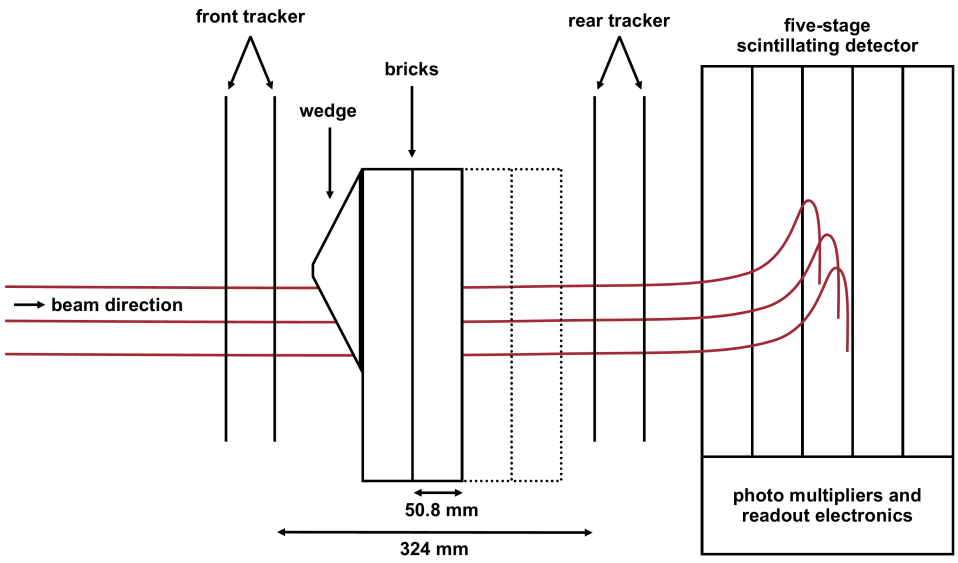

Figure 7.1: True–to–scale schematic drawing of the pCT prototype scanner with front and rear tracking modules and the five–stage scintillating energy detector. Additionally, a wedge–shaped calibration phantom together with two bricks is in place. Up to four bricks can be placed in addition to the wedge as indicated by dashed lines. Three schematic proton representations indicate how the wedge in combination with the bricks scans the dynamic range of the detector (Bragg peak is not to scale and multiple Coulomb scattering is ignored).

Experimental data were acquired at the Northwestern Medicine Chicago Proton Center by wobbling a narrow beam (FWHM $\approx 40\,\text{mm}$) of 200 MeV protons over a FWHM area of $80\,\text{mm} \times 200\,\text{mm}$ for phantom runs and $80\,\text{mm} \times 300\,\text{mm}$ for calibration runs, respectively. During phantom runs, the object was rotated continuously at a speed of 1 rpm to acquire tomographic data for 6 min at a rate of about 10^6 registered protons per second.

7.2.2 Calibration and reconstruction

For each proton, the energy deposit above 1 MeV to the furthest stage (referred to as stopping stage) was mapped to a WEPL using a two–step calibration procedure. In a first step, each stage's channel numbers of the analog–to–digital converter (ADC numbers) obtained in a degrader–free acquisition were mapped to pre–calculated energy values to compensate spatial non–uniformities of the detector response and to achieve an absolute energy measurement. These five energy values E_n^{G4} were obtained in a previous study by Bashkirov et al. (2016). In a second step, data of a wedge–shaped calibration phantom (described in the next section) were used to map the stopping stage's energy deposit to a known WEPL. During subsequent acquisitions, this lookup–table was used to calculate unknown WEPL from measured energy deposits to the stopping stage.

With the spatial information of the tracking modules, a most likely path as described in Schulte et al. (2008) was estimated that then was used to reconstruct tomographic images by binning the WEPL data to virtual projections at every depth of interaction. This process of distance–driven binning is detailed in Rit et al. (2013) and aims at improving spatial resolution in pCT images reconstructed using filtered backprojection. Prior to reconstruction, we performed a rejection of proton histories, referred to in the field as cuts. Protons were rejected if their WEPL or direction information was outside of 3σ boundaries around the median value for bins defined by a 2D grid based on the front tracker position information. This is a standard procedure for pCT (Schulte et al. 2008, Rit et al. 2013).

For reconstruction distance–driven binning was performed at a grid of 180×50 bins laterally and at 180 depths longitudinally with a uniform voxel size of 1 mm in all dimensions. Bins for the calculation of cuts were calculated on a uniform grid of 2 mm by 2 mm. Both RSP and noise maps were reconstructed to a volume of $180 \times 180 \times 50$ voxels with a uniform voxel size of 1 mm in all dimensions. Both the projection grid and the reconstruction grid were centered at the isocenter.

7.2.3 Phantoms

The following phantoms were used in this study. Wherever RSP values for physical phantoms are stated with uncertainties, these were experimentally determined using measurements with a variable water column.

To calibrate the setup, a wedge–shaped calibration phantom made from polystyrene (RSP $= 1.030 \pm 0.003$, Piersimoni et al. (2017)) was used. The flattened peak of the wedge faces the front tracker as shown in 7.1. It has a maximum longitudinal thickness of 50.8 mm while the lateral width is 209 mm. Between zero and four polystyrene bricks with a thickness of 50.8 mm each were placed behind the wedge to cover the whole dynamic range of the detector up to 254 mm.

For tomographic acquisitions, a water phantom, the sensitometric CTP404 phantom (*Phantom Laboratory*, New York, USA) as well as a pediatric head dosimetry phantom (ATOM®, Model 715 HN, *CIRS Inc.*, Norfolk, USA) were used. The water phantom consists of a water–filled PMMA cylinder (RSP $\approx$ 1.17) with an outer diameter

of 150.5 mm and a wall thickness of 6.35 mm. The CTP404 phantom's cylindrical body made from epoxy (RSP $= 1.144 \pm 0.001$, Giacometti et al. (2017a)) has a diameter of 150 mm and multiple cylindrical inserts with RSP ranging from air (RSP < 0.01) to Teflon (RSP $= 1.790 \pm 0.002$, Giacometti et al. (2017a)). The pediatric head phantom is a realistic anatomical model of a 5–year–old child built from tissue–equivalent materials and was used in previous pCT studies (Giacometti et al. 2017b). Since the height of the head phantom was larger than the height of the detector aperture, it needed to be scanned in two consecutive acquisitions with an overlap of several millimeters.

7.2.4 Simulation platform

To simulate acquisitions, a dedicated Monte Carlo simulation platform was used that models the complete geometry of the detector. The platform was described and validated for its RSP fidelity in Giacometti et al. (2017a). It is based on the GEANT4 framework, version 10.2.p01, as presented in Agostinelli et al. (2003). The reference physics list `QGSP_BIP_HP` was used for the simulation of the interaction of particles with matter. For a highly accurate description of electromagnetic interactions, the `G4EmLivermorePhysics` model was used for electrons and photons. The cut for secondary particle production in the energy detector was defined at 6 µm while for the rest of the scanner at 1 mm.

In this work, we extended the platform to model non–linear effects of light production in the scintillator (7.2.6) and to incorporate a realistic beam model (7.2.7). Moreover, we aimed at quantifying the accuracy of image noise (sections 7.2.8 and 7.2.9). All phantoms mentioned in the previous section were simulated as analytical phantoms based on their known geometry and materials, except for the pediatric head phantom, for which a high–resolution voxelized implementation based on x–ray CT scans as presented in Giacometti et al. (2017b) existed. For all phantoms, RSP accuracy was ensured by fine–tuning the I–value in the simulation such that agreement to reference values was better than 0.1 % for protons with an energy of 150 MeV. The I–value of water was 78.0 eV, according to the latest ICRU recommendation.

7.2.5 Verification of noise reconstruction

In order to calculate image noise maps, we used a noise reconstruction method developed for use with distance–driven binning by Rädler et al. (2018). For every bin p with a given mean WEPL in the distance–driven projection, the expected standard deviation of the mean for n protons intersecting that bin was calculated as

$$\sigma_p = \sigma_{\mathrm{WEPL}}/\sqrt{n}, \tag{7.1}$$

where σ_{WEPL} is the per–proton standard deviation of the underlying set of WEPL and σ_p is the expected standard deviation of the mean of n WEPL. This yielded a three–dimensional distance–driven noise projection, which was reconstructed after convolving it with a special noise filter, consisting of the square of the reconstruction filter (e.g.

the filter of Ramachandran & Lakshminarayanan (1971) in Rädler et al. (2018)). As described in Rädler et al. (2018), effects of projection value interpolation (as part of filtered backprojection) on the image noise can be calculated per voxel or as an effective mean over the entire volume. In this study, the second option was chosen as it is computationally more efficient and the high–frequency voxel–wise components are not relevant for a future application using fluence modulation, since they are smaller than any clinical proton pencil beams.

In the publication of Rädler et al. (2018), the method was only applied to simulated ideal pCT data with a parallel beam. Accuracy was only evaluated using annular ROIs, and not on a voxel–by–voxel basis, which is standard (see for example Wunderlich & Noo (2008)). Therefore, in a first noise reconstruction validation study, we simulated $N = 40$ independent noise realizations for a tomographic scan of the water phantom. Each of the noise realizations was reconstructed independently, which allowed us to calculate a voxel–wise standard deviation over N samples, which is referred to as batch method. This served as the ground truth to be compared to the distance–driven noise reconstruction of a single noise realization.

7.2.6 Non–linearities in the detection process

When irradiated by high–LET particles, such as protons of low energy, the light production in the scintillators becomes a non–linear function of the deposited energy due to the effect of quenching as investigated by Birks (1951). To allow for a quantitative prediction of the measurement uncertainty, the simulation needed to account for this non–linearity. In the case of the five–stage energy detector and using the theory of Birks (1951), a proton stopping in stage n and with a residual range R_n within that stage will produce the distorted energy measurement

$$E'_n(R_n) = S_n \cdot \int_{R_n}^{0} \mathrm{d}x \, \frac{\mathrm{d}E/\mathrm{d}x}{1 + k_B \cdot \mathrm{d}E/\mathrm{d}x} \,, \tag{7.2}$$

where k_B is the empirical Birks' factor and S_n is a scaling factor specific for each stage, which is fixed during the calibration. In appendix B.0.1 equation B.5, we calculate the scaling factor taking into account the specific calibration procedure. We find a formulation for $E'_n(\mathrm{WEPL})$ as a function of the WEPL instead of the range. This formulation depends only on the Birks' factor k_B as well as the thickness of each scintillating stage l_{stage} and the residual range in the detector material at the entrance of the detector R_0. The required material specific stopping power $\mathrm{d}E/\mathrm{d}x$ was calculated as a function of the residual range R for the correct material composition using the Geant4 platform.

At the same time, we were able to obtain energy measurements $E^*_n(\mathrm{WEPL})$ from experimental data using the wedge phantom, which covers the whole dynamic range $\mathrm{WEPL} \in [0, 254]\,\mathrm{mm}$ of the detector. For each proton the energy deposit E^*_n was calculated as well as the WEPL based on the tracking information and the known geometry.

For the whole dataset, this yielded a 2D–histogram of E_n^* against WEPL, from which the relationship $E_n^*(\mathrm{WEPL})$ was deduced by finding the most frequent energy deposit for each WEPL.

This allowed us to find the estimation of the three unknowns k_B, R_0 and l_{stage} by minimizing the cost function

$$C = \sum_n \left[\sum_{\mathrm{WEPL}} \left(E'_n(\mathrm{WEPL}) - E_n^*(\mathrm{WEPL}) \right)^2 \right]^2 \tag{7.3}$$

using the quasi–Newton method of Broyden (1970). The Birks' factor k_B was then used to modify each incremental energy deposit in the simulation as

$$\frac{\mathrm{d}\,\mathrm{ADC}}{\mathrm{d}x} \propto \frac{\mathrm{d}E'}{\mathrm{d}x} \propto \frac{\mathrm{d}E/\mathrm{d}x}{1 + k_B \cdot \mathrm{d}E/\mathrm{d}x}, \tag{7.4}$$

where d ADC is the increment of the simulated ADC number. The scaling factor S_n was neglected here due to the arbitrary scaling of the ADC number.

7.2.7 Realistic beam model

To realistically model the initial positions and directions of protons in the simulation, we exploited the fact that such measurements are directly available from experimental data. Each proton's position was projected to a point at 400 mm in front of the isocenter, and, therefore, 232.8 mm in front of the first tracking module, along a straight line according to its direction vector. To avoid interplay–effects between the proton placement and the tracking strip, the position information was blurred with a random number normal–distributed around zero and with a standard deviation equal to the distance between two tracking strips.

The experimental dataset, on which this study was based, used different field widths for the calibration and the subsequent phantom acquisitions. Modeling noise of phantom acquisitions correctly required to use the smaller field width of these measurements. Unfortunately, degrader–free runs were only available for the calibration runs with the larger field width. However, acquisitions with a the water phantom in the beam path were used to create a beam model. We therefore accounted for the additional attenuation due to the phantom by randomly selecting protons with a probability anti–proportional to the transmittance.

The initial energy of each proton could not be measured directly, because in the degrader–free run the calibration forces it to be $E_0 = \sum_n E_n^{\mathrm{G4}} + \Delta E_{\mathrm{tracker}} = 200\,\mathrm{MeV}$, where $\Delta E_{\mathrm{tracker}} = 3.65\,\mathrm{MeV}$ is the assumed mean energy loss in the tracking modules and air. It was, however, possible to determine the beam energy spread σ_{beam}, which is likely to have a strong impact on the resulting image noise. Assuming that electronic readout noise is negligible compared to σ_{beam}, we performed degrader–free simulations without attenuation correction and for values of $\sigma_{\mathrm{beam}} \in \{0.5, 0.6, \ldots, 1.0\}\,\mathrm{MeV}$. We calculated the resulting spread σ_{E_5} of the energy deposit in the last stage using Gaussian fits. Since the

simulations considered the optimal Birks' factor as described in the last section, it was possible to compare this σ_{E_5} to the one obtained from a measurement and choose σ_{beam} such that σ_{E_5} matches the experimental value.

7.2.8 Comparison to experimental data in the projection and noise contributions

To ascertain that the modeling in sections 7.2.6 and 7.2.7 resulted in accurate noise predictions, we simulated a projection of the wedge–shaped calibration phantom with zero to three bricks and compared them to experimental data in terms of standard deviation. Due to the larger extent of the phantom compared to reconstructions, projections were binned to 250×50 bins laterally and at 250 depths longitudinally. The voxel size and all other parameters are described in section 7.2.2.

After applying the WEPL calibration and cuts, two sets of distance–binned projections were generated from the data. The first was the average calibrated WEPL. The second was the per–proton WEPL standard deviation $\sigma_{\text{WEPL}} = \sigma_p \cdot \sqrt{\bar{n}}$, where σ_p is the standard deviation according to equation 7.1 and $\bar{n}$ is the average fluence in the projection. The normalization to a fluence of one proton per bin makes results comparable to, e.g. Bashkirov et al. (2016). After rejecting data outside the hull of the wedge, this resulted in a list of WEPL–noise–pairs, with every pixel in the projection giving one entry in the list. For each binning depth, these pairs were binned by WEPL to multiples of 1.5 mm and for each bin the median noise value was calculated.

We hypothesize that noise in the projection can be attributed to five different contributions as listed in table 7.1. In order to study these contributions to image noise, noise calculations were performed for simulations using the following scoring methods, which disentangle each of these effects.

1. *WEPL scoring.* Dynamic scoring in GEANT4 of each proton's exact WEPL by multiplying the stepping length with the current material's RSP calculated for the current proton energy. Exact coordinates when crossing the tracking planes were recorded. Noise is caused only by scattering of protons with different paths into the same distance–driven bin.

2. *Energy scoring.* Scoring of the proton's energy at tracking planes before and after the phantom and conversion of the energy loss to WEPL. Exact coordinates when crossing the tracking planes were recorded. Noise is caused by multiple Coulomb scattering as well as energy straggling within the object.

3. *Energy scoring (realistic position).* Energy scoring, but using the position information estimated by the tracking modules. Noise is increased due to the additional uncertainty of the tracking information.

4. *Realistic scoring ($\sigma_E = 0\,\text{MeV}$).* Fully realistic simulation, but setting the beam energy spread to 0 MeV. Noise is increased compared to *energy scoring* due to energy

straggling in the detector and the calibration process, but it is lower compared to *realistic scoring*, which includes the beam energy spread, and its impact can be quantified.

5. *Realistic scoring.* Fully realistic simulation, including the described beam model with the determined beam energy spread. This model contains all known noise contributions.

Table 7.1: Calculation of the five noise contributions as a difference of the five scoring techniques. Numbers in brackets refer to the listing of scoring techniques in 7.2.8.

noise contribution	difference of scoring techniques
scattering	(i)
energy straggling (ES) in the object	(ii) – (i)
tracking	(iii) – (ii)
energy detection process	(iv) – (iii)
beam energy spread	(v) – (iv)

To quantify the relative noise contribution of scattering, energy straggling in the object, tracking, the energy detection process and the beam energy spread, we calculated the difference in projection noise values for the five scoring techniques as indicated in 7.1. Since the cuts of proton histories depend on the underlying standard deviation, the efficiency of data filtering may vary across datasets and impair this analysis as a different set of protons was used to calculate each dataset. However, considerable differences of the cuts efficiency were only observed at the edge of the phantom and only for the *WEPL scoring* technique, where the analyses including this scoring technique may not be quantitative. To correctly assess the relative noise contributions, this part investigates variance values, which are the square of the standard deviation values we calculated so far. This is because variance contributions add up linearly, while standard deviation contributions need to be added quadratically. The variance difference between the scoring techniques was calculated for each WEPL bin and normalized to the variance using *realistic scoring*.

The previous analyses were performed on very homogeneous data of the continuous part from the wedge–shaped calibration phantom. To demonstrate the impact of heterogeneities, we calculated standard deviation profiles for the steep edge of the calibration phantom. The evaluation of this dataset is shown in appendix B.0.2.

7.2.9 Comparison to experimental data with heterogeneous and anthropomorphic phantoms

To investigate the performance of the Monte Carlo simulation for predicting image noise maps, we simulated pCT data for the water phantom, the CTP404 phantom as well as the pediatric head phantom (see section 7.2.3) using the *realistic scoring* technique. For the same phantoms, experimental data were acquired using the prototype pCT scanner. We reconstructed both RSP and noise images. All image noise maps were normalized to an average projection fluence of $f_0 = 20\,\mathrm{mm}^{-2}$. For the water phantom and the CTP404 phantom, which are symmetric in the z–direction, 16 slices were averaged after noise reconstruction.

In a first step, we manually selected two corresponding slices and registered the 2D–RSP image of the simulation onto the experimental reconstruction using a rigid registration allowing translation and rotation. This registration was then applied to the corresponding slices of the image noise maps. In the following evaluation, only pixels inside the object's outer hull as determined by an RSP threshold of 0.15 were considered.

The registration allowed a calculation of the pixel–wise relative error map $\Delta\sigma$ in terms of an estimation of the standard deviation as

$$\Delta\sigma = \frac{\sigma_{\mathrm{RSP,sim}} - \sigma_{\mathrm{RSP,exp}}}{\sigma_{\mathrm{RSP,exp}}}, \tag{7.5}$$

where σ_{RSP} is the registered data of the image noise reconstruction for experimental and simulated data, respectively.

Using the *WEPL scoring* technique (see 7.2.8), we reconstructed images showing the simulated scatter–only standard deviation $\sigma_{\mathrm{scatter}}$. The non–scatter standard deviation $\sigma_{\text{non-scatter}}$, i.e. the expected standard deviation in absence of scatter, was calculated as

$$\sigma_{\text{non-scatter}} = \sqrt{\sigma^2_{\mathrm{RSP}} - \sigma^2_{\mathrm{scatter}}}, \tag{7.6}$$

where σ_{RSP} is the total standard deviation. Again, a varying cuts efficiency may impair this analysis in particular at the edges of objects.

Using standard GEANT4 functionality of the simulation platform, the imaging dose was scored for every projection angle and then summed. Mean imaging dose values for individual slices were calculated within the phantom.

As a sanity check, we calculated RSP and standard deviation histograms of the whole volume of the head phantom. Results of this verification are presented in appendix B.0.3.

7.2.10 Application: a bow–tie filter for proton CT

To demonstrate the feasibility of using Monte Carlo simulated noise data for fluence modulation, we calculated a fluence modulation profile for the homogeneous water phantom which has the same effect as a bow–tie filter in x–ray CT. The simplest concept of an x–ray bow–tie filter aims at homogeneous noise at the detector as described in Harpen (1999) or

Graham et al. (2007a). While in x–ray CT this can be achieved by aiming at homogeneous fluence at the detector, this does not hold true for pCT because of the impact of multiple Coulomb scattering as discussed in Rädler et al. (2018) for an idealized scanner. Instead, a modulated fluence is required that yields a constant noise level at the detector, which in turn will make noise flat in the image, as can be derived from the variance reconstruction formula presented in Rädler et al. (2018).

The artificial fluence modulation profile was calculated solely based on simulated data. It was then applied to experimental data by randomly selecting protons with an acceptance probability $p(u)$ as a function of the lateral coordinate u (no modulation was needed orthogonal to that due to the symmetric phantom). Given a simulated image noise profile $\sigma_{p,\mathrm{sim}}(u)$ in the projection binned at the isocenter, we calculated a fluence modulation

$$p(u) = \min\left(\frac{\sigma^2_{p,\mathrm{sim}}(u)}{\sigma_0^2}, 1\right) \tag{7.7}$$

that forces the standard deviation in the projection to the value σ_0. Note that the fluence modulation is proportional to variance and, therefore, to the square of the standard deviation. In the given case, it was only possible to reduce fluence, and, therefore, $\sigma_0 \geq \max_u \sigma_{p,\mathrm{sim}}(u)$. However, this condition could also be violated at the cost of a resulting non–homogeneous fluence where noise is highest in the original acquisition. A violation is required at the object's hull, where noise tends to be elevated. In these regions, an unreasonably high skin dose would result from forcing the noise level to be the same as inside the object. In this study, we prescribed a projection standard deviation of $\sigma_0 = 5.48\,\mathrm{mm}$, and, therefore, a variance of $\sigma_0^2 = 30\,\mathrm{mm}^2$.

7.3 Results

7.3.1 Verification of noise reconstruction

Figure 7.2 shows the standard deviation ground truth calculated with $N = 40$ noise realizations as well as the noise reconstruction for the water phantom simulation. Both image noise maps agreed well. The residual difference was caused by the approximation of the effect of projection value interpolation in our noise reconstruction and displayed as a star–shaped increase in noise that was centered at the center of rotation and that spanned the whole reconstruction volume. In the relative difference map shown in (c) the star–shaped pattern reoccurred as a under–representation of the noise reconstruction. The agreement between the ground truth and the noise reconstruction was good with the mean error over the whole phantom being $-2.5\,\%$ and the root–mean–square error $4.1\,\%$. Also the profile plots in (d) revealed that there was a slight underestimation of noise from the noise reconstruction.

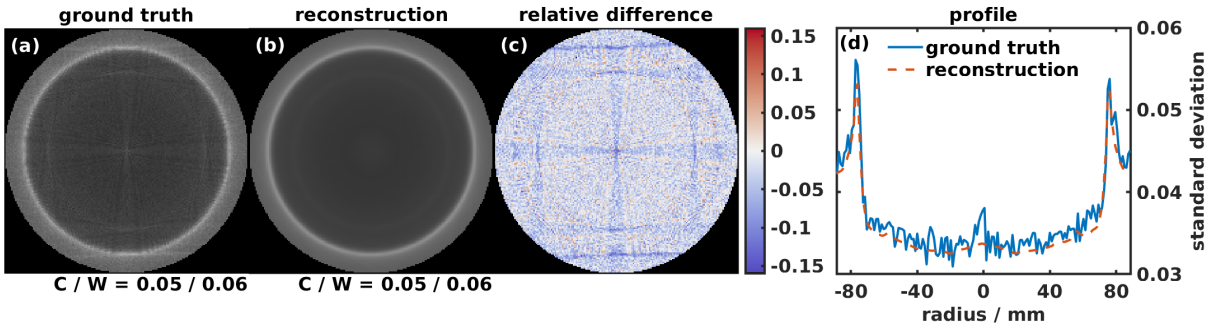

Figure 7.2: Standard deviation images calculated for the water phantom (a) with $N = 40$ noise realizations as a ground truth, (b) with one noise realization using the noise reconstruction; (c) the relative difference to the ground truth (reconstruction - ground truth)/ground truth; (d) diagonal profiles across (a) and (b). Grayscale center (C) and window (W) settings were the same in (a) and (b).

7.3.2 Non–linearities in the detection process

Figure 7.3 (a) shows a two–dimensional histogram of the relative proton counts for WEPL versus energy deposit in the stopping stage. Data were obtained from an experimental acquisition of the wedge–shaped calibration phantom, for which the WEPL was geometrically calculated for every proton. Each of the five lines in the histogram corresponds to one of the five stopping stages. For example, the line at the lowest WEPL values corresponds to the fifth stage as protons with a low WEPL can penetrate all other stages before stopping in the last one. With increasing WEPL, the energy deposit decreases until suddenly the stopping stage changes and the energy deposit is high again.

For each stage, there was some WEPL range at multiples of the WEPL of one brick $\mathrm{WEPL}_{\mathrm{brick}} = l \cdot \mathrm{RSP} = 52.3\,\mathrm{mm}$ for which counts decreased and which is visible in the histogram as a vertical line of lower intensity. At this point, data need to be merged between protons going through n bricks and the center of the wedge, and protons going through $n + 1$ bricks, but missing the wedge (see figure 7.1). These flat regions of the calibration phantom and in particular divergent protons leaving the lateral edge of bricks cause the observed change in statistics. However, since each WEPL bin was evaluated individually to obtain the relationship $E_n^*(\mathrm{WEPL})$, this did not affect the evaluation.

Table 7.2 reports the three free parameters, which were optimized by minimizing the cost–function in equation 7.3. Note that these parameters were found solely from experimental data.

In figure 7.3 (b), the same histogram is shown for a simulated dataset and for $k_B = 0\,\mathrm{mm\,MeV^{-1}}$ and, therefore, neglecting quenching effects. For a given WEPL, energy deposits were considerably higher compared to the experimental data. Figure 7.3 (c) shows simulated data using the optimized Birks' factor. The energy deposits decreased and agreed with experimental data shown in figure 7.3 (a). For experimental data, there was a cut–off related to the measurement process below an energy deposit of around 20 MeV for the first

stage (WEPL around 250 mm), which was not present in the simulated data.

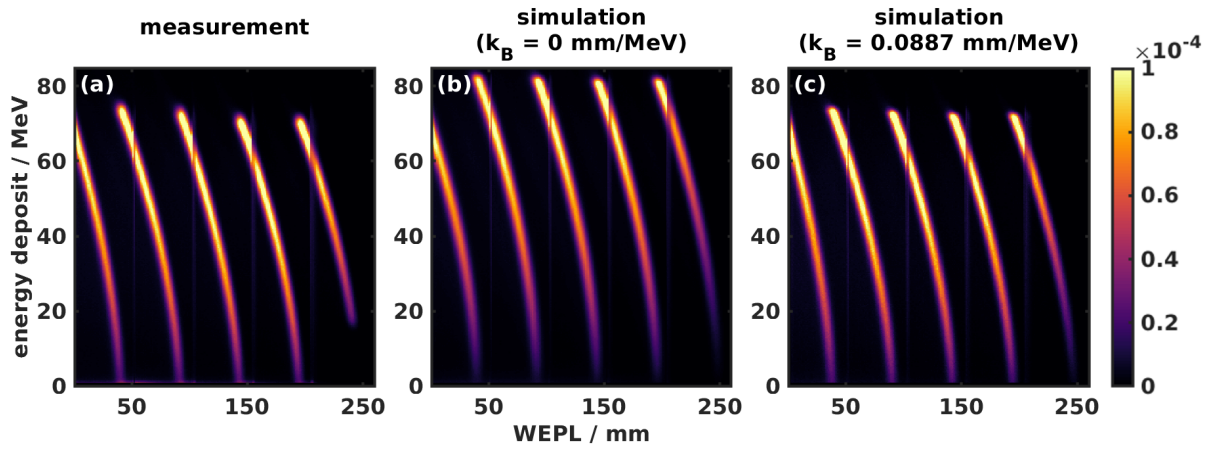

Figure 7.3: Two–dimensional histograms showing relative proton counts for water–equivalent pathlength versus energy deposit in the stopping stage based on data of the wedge–shaped calibration phantom for (a) experimental data, (b) the standard simulation without modeling quenching, and (c) the extended simulation with correct modeling of quenching. The sum of all counts is normalized to unity.

Table 7.2: Optimization result of the three free parameters to model quenching effects in the detector.

parameter	fit estimate
Birks' coefficient k_B	0.0887 mm/MeV
residual range at entrance R_0	237.9 mm
stage thickness l_{stage}	49.7 mm

Figure 7.4 shows calibration curves that were used to calculate a WEPL given the energy deposit to the stopping stage for each individual proton. These curves were obtained for each stage during the calibration process. Three calibrations were calculated based on data shown in figure 7.3 (a) - (c). As expected from the raw data, the calibration curve of the simulation without quenching disagreed with the one of the experimental data and the simulation using the optimized Birks' factor.

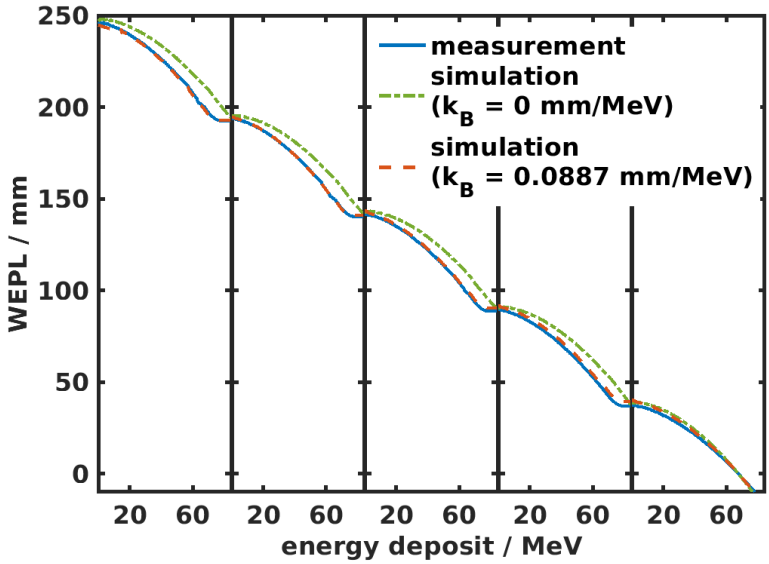

Figure 7.4: Calibration curves mapping a given energy deposit in the stopping stage to a water–equivalent path length for experimental data, the standard simulation without modeling quenching, and the extended simulation with correct modeling of quenching.

7.3.3 Realistic beam model

Figure 7.5 shows the resulting spread of the last stage's calibrated energy deposit σ_{E_5} as a function of σ_{beam} together with a quadratic fit to this relationship. In experimental data, where σ_{beam} cannot be measured directly, we determined the energy spread of the last stage to be $\sigma_{E_5} = (3.47 \pm 0.02)\,\text{MeV}$. According to the quadratic fit, this corresponded to a beam energy spread of $\sigma_{\text{beam}} = (0.66 \pm 0.02)\,\text{MeV}$. The uncertainty of σ_{E_5} is given by a one standard deviation confidence interval of the fit parameter and the uncertainty of σ_{beam} was calculated as an error propagation thereof through the quadratic fitting function. Both uncertainties are visualized in figure 7.5 as shaded areas.

7.3.4 Comparison to experimental data in the projection domain and noise contributions

Figures 7.6 (a) - (c) shows the WEPL standard deviation as a function of the WEPL for projection data of the wedge–shaped calibration phantom. Data are shown at three different binning depths: at the front tracker position (a), at the isocenter (b), and at the rear tracker position (c). While the first five color–coded curves show the noise level for simulations with different scoring techniques at increasing complexity, the last curve was obtained from experimental data. See section 7.2.8 for the definition of terms printed in italic type in the following paragraphs.

The lowest noise level was recorded for *WEPL scoring*. While for binning at the front

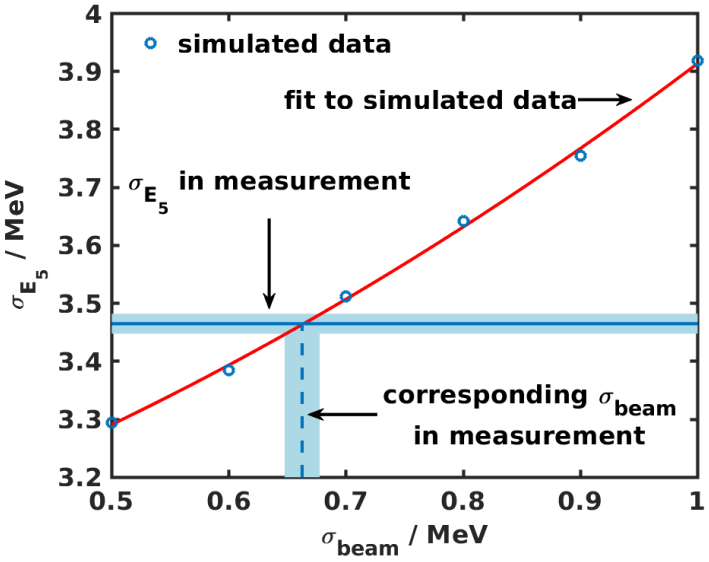

Figure 7.5: Relationship between the beam energy spread σ_{beam} and the spread of the energy deposit of the last stage σ_{E_5} for a degrader–free measurement. Circles indicate simulated data to which a quadratic model was fitted. Shaded areas indicate a one standard deviation confidence interval. The intersection of this fit with the σ_{E_5} of a measurement was chosen to be the correct σ_{beam} to be used in simulations.

tracker standard deviation was below 0.2 mm, it increased to values between 1 mm and 2 mm for binning at the rear tracker with a general increase of noise with the WEPL. Additionally, at multiples of the brick thickness (indicated by the dashed lines in figure 7.6), an abrupt decrease of standard deviation of about 0.25 mm was observed for binning at the isocenter and at the rear tracker, which will be explained in section 7.4.4.

For *energy scoring*, noise increased with increasing WEPL and pronounced effects at multiples of the brick thickness were only observed for rear tracker binning. Use of *energy scoring (realistic position)* had only a minor effect on the standard deviation compared to *energy scoring*.

For *realistic scoring*, using the same calibration method as in a measurement, the dependency of the standard deviation on WEPL became less pronounced. Table 7.3 presents mean WEPL standard deviation over the whole WEPL range. There is a considerable increase in standard deviation due to the beam energy spread. The curve for the fully realistic simulation, as well as its mean value, agreed within their uncertainty with the experimental data.

In figures 7.6 (d) - (f), the relative contributions to WEPL variance are shown, and can be attributed to the following effects (see section 7.2.8): scattering, energy straggling in the object, tracking, the energy detection process and the beam energy spread. While scattering was negligible at the front tracker (d), the accumulated contribution at the rear tracker (f) reached about 20 %. The contribution of the tracking uncertainty was

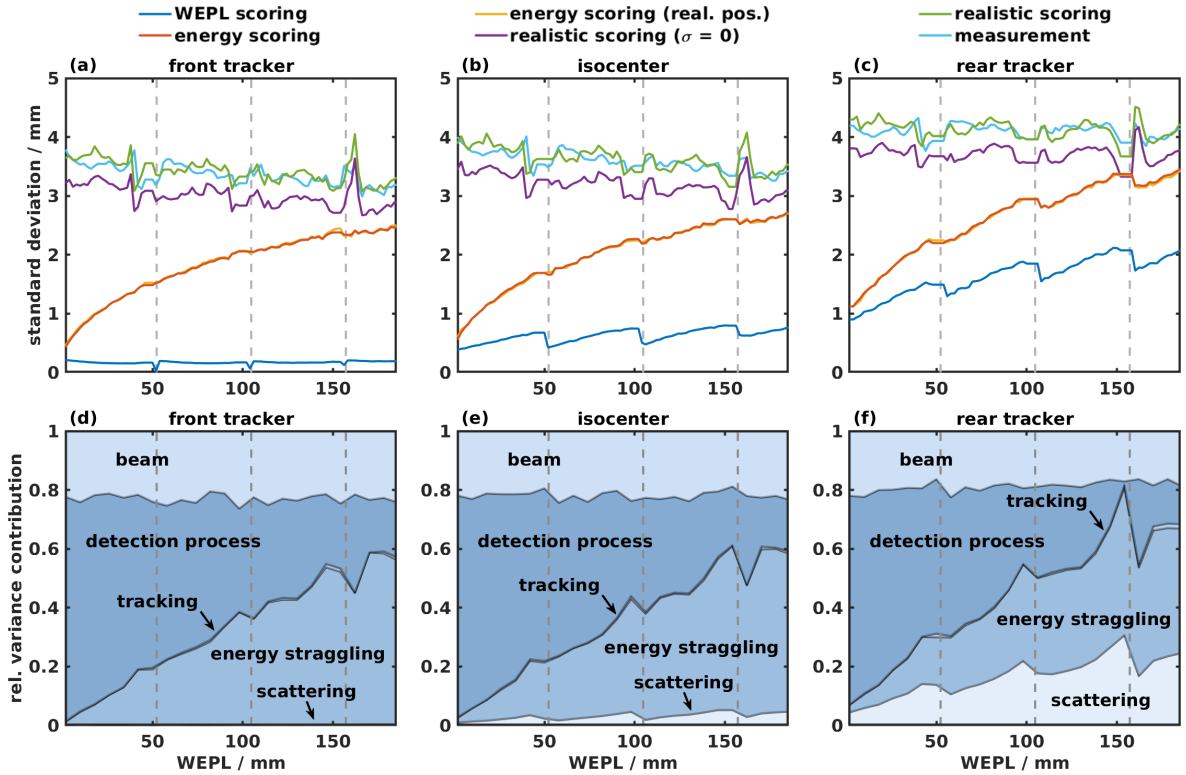

Figure 7.6: (a) - (c) Standard deviation level in the projection domain for different scoring techniques used in a simulation and comparison to the experimentally determined noise level; (d) - (e) relative variance contributions attributed to physical effects as calculated by subtraction of data in (a). All data are shown at three different depths of the distance–driven binning: at the front tracker for (a) and (d), the isocenter for (b) and (e) and the rear tracker position for (c) and (f). Dashed lines indicate multiples of the brick thickness. Note that the legend in (a) - (c) spans across all three plots.

Table 7.3: Mean WEPL standard deviation over the 0 mm to 184 mm WEPL range for two different scoring techniques as well as for experimental data binned at three different depths.

scoring	mean standard deviation / mm		
	front tracker	isocenter	rear tracker
realistic scoring ($\sigma_E = 0$)	2.99 ± 0.18	3.19 ± 0.19	3.71 ± 0.14
realistic scoring	3.41 ± 0.20	3.61 ± 0.21	4.12 ± 0.16
measurement	3.35 ± 0.19	3.57 ± 0.18	4.09 ± 0.11

negligible over the whole WEPL range. The main source of noise in the projection of the wedge phantom was the detection process as well as energy straggling in the object.

Both contributions had an approximately constant sum, while the contribution of energy straggling in the object increased proportionally to the WEPL. The variance added by the beam energy spread had a constant value of about 20 % independent of the WEPL.

7.3.5 Comparison to experimental data with heterogeneous and anthropomorphic phantoms

In figure 7.7, RSP and image noise reconstructions of experimental and simulated data, for three different phantoms are shown: the homogeneous water phantom, the sensitometric CTP404 phantom with tubular inserts as well as three slices of the CIRS pediatric head phantom. Figure 7.7 (a) shows the RSP reconstruction of the experimental data. Standard deviation maps shown in figure 7.7 (b) and (c) all show an increased noise level at the hull of the object. In general, experimental and simulated data agreed well visually and subtle noise features near heterogeneities of the CTP404 phantom and the head phantom were captured.

In the RSP images of the homogeneous water phantom, ring artifacts were apparent at the center and one at a larger radius. These were caused by inaccuracies of the calibration at multiples of the brick thickness and are known to degrade RSP accuracy as reported in Sadrozinski et al. (2016). These ring artifacts became more smeared out in the more heterogeneous phantoms such as the CTP404 and are barely visible in the head phantom, except for the very homogeneous region in the brain (slice 3).

The overall good agreement between simulation and measurement was supported by the difference maps relative to the noise level of the experimental data in figure 7.7 (d). The relative error was increased primarily in regions where ring artifacts also impact the RSP accuracy, as it was seen for the water phantom, but also for the CTP404 phantom and the homogeneous part of the head phantom (slice 3). Table 7.4 summarizes the mean and the root mean square error for the simulated versus the experimentally acquired standard deviation maps. While the water phantom and the CTP404 phantom both had a negative offset, the head phantom showed both slices where the simulated noise maps over– and under–estimated the experimental level. Root mean square errors are below 10 % for all acquisitions.

In figure 7.7 (e), profile plots across the standard deviation maps going horizontally through the isocenter are shown for simulation and measurement. The position coordinate was normalized from the phantom entrance to the exit point. For all phantoms, subtle spatial fluctuations in noise level were captured. The absolute noise level in the center was lower for the homogeneous water phantom and slice 3 of the head phantom (around 0.025). For the more heterogeneous CTP404 phantom and slice 2 of the head phantom, the standard deviation was increased to about 0.03, while for the most heterogeneous slice 1, the noise level exceeded 0.05, i.e. a two–fold increase compared to the homogeneous phantoms. Noise levels as the mean standard deviation within a centered circular region with a radius of 20 mm are summarized in table 7.5. The scatter–only noise contribution obtained with *WEPL scoring* and shown as a third profile, behaves correspondingly. For

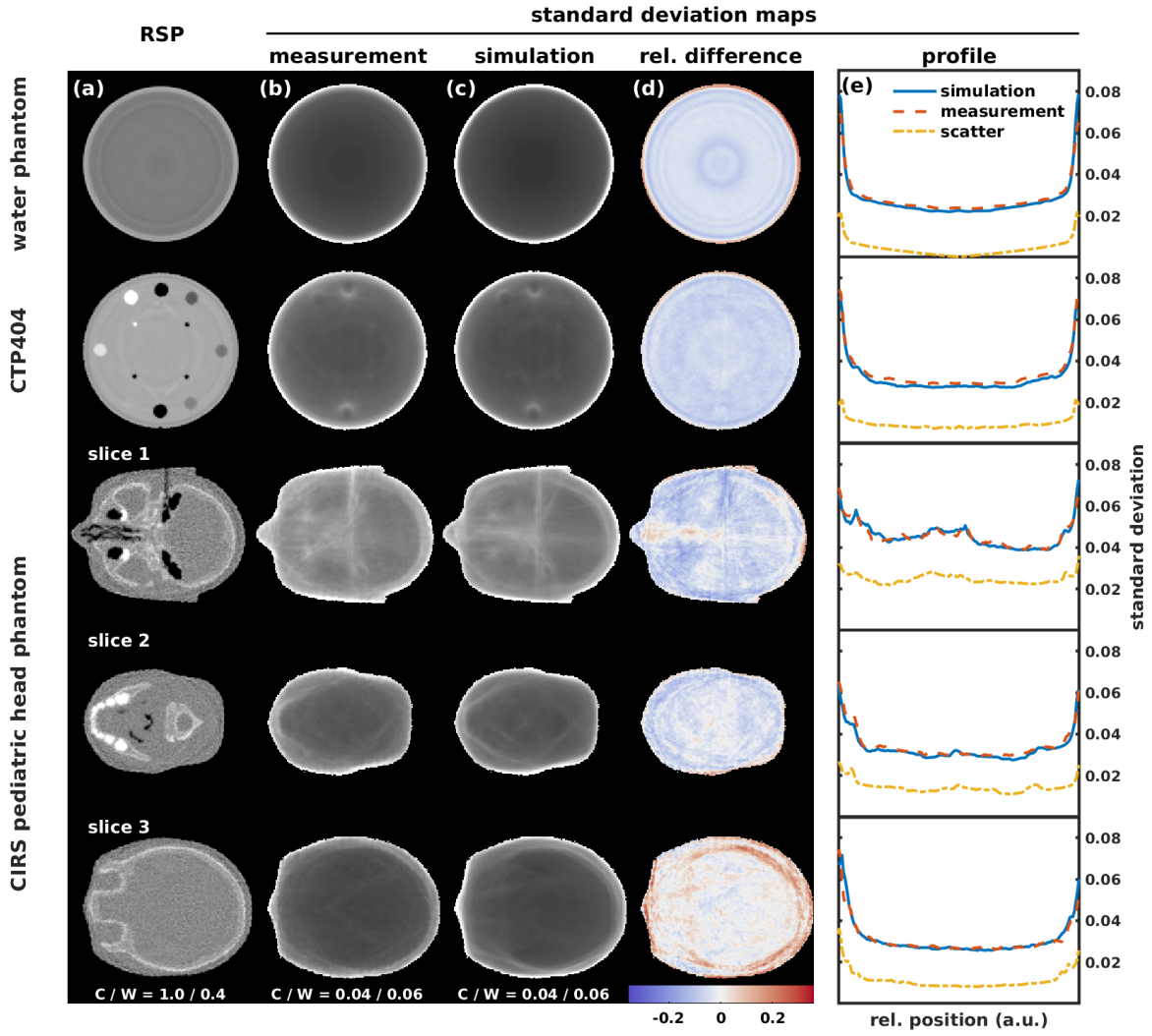

Figure 7.7: RSP and image noise reconstructions of experimental and simulated data for three phantoms. The RSP reconstruction of the experimental data (a), standard deviation maps for experimental (b) and simulated (c) data, a relative difference map between measurement and reconstruction (d), as well as a horizontal profile plot through the noise reconstructions (e) are shown. Profiles in (e) also include a simulated scatter–only reconstruction. The diameter of the three head phantom slices was 171 mm, 146 mm and 170 mm, respectively, and 150 mm for the two other phantoms. Grayscale center (C) and window (W) settings are given for each display.

the homogeneous water phantom the scatter contribution went to zero at the center and increased towards the edges. For the CTP404 phantom and slice 3 of the head phantom it was 0.01 at the center and had a similar, but less pronounced increase towards the edges. While for slice 2 of the head phantom it reached 0.015, for slice 1 the scatter–only standard deviation was above 0.025. The non–scatter contribution $\sigma_{\text{non-scatter}}$ according

Table 7.4: Relative mean and root mean square noise prediction errors for the phantom slices shown in figure 7.7. Errors are the pixel–wise difference between the simulated standard deviation versus the experimental data, relative to the experimental data.

phantom (slice)	mean error / %	root mean square error / %
water phantom	−2.9	6.7
CTP404 phantom	−5.6	6.4
head phantom (1)	−4.6	6.8
head phantom (2)	−3.6	5.3
head phantom (3)	3.2	6.2

to equation 7.6 was mostly constant for all phantoms. In the central region excluding the edges it was $\sigma_{\text{non-scatter}} = 0.028 \pm 0.004$, where the uncertainty is calculated over all phantoms.

Table 7.5: Noise levels (mean standard deviation) within a centered circular region and mean imaging doses calculated using the simulation platform.

phantom (slice)	noise level (simulation)	noise level (experiment)	imaging dose / mGy (simulation)
water phantom	0.022 ± 0.001	0.024 ± 0.001	0.85 ± 0.04
CTP404 phantom	0.027 ± 0.001	0.029 ± 0.001	0.86 ± 0.04
head phantom (1)	0.044 ± 0.003	0.046 ± 0.002	0.82 ± 0.04
head phantom (2)	0.030 ± 0.001	0.030 ± 0.001	0.84 ± 0.04
head phantom (3)	0.026 ± 0.001	0.027 ± 0.001	0.88 ± 0.04

To allow for a comparison of noise levels in other and future studies, table 7.5 reports average imaging doses. Experimental imaging doses are expected to be similar as the number of protons was matched and the effect of pile–up was found to be small in previous studies (Dedes et al. 2018). Reconstruction parameters that impact image noise, such as the voxel size, are given in section 7.2.2.

7.3.6 Application: a bow–tie filter for proton CT

In figure 7.8 (a), standard deviation maps are shown for experimental data of the water phantom without post–processing fluence modulation, and in figure 7.8 (b) with a fluence modulation according to equation 7.7 that aims at equalizing noise in the projection domain and, therefore, also in the image. Since this part was based on experimental data, fluence modulation was done by rejecting events and only fluence reduction was possible.

Therefore, the modulated image noise map in figure 7.8 (a) had a higher average noise level compared to the unmodulated map in figure 7.8 (b). While standard deviation was clearly lower in the center of figure 7.8 (a) compared to the outside, the virtual bow–tie modulation succeeded in having a homogeneous noise level. At the outer hull of the phantom, noise is purposely kept high as discussed in section 7.2.10. In figure 7.8 (c), horizontal profile plots across the image noise maps are shown, which demonstrated again the homogeneous profile that can be reached with the virtual bow–tie filter. Figure 7.8 (d) shows the fluence at the front tracker. While fluence was homogeneous in the unmodulated data, it was reduced in the center of the object in order to achieve homogeneous image noise. RSP accuracy of the fluence–modulated reconstruction was not degraded compared to the non–modulated reconstruction.

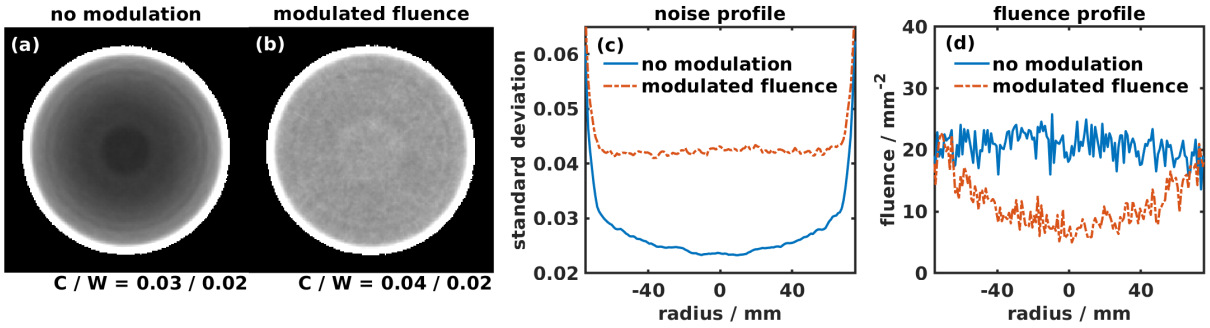

Figure 7.8: Fluence modulation to reach a homogeneous noise level in the water phantom: (a) the unmodulated standard deviation map and (b) the standard deviation map after applying the fluence modulation, (c) horizontal profiles through the standard deviation maps, (d) fluence profiles at the front tracker. Note that the grayscales in (a) and (b) have a different center (C), but the same window (W) setting.

7.4 Discussion

7.4.1 Verification of noise reconstruction

Using a ground truth generated from independent noise realizations, we were able to show that noise reconstruction predicts the image standard deviation with a root–mean–square error of only a few percent. Accurate image noise maps for FMpCT may be calculated very efficiently at computational costs similar to a standard filtered backprojection reconstruction and without the need to simulate multiple independent noise realizations for every fluence pattern. The remaining star–shaped discrepancy between the noise reconstruction and the ground truth is systematic and due to approximation of interpolation effects, which was extensively discussed in Rädler et al. (2018). This pattern is independent of the imaged object and did not impact evaluations in this study, as the same reconstruction algorithm was applied to both experimental and simulated data and a systematic error

would cancel out in a relative comparison. Furthermore, for FMpCT, such high–frequency components were not relevant in the given scope, as fluence can only be modulated within the extension of one pencil beam.

7.4.2 Non–linearities in the detection process

By fitting Birks' law to experimental data, we obtained a Birks' coefficient allowing the simulation of quenching effects in the energy detector's five scintillators. The initial range of protons at the entrance of the detector and the stage thickness were also fitting parameters. The CSDA range of 200 MeV protons in polystyrene is 250.0 mm as calculated from the *PSTAR* database (Berger et al. 2005). This value needs to be reduced by an equivalent of the physical tracker thickness of 2×0.4 mm in silicon (see Johnson et al. (2016)) and the energy loss in air, and, therefore, the fitting result given in table 7.2 were reasonable. The physical stage thickness was 51 mm and thus also this fitting result was in the correct range. Uncertainties in the knowledge of the precise beam energy, the trackers' and the scintillators' RSP, as well as the impact of wrapping materials around the scintillator stages, did not allow to precisely calculate these two values. Given that the fit resulted in values close to the expected ones, it confirmed that leaving them as free parameters is a valid procedure.

The determined Birks' factor in table 7.2 multiplied by the density of polystyrene ($\rho = 1.06\,\mathrm{g/cm^3}$) was $k_B \cdot \rho = 9.4 \times 10^{-3}\,\mathrm{gMeV^{-1}cm^{-2}}$, which agreed with values for polystyrene published in literature of $k_B \cdot \rho = 9 \times 10^{-3}\,\mathrm{gMeV^{-1}cm^{-2}}$ in Tretyak (2010) or $k_B \cdot \rho = 14 \times 10^{-3}\,\mathrm{gMeV^{-1}cm^{-2}}$ in Reichhart et al. (2012).

Using the optimized Birks' factor, the agreement between the simulation and experimental data increased considerably. A subtle remaining difference between the simulation and the experimental data was that no protons are recorded with an energy deposit of less than 20 MeV in the first stage. This was because the first stage is needed for triggering recording of an event in experiments and, therefore, protons below this threshold are generally not considered. In the simulation, each proton was treated individually and a threshold was not needed nor considered. This, however, was no problem, because none of the phantoms in this study had WEPL above 200 mm.

The parametrization of Birks' law came at almost no additional computational cost for data simulation while a full–scale simulation of the optical photon transport, which is possible with Geant4, would be unfeasible in terms of computational speed considering the amount of data needed for pCT applications.

In particular, the shape of the calibration curve was highly impacted by quenching. For WEPL accuracy, the exact shape of the calibration curve is irrelevant, since in any case the energy deposit gets assigned the correct WEPL. Therefore, in Giacometti et al. (2017a) a high RSP fidelity was achieved even without a quenching model. However, non–linearities in the detection process may impact accuracy of the WEPL image noise prediction. This is because the change in WEPL caused by a given change in the energy deposit (e.g. due to range straggling in a prior stage) is proportional to the inclination of the calibration curve, which is incorrect when neglecting quenching. Furthermore, having

calibrated energy deposits of the simulation agree with those of a measurement, allowed to fine–tune the beam energy spread in the next section.

7.4.3 Realistic beam model

By exploiting tracking information from experimental data, we created a virtual beam model to be used in simulations. The estimated beam energy spread given in section 7.3.3 was lower than the value typically expected for a proton treatment facility of about 0.5 % to 1 % of the initial beam energy as described in Schippers (2017), which would be around 1 MeV to 2 MeV in the case of 200 MeV protons. However, in order to reduce fluence to a level that can be handled by the pCT scanner, momentum slits of the accelerator needed to be closed down to a minimum which is atypical for treatment operation, and which may explain the lower energy spread.

7.4.4 Comparison to experimental data in the projection domain and noise contributions

Using different scoring techniques, we disentangled noise contributions in simulated data. For the continuous part of the wedge–shaped calibration phantom, we observed that the scatter–only component was low for front tracker binning and highly elevated for rear tracker binning. This was because of the lateral widening of the beam envelope as protons traverse the phantom. Due to the geometry of the calibration phantom shown in figure 7.1 with the wedge facing the front tracker, for a given front tracker location, protons will go through a small distribution of WEPL. However, for a given rear tracker location, protons can have scattered from a larger cone due to the drift in air (a passage with a low scattering probability) after the phantom and thus the WEPL distribution was broader. Because the accumulated mean scattering angle increases with the WEPL there was an increased WEPL uncertainty towards higher WEPL.

Moreover, an abrupt decrease of the standard deviation was observed at multiples of the brick thickness. This has a similar geometrical reason as the region of reduced statistics described in section 7.3.2. Assume a proton going through the center of the wedge phantom. By adding a brick and moving to the periphery of the wedge, the WEPL stays almost constant and so does the accumulated mean scattering angle. However, the geometrical distance of the drift in air is reduced and, therefore, noise decreases (in figure 7.1 compare the upper proton with the lower proton with one additional brick).

The main contribution of projection variance came from energy straggling, which occurs both within the imaging object as well as in the scintillating detector. Energy straggling in the object increased proportionally to the WEPL. Energy straggling in the detector decreased accordingly such that the total contribution of energy straggling stayed constant. Apart from energy straggling in the scintillator, noise in the detection process was also increased due to the impact of the calibration curve, which itself is noisy and of varying slope. Further noise contributions to the detection process, such as the energy resolution

of the scintillator and electronic noise, are assumed to be small and independent of the WEPL. They were discussed in Bashkirov et al. (2016), where the design of the detector was chosen such that WEPL noise is close to the energy straggling limit. Their remaining contribution may be covered by the beam energy spread estimated in section 7.2.7.

We showed that the contribution of the tracking uncertainty due to the finite resolution of the location and direction measurement is negligible compared to other sources of uncertainty. A considerable contribution, however, came from the beam energy spread, which contributed a fraction of about 20 % despite the fact that in section 7.2.7 it was estimated to be lower than the energy spread during treatment. This completes the work of Rädler et al. (2018) and is the first noise simulation of a realistic detector system considering all relevant contributions.

7.4.5 Comparison to experimental data with heterogeneous and anthropomorphic phantoms

By calculating image noise maps for various phantoms with different amounts of heterogeneities and comparing them to experimental data, we demonstrated the feasibility of using Monte Carlo simulated data to predict image noise in complex geometries. Remaining errors were below 10 % and, therefore, small compared to the overall fluctuation of image noise for the anthropomorphic head phantom which varied within a factor of 2.3 in experimental data (see appendix B.0.3). Therefore, fluence modulation based on the simulated noise maps should be feasible.

Furthermore, we showed that the absolute image noise level depends on the heterogeneity of the phantom. The increase of the noise level in heterogeneous phantoms is driven by multiple Coulomb scattering along heterogeneities, which were shown to be considerably different even between two slices of the same (anthropomorphic) phantom. This means that in a clinical setting, the imaging dose advantage of pCT over conventional x–ray CT might be less than expected, as the previous study investigating the density resolution of pCT (Schulte et al. 2005) used a phantom with homogeneous materials and thus may have predicted a reduced noise level compared to a clinically relevant geometry.

The non–scatter noise contribution was found to be comparable for all phantoms (excluding the edge). This agrees with section 7.4.4 and experiments of Bashkirov et al. (2016), who found WEPL–variance to be constant over the whole WEPL–range for a homogeneous phantom without heterogeneities. The impact of a single heterogeneity in a controlled setting was demonstrated in appendix B.0.2.

7.4.6 Application: a bow–tie filter for proton CT

We calculated the profile of a virtual bow–tie filter, which makes noise flat at the detector level. The modulation profile was found based on the simulated noise prediction. The resulting noise profile in image domain was flat as desired and mean RSP accuracy was maintained. This result showed the feasibility of using Monte Carlo simulated noise predictions to modulate the fluence in experimental scans according to a given noise prescription.

The virtual bow–tie fluence profile was fundamentally different compared to the fluence profile expected in x–ray CT, which is high in the center of the patient to compensate for the stronger attenuation. For proton CT instead, fluence must be lower in the center and higher at the periphery to compensate for noise introduced due to scattering at the object's hull. However, unlike in fan–beam x–ray CT, where imaging is based on the straight path of primary photons, we cannot expect the shape of a bow–tie filter based on a homogeneous water–equivalent model to be generally valid for proton CT. As we showed, noise maps for the pediatric head phantom differ considerably from the ones observed for the water phantom. Therefore, this virtual bow–tie filter is only valid for a given phantom and is presented here to demonstrate the capability of predicting, prescribing and creating a specific noise pattern.

7.5 Conclusion

We demonstrated the feasibility of using image noise reconstruction to generate tomographic image noise maps using the cone–beam geometry of a prototype pCT scanner by comparing data to a simulated ground truth. By modeling quenching effects, we were able to match calibrated energy deposits of the five–stage energy detector to Monte Carlo simulated data. This allowed us to determine the beam energy spread of the given incident proton beam. Together with experimental tracking data, this allowed us to create an accurate beam model to be used in the simulation that matches the experimental beam in terms of positions, direction vectors, energy and energy spread. In the projection domain, we compared the Monte–Carlo–predicted noise levels with experimental data and disentangled noise contributions. While the contribution of scattering to noise is negligible in the center of a homogeneous phantom, it becomes a dominating source of noise around heterogeneities for an anthropomorphic head phantom. This is a novel finding for a realistic detector model compared to a previous study (Rädler et al. 2018) that investigated the impact of multiple Coulomb scattering at the edge of a homogeneous phantom for an idealized detector.

In conclusion, we improved a Monte Carlo simulation to yield accurate image noise maps, which we compared to experimental data. The accuracy of predicted noise is better than the expected fluctuations within a head phantom. Therefore, fluence–modulated proton CT based on Monte Carlo simulated image noise maps should be feasible. At the same time, it was shown that a simple ray–tracing model mapping a WEPL to its uncertainty would not be feasible. To demonstrate the use of calculated image noise maps, we obtained a virtual fluence modulation profile to achieve a bow–tie–like homogeneous noise level in the image. While this was successful for a homogeneous water phantom, such fluence modulation profiles will depend on the patient geometry and cannot be generalized. This suggests that patient–specific fluence–modulation is a crucial component for dose–efficient pCT imaging.

Acknowledgments

Dr. Valentina Giacometti is gratefully acknowledged for developing and sharing the GEANT4 simulation platform used in this study. We thank Nick Detrich for support during data acquisition. This work was supported by the German Research Foundation (DFG) project #388731804 "Fluence modulated proton computed tomography: a new approach for low–dose image guidance in particle therapy" and the DFG's Cluster of Excellence Munich–Centre for Advanced Photonics (MAP), by the Bavaria–California Technology Center (BaCaTeC) project Nr. 28 [2015-2] and by the Franco–Bavarian university cooperation center (BayFrance).

Chapter 8

Optimization algorithm for fluence–modulated proton CT

The investigation presented in this chapter was published in Medical Physics, volume 47, issue 4, pages 1895–1906, in April 2020, with the title *An optimization algorithm for dose reduction with fluence–modulated proton CT*, by Dickmann et al. (2020) – (senior author: Dedes).

8.1 Introduction

Cancer treatment using intensity–modulated proton and heavier ion therapy is effective, and comes at a low risk of side–effects for the patient compared to conventional treatment modalities using x–rays. The good tolerance is believed to be linked to the low dose to normal tissue when using protons for treatment (Weber et al. 2012, Verma et al. 2016, Nakajima et al. 2017, Kim et al. 2019). At the same time, low–dose, frequent and accurate imaging, ideally at the treatment site, is required to ensure a safe delivery of the therapeutic doses (Landry & Hua 2018, Nenoff et al. 2019). Proton therapy treatment planning requires a spatial map of the relative stopping power (RSP), which in current clinical practice is acquired through a conversion from x–ray CT images (Hudobivnik et al. 2016, Wohlfahrt et al. 2017b, Taasti et al. 2018). X–ray CT images are typically not acquired in treatment position and not prior to every treatment fraction, in order to keep treatment time short and imaging dose low enough that they do not compromise the dose benefit of proton therapy (Alaei & Spezi 2015). Direct imaging of RSP using proton CT (pCT) (Rinaldi et al. 2013, Schulte et al. 2005, Johnson et al. 2016, Esposito et al. 2018, Volz et al. 2018, Meyer et al. 2019) has been proposed to increase accuracy and to allow for a frequent, dose efficient acquisition in treatment position. Accuracies achievable with current prototypes are comparable to state–of–the art clinical dual energy x–ray CT (Yang et al. 2012, Hudobivnik et al. 2016, Hansen et al. 2015, Dedes et al. 2019).

A further reduction of imaging dose can be achieved by modulating the imaging fluence field during the acquisition and thereby achieving a task–specific image quality. Fluence–

modulated scans (Graham et al. 2007b) can either aim for homogeneous variance across the whole volume, or for region–of–interest imaging, where only the relevant part of the image is acquired at low noise and imaging dose is reduced elsewhere. Algorithms (Bartolac et al. 2011, Hsieh & Pelc 2014, Gang et al. 2019, Wang et al. 2019) and experimental prototypes (Bartolac & Jaffray 2013, Szczykutowicz & Mistretta 2014, Szczykutowicz et al. 2015, Stayman et al. 2016, Huck, Fung, Parodi & Stierstorfer 2019) have been developed for fluence modulation in x–ray CT. Recently, fluence–modulated proton computed tomography (FMpCT) has also been proposed (Dedes et al. 2017) and its initial experimental feasibility using pencil beam scanning was investigated (Dedes et al. 2018). The best achievable dose efficiency through fluence modulation or other techniques is a key requirement for x–ray CT (Lell & Kachelrieß 2020) and most likely will be for pCT as it moves closer to the clinics. Moreover, region–of–interest imaging is of high interest for particle therapy treatment planning and dose verification, where only a fraction of the image volume (the treatment beam path) is of relevance (Dedes et al. 2017). A challenge for FMpCT is that simple Poisson noise modeling is not sufficient, as image variance for pCT depends on the object's heterogeneity, and several contributions, including multiple Coulomb scattering, have to be taken into account for fluence–modulation (Rädler et al. 2018, Dickmann et al. 2019).

In this work, we present a method for fluence–field optimization in pCT using pencil beam scanning. We employ a pCT scanner–specific Monte Carlo simulation (Giacometti et al. 2017a), which was shown to reproduce experimental variance levels for a typical fluence field (Dickmann et al. 2019). The problem of finding relative modulation factors for each pencil beam such that the summed fluence pattern results in a prescribed image variance map is a computationally expensive optimization problem which generally requires alternating between the reconstructed image domain (where the variance prescription is defined) and the projection domain (the detector data at each projection angle from which the image is reconstructed, and where the fluence modulation is defined). Therefore we separated the problem into first solving for the projection domain variance yielding a given prescribed variance in the image domain and subsequently optimizing pencil beam weights leading to this projection domain variance. To realistically describe pencil beams in the optimization and in simulations, we established a pencil beam model based on experimental data. In a simulation study, we estimated dose savings for fluence–modulated pCT using three different phantoms, and compared our proposed solution with a straightforward intersection–based fluence modulation (Dedes et al. 2018). We also verified that the resulting variance map approaches the target variance. Both a constant variance target as well as two regions–of–interest (ROIs) following typical treatment beam paths were investigated.

8.2 Materials and methods

8.2.1 Simulation framework

The Monte Carlo simulation framework (Giacometti et al. 2017a) used in this study is a detailed model of the phase II pCT prototype scanner (Johnson et al. 2016). It is based on the GEANT4 toolkit (Agostinelli et al. 2003) version 10.2.p01. Details about the modeling of physics processes can be found in literature, where the platform was validated for its fidelity in terms of RSP (Giacometti et al. 2017a, Dedes et al. 2019). A previous study (Dickmann et al. 2019) improved the platform for reproducing variance levels of experimental scans. With respect to that work, the beam model was modified, and is described below. Imaging doses, in the form of absorbed physical dose, were scored on a centered grid of $125 \times 125 \times 35$ voxels with a uniform voxel size of 2 mm and summed for all projection angles.

The simulation framework outputs data in the same format as the prototype scanner. It records position and direction information of individual protons before and after the object, as well as the proton's residual energy. Using a calibration (Bashkirov et al. 2016), the residual energy can be mapped to a WEPL, which is the line integral over the RSP of the object along the curved path of the proton. Because measurements are available for every detected proton, these data are referred to as "list–mode."

8.2.2 Image reconstruction

To reconstruct RSP images from the list–mode data, a most likely path (Schulte et al. 2008) is estimated for every proton from the tracking information. The path information is taken into account by performing distance–driven binning and applying a special cone–beam filtered backprojection algorithm(Rit et al. 2013). In total, 90 projections from rotation angles covering a full rotation were used. This relatively low number of projections was chosen to satisfy experimental timing constraints and to allow for a future experimental validation of this work. Reconstructions were performed on a grid of $250 \times 250 \times 70$ voxels with a uniform size of 1 mm. For performing data cuts,(Schulte et al. 2005, Schulte et al. 2008, Rit et al. 2013) the grid was 125×125 pixels with a uniform size of 2 mm. Binning of data into distance–driven projections was performed on a grid of $250 \times 250 \times 70$ voxels with a uniform size of 1 mm. All grids were centered on the isocenter.

Assume a voxel centered in (u, v, d) in the three–dimensional distance–driven projection, where d is the binning depth and u and v are the coordinates normal to it. We can identify a set of protons such that the most likely path of every proton crosses the voxel volume around (u, v, d). The number of protons in that set is then referred to as the "counts" $C(u, v, d)$. These counts only consider protons used for image reconstruction and therefore are reduced compared to the incident protons due to interactions with the object and cuts on the data. In contrast to that, counts in the absence of interactions and cuts are referred to as $F(u, v, d)$ throughout the paper. The point $u = v = d = 0$ is the location of the isocenter, where the rotation axis is located.

8.2.3 Phantoms

In the simulation study, three different phantoms with a physical counterpart were used. The water phantom is a cylindrical container made from polystyrene (outer diameter 150.5 mm, wall thickness 6.35 mm) and filled with distilled water. The CTP404 phantom (*Phantom Laboratory*, New York, USA) is a commercial sensitometric phantom with a cylindrical shape (diameter 150 mm) and several tissue–equivalent inserts and two cylinders filled with air. Both phantoms were implemented in the simulation as analytical models. The pediatric head phantom (ATOM®, Model 715 HN, *CIRS Inc.*, Norfolk, USA) models the anatomy of a 5–year–old child and was implemented as a voxelized phantom in the simulation (Giacometti et al. 2017b). Previous publications (Dickmann et al. 2019, Giacometti et al. 2017a, Giacometti et al. 2017b) can be consulted for details about the phantoms.

8.2.4 Gaussian pencil beam model

To allow for the flexible simulation of FMpCT data, an analytical pencil beam model was derived from experimental tracking data acquired at the pencil beam scanning beamline of the Northwestern Medicine Chicago Proton Center without phantom. Using the timing information of the scanner, a count rate was calculated in steps of 0.8 ms, allowing for the separation of individual pencil beams as the count rate dropped in between two spots. The separated data were processed individually by estimating most likely paths and performing distance–driven binning (Rit et al. 2013).

For each pencil beam b, this resulted in a three–dimensional experimental counts map $C_b(u, v, d)$. We fitted the Gaussian model

$$G(u, v, d) = \frac{N_0}{2\pi\sigma'_u\sigma'_v} \cdot \exp\left(-\frac{(u - u'_0(d))^2}{2\sigma'^2_u} - \frac{(v - v'_0(d))^2}{2\sigma'^2_v}\right) \tag{8.1}$$

to each pencil beam's C_b, where N_0 is the total number of protons per pencil beam, and $(u'_0(d), v'_0(d))$ is the pencil beam center at depth d. The pencil beam center is assumed to diverge linearly with the binning depth, such that $u'_0(d) = u_0 \cdot (1 + \delta_u \cdot d)$ and $v'_0(d)$ analogously, where (u_0, v_0) is the pencil beam center at $d = 0$ and δ_u and δ_v are the linear divergence factors. By construction, the isocenter–beam for $u_0 = v_0 = 0$ is parallel to the d–axis. The $\sigma'_u = \sigma_u \cdot \sqrt{1 + \delta_u^2 u_0^2}$ and σ'_v analogously are the beam widths projected to a plane normal to the d–axis while σ_u and σ_v are the actual beam widths. This resulted in a fit with seven open parameters (N_0, u_0, v_0, σ_u, σ_v, δ_u, δ_v), which was performed for each individual pencil beam by minimization of the squared deviation. The parameters σ_u, σ_v, δ_u and δ_v were not specific to one pencil beam, and estimates for them were therefore found as the mean value over all pencil beams. N_0, v_0 and u_0 were open parameters specific to a given pencil beam, but overwritten in subsequent simulations of different pencil beam patterns. They are therefore not reported.

Simulation of pencil beams

All datasets were generated by shooting a regular grid of simulated proton pencil beams. At $d = 0$, neighboring pencil beams were interspaced by $\Delta_{\mathrm{PB},u} = 12\,\mathrm{mm}$ along u and $\Delta_{\mathrm{PB},v} = 8\,\mathrm{mm}$ along v. The pencil beam grid was offset in u by $\Delta_{\mathrm{PB},u}/4 = 3\,\mathrm{mm}$ so that the summed fluence from two opposing angles was homogeneous. This helped to reduce the total number of pencil beams and thereby reduce the complexity of the optimization. In the simulation platform, protons were emitted from a point $(u_0 \cdot (1+\delta_u \cdot d_0), v_0 \cdot (1+\delta_v \cdot d_0), d_0) + (r_u, r_v, 0)$, where $d_0 = -400\,\mathrm{mm}$ and r_u and r_v are normally distributed random numbers with a standard deviation of σ_u and σ_v respectively. The point d_0 is just before the front tracker and was chosen in agreement to previous investigations (Dickmann et al. 2019). Protons were assumed to have an initial direction vector of $(u_0\delta_u, v_0\delta_v, 1)$. The beam centers (u_0, v_0) were chosen according the pencil beam grid defined above. For non–modulated scans, N_0 was set to a default value $N_0 = N$ for all pencil beams. For modulated scans it was $N_0 = m_b^\alpha N$ for a pencil beam modulated with a factor m_b^α. The proton's initial energy was set to $(200.00 \pm 0.66)\,\mathrm{MeV}$, which is the standard mean energy used experimentally. The energy spread was determined in a previous study (Dickmann et al. 2019) and agrees with experimental data acquired at the beamline at Northwestern Medicine Chicago Proton Center, albeit with a wider spot size setting.

Pencil beam reference counts

To optimize pencil beam weights, a reference of the proton counts for every pencil beam is needed. This reference serves as a basis function for the fluence modulation and should not take into account interactions with the object. It can be generated for every pencil beam b using the Gaussian model

$$F_b(u, v, d) = G(u, v, d)\Big|_{N_0=N, u_0=u_b, v_0=v_b} \tag{8.2}$$

assuming a pencil beam center (u_b, v_b) according to the regular grid and a constant number of protons N which is equal for all pencil beams.

Optimization of pencil beam weights

Using the F_b as basis functions, it is possible to generate an arbitrary counts field C^α for rotation angle α by finding weights w_b^α, such that C^α is expressed as a linear combination of the reference counts F_b from equarion 8.2. Weights were found by minimizing the squared deviation, and therefore

$$w_b^\alpha(C^\alpha) = \arg\min_{w_b^\alpha} \iint \mathrm{d}u\,\mathrm{d}v \left(C^\alpha(u, v, 0) - \sum_b w_b^\alpha F_b(u, v, 0) \right)^2 . \tag{8.3}$$

Integration was performed over u and v, but only the isocenter binning depth $d = 0$ was considered. Optimization was performed using the method of Nelder and Mead (Nelder & Mead 1965).

8.2.5 Proposed algorithm for fluence field optimization

Fluence field optimization requires finding a set of fluence modulation factors $m_b^\alpha \in [0,1]$ for pencil beam b at rotation angle α, such that the resulting pCT reconstruction best achieves a given image variance target $V_{\text{target}}(x,y,z)$. The proposed method for fluence field optimization is performed in the projection domain (denoted by coordinates (u,v,d) and the rotation angle α) instead of the image domain (denoted by coordinates (x,y,z)). The method is sketched in figure 8.1 and consists of the following three steps, which will be detailed in the next sections:

1. For a given phantom, find the resulting variance $V_{\text{unit}}^\alpha(u,v,d)$ in the projection domain for a unit fluence simulation with $m_b^\alpha = 1$ for all pencil beams.
2. For a given image variance target $V_{\text{target}}(x,y,z)$, find a stack of variance levels in the projection domain $V_{\text{target}}^\alpha(u,v,d)$ that yields the image variance target.
3. Calculate the pixel–wise counts target $C_{\text{target}}^\alpha(u,v,d)$. Then, optimize weights that yield the counts target according to equation 8.3.

The algorithm extends ideas from literature for x–ray CT (Bartolac et al. 2011, Hsieh & Pelc 2014) to requirements of pCT such as the three–dimensional projections due to distance–driven binning (Rit et al. 2013) and a more complex noise model (Rädler et al. 2018, Dickmann et al. 2019). It is, to our knowledge, not equivalent to any existing approach as it is performed in projection domain and computationally feasible without simplification to a parallel–beam geometry.

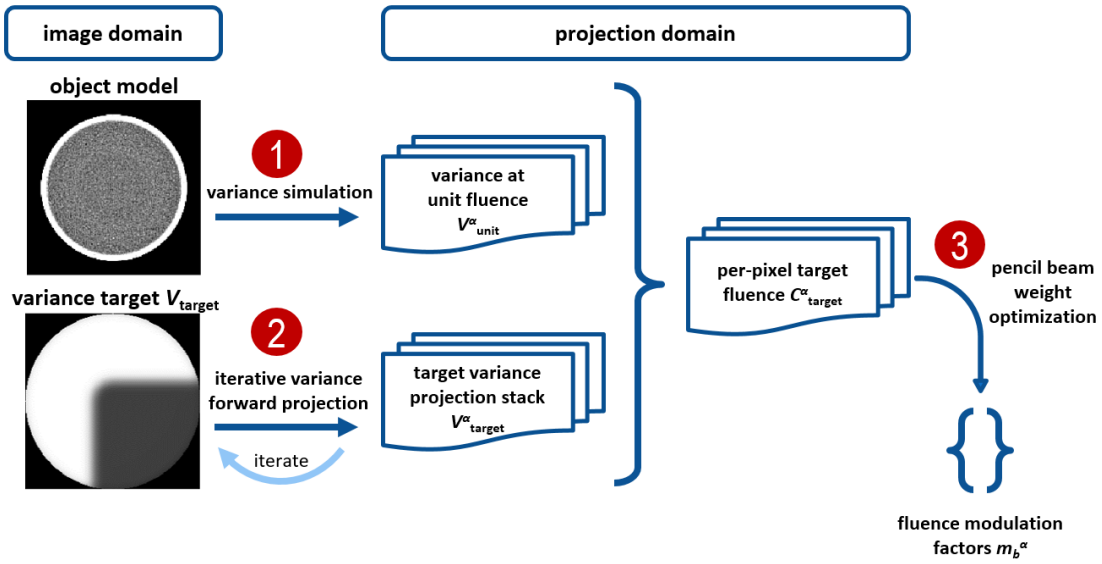

Figure 8.1: Workflow for optimization of fluence modulation factors m_b^α, given an object model and a variance target V_{target}.

Step 1: Variance at unit fluence prediction

To find variance levels at unit fluence for a given phantom, we employed a Monte Carlo simulation using the beam model described in section 8.2.4 and $m_b^\alpha = 1$ for all pencil beams and rotation angles. This step requires an object model according to section 8.2.3 and resulted in counts $C^\alpha_{\text{unit}}(u,v,d)$, which were reduced compared to the reference counts $F_{\text{unit}}(u,v,d) = \sum_b F_b(u,v,d)$ due to interactions with the object. For every point (u,v,d) in the projection, a set of $n = C^\alpha_{\text{unit}}(u,v,d)$ WEPL, $\{p\}$, was found such that the voxel around (u,v,d) was crossed by the most likely path of each of the selected protons (Rit et al. 2013). The unit fluence variance was then the squared error of the mean

$$V^\alpha_{\text{unit}}(u,v,d) = \text{Var}[\{p\}]/C^\alpha_{\text{unit}}(u,v,d). \tag{8.4}$$

Given a variance projection stack $V^\alpha_{\text{unit}}(u,v,d)$ the corresponding image variance $V_{\text{unit}}(x,y,z)$ can be calculated analytically as reconstruction was performed using filtered backprojection. Please refer to previous publications (Dickmann et al. 2019, Rädler et al. 2018) for details about variance calculations for pCT and for variance reconstruction in general (Wunderlich & Noo 2008).

Step 2: Iterative variance forward projection

Finding a stack of variance projections $V^\alpha_{\text{target}}(u,v,d)$ whose variance reconstruction (Rädler et al. 2018) yields a given image variance target $V_{\text{target}}(x,y,z)$ is a problem with a large set of solutions. We therefore aimed to find the inverse operation of variance reconstruction (Wunderlich & Noo 2008), i.e. a "variance forward projection." An initial guess $V^\alpha_0(u,v,d)$ could be obtained by performing ray–tracing (Joseph 1982) through the image variance target $V_{\text{target}}(x,y,z)$ followed by a ramp–filtration. The additional filtration was motivated by the fact that variance reconstruction is very close to a simple unfiltered backprojection (Wunderlich & Noo 2008). Since ray–tracing is the inverse operation to filtered backprojection, an additional ramp–filtration was required. While such forward– and backprojection yield V_{target} again, this often yields unphysical negative variance projection values and amplifies noise. Therefore, a median filter was applied to the ramp–filtered projections followed by thresholding to positive values.

To minimize the error introduced by thresholding, we employed an iterative approach by applying variance reconstruction to the i–th set of variance projections $V^\alpha_i(u,v,d)$ yielding a variance volume $V_i(x,y,z)$. Again, using ray–tracing, the difference volume $V_{\text{target}}(x,y,z) - V_i(x,y,z)$ was forward–projected and added to the current stack of variance projections. In every iteration, variance projection values were forced to be positive. This will converge to a set of physical (i.e. strictly positive) variance projections that yield an image variance approaching $V_{\text{target}}(x,y,z)$.

Step 3: Fluence optimization

By definition, the variance projection values in equation 8.4 are inversely proportional to the number of contributing protons C. Therefore, the pixel–wise counts required to achieve

the variance target could be calculated as $(V^{\alpha}_{\text{unit}}/V^{\alpha}_{\text{target}}) \cdot C^{\alpha}_{\text{unit}}$. However, for low counts, we need to consider that C follows a Poisson distribution (contrary to a normal distribution at sufficiently high counts) and therefore an additional correction function

$$k(C) = C \cdot \sum_{n'=1}^{\infty} P_C(n') \cdot \beta_{n',C} = C^2 \sum_{n'=1}^{\infty} \frac{P_C(n')}{n'} \tag{8.5}$$

needs to be introduced, where $P_C(n') = C^{n'} \exp(-C)/n'!$ is the Poisson probability of detecting n' protons instead of the expectation value of C and $\beta_{n',C} = C/n'$ is the relative change in variance if n' instead of C protons were detected. The function $k(C)$ was stored in a lookup table for all relevant integer values of C up to 300 by numerically calculating the infinite sum for 1000 summands. Since $\lim_{C \mapsto \infty} k(C)/C = 1$ and $k(300)/300 = 1.0033$ we assumed $k(C) = C$ for all $C > 300$. Furthermore, $k(C)$ was thresholded to return at least $C_{\min} = 8$ protons to avoid detector elements with missing information.

We used an optimization according to equation 8.3 to find pencil beam weights $w^{\alpha}_{b}(C^{\alpha}_{\text{target}})$ which achieve the pixel–wise projection counts target of

$$C^{\alpha}_{\text{target}}(u,v,d) = k\left[\frac{V^{\alpha}_{\text{unit}}(u,v,d)}{V^{\alpha}_{\text{target}}(u,v,d)} \cdot C^{\alpha}_{\text{unit}}(u,v,d)\right]. \tag{8.6}$$

Due to the fact that C^{α}_{unit} and $C^{\alpha}_{\text{target}}$ are both affected by interactions with the object, the optimization also needed to be performed for unit fluence allowing for an elimination of the effect of attenuation and scattering. This resulted in fluence modulation factors

$$m^{\alpha}_{b} = \frac{w^{\alpha}_{b}(C^{\alpha}_{\text{target}})}{w^{\alpha}_{b}(C^{\alpha}_{\text{unit}})} \tag{8.7}$$

with numerator and denominator as defined in equation 8.3. Due to the normalization, these factors were corrected for interactions with the object and thus could be used to simulate an FMpCT scan according to section 8.2.4.

Reference approach

A simpler approach for fluence field optimization, which was used in previous works (Dedes et al. 2018), is to perform a binary modulation with two fluence levels. In image domain, a ROI is defined as a set of voxels that should be imaged with high fluence. A pencil beam is assigned a high imaging fluence if its central axis intersects the ROI, and a low imaging fluence otherwise. The fluence modulation factors were

$$m^{\alpha}_{b} = \begin{cases} 1 & \text{if intersecting} \\ \gamma & \text{otherwise} \end{cases}, \tag{8.8}$$

where $0 < \gamma < 1$ is the modulation strength, which was chosen to be equal to the contrast of the variance prescription of the proposed method.

8.2.6 Simulation study

In a simulation study we prescribed three different image variance targets, which can be appreciated in figure 8.2: (1) constant variance V_{ROI} throughout the imaged object; (2) FMpCT prescription A (variance V_{ROI} inside one quadrant of the imaged object and $4 \cdot V_{\mathrm{ROI}}$ elsewhere); and (3) FMpCT prescription B (V_{ROI} inside a central rectangular region and $4 \cdot V_{\mathrm{ROI}}$ elsewhere). Variance targets are used in step 2 of the proposed algorithm, and therefore independent of the imaged object. In agreement to previous investigations (Dedes et al. 2018) the prescription contrast of 4 was chosen such that it is higher than the variance dynamic range of a unit fluence scan (Dickmann et al. 2019), but reasonably achievable without expecting regions with vanishing counts or distortions of RSP accuracy.

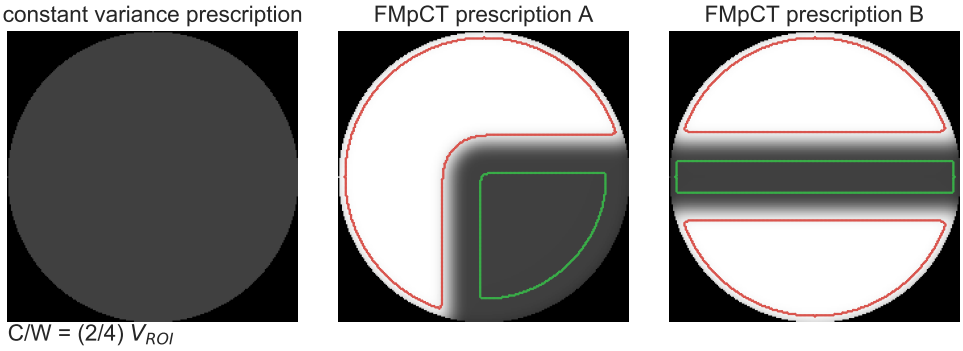

Figure 8.2: Three different image variance targets for the simulation study. The ROI region and the out–of–ROI region are indicated in green and red respectively. The display center (C) and window (W) is noted below the figure.

Previous investigations (Dickmann et al. 2019) have shown that a uniform fluence does not yield a constant variance for pCT. Therefore, the constant variance prescription is the most dose–efficient image, if the complete image is required for diagnosis. Prescriptions A and B model two treatment scenarios, where the treatment beam path is coming from 0 and 90 degrees in A and from 90 and 180 degrees in B. Prescriptions were slightly blurred as sharp gradients in image variance cannot be achieved due to the ramp filtration involved in reconstruction. Throughout this work we use the nomenclature “constant”, “A” and “B” to refer to the three prescriptions.

For all phantoms we first simulated a high dose unit fluence dataset with $m_i^{\alpha} = 1$. The mean incident proton fluence was chosen to be $133\,\mathrm{mm}^{-2}$ such that it yielded a typical imaging dose for pCT of about 1.4 mGy (Schulte et al. 2005), when summed over all projections. We then chose the value of V_{ROI} for each phantom as the 95^{th}–percentile value of the variance in the unit fluence scan. For the water phantom this was $V_{\mathrm{ROI}} = 4.61 \times 10^{-4}$, for the CTP404 phantom $V_{\mathrm{ROI}} = 5.89 \times 10^{-4}$, and for the head phantom $V_{\mathrm{ROI}} = 11.96 \times 10^{-4}$. These values are consistent with previous studies (Dickmann et al. 2019).

For the CTP404 phantom RSP values of the phantom body and of two inserts inside the

ROI were evaluated and compared to the unit fluence scenario. The body consisted of epoxy (RSP = 1.144), and inserts were made from Teflon (RSP = 1.791) and polymethylpentene (RSP = 0.883). RSP values were calculated with GEANT4 at a proton energy of 150 MeV and agreed with previous experiments (Giacometti et al. 2017a).

For a fair comparison of imaging doses, we computed the 95$^{\text{th}}$–percentile variance value v_{95}^{ROI} inside the ROI (inside the whole phantom for unit fluence) and calculated a linear correction factor $\eta = v_{95}^{\text{ROI}}/V_{\text{ROI}}$. Doses and counts were multiplied by η, variances were multiplied by $1/\eta$. The choice of the 95$^{\text{th}}$–percentile value over the mean or the maximum value is a compromise between the requirement that variances should be at V_{ROI} or lower, and tolerating outliers. As the water and the CTP404 phantom were thin, the percentile value was calculated only within the displayed central slice. For the head phantom, which covered the entire height of the detector aperture, it was calculated over the full volume. To avoid the variance evaluation being dominated by increased noise at the hull of the phantom as discussed in previous works (Dickmann et al. 2019, Rädler et al. 2018), we determined the shape of the hull by setting an RSP threshold of 0.5 and eroding the hull by 7 mm. Values outside this region were disregarded. The ROI region and the out–of–ROI region are indicated in figure 8.2 for fluence modulations A and B.

8.3 Results

8.3.1 Gaussian pencil beam model

In an experimental dataset without phantom we determined the beam spreads of the Gaussian beam model to be $\sigma_u = (4.04 \pm 0.08)\,\text{mm}$ and $\sigma_v = (5.24 \pm 0.09)\,\text{mm}$. The divergence was $\delta_u = (5.2 \pm 0.6)\,10^{-4}\text{mm}^{-1}$ and $\delta_u = (5.8 \pm 1.4)\,10^{-4}\text{mm}^{-1}$. The beam spread in the u direction was significantly smaller compared to the beam spread in v direction. Divergence in the u and v direction did not differ outside of the uncertainty bounds. The distances from the isocenter to a virtual source were $1/\delta_u = (1.9 \pm 0.2)\,\text{m}$ and $1/\delta_v = (1.7 \pm 0.4)\,\text{m}$, which agrees with the position of the scanning magnets, which is 1.8 m from the isocenter. The stated parameters were used in all following evaluations.

8.3.2 Variance optimization

Iterative variance forward projection

For step 2 of the proposed method, figure 8.3 (a) shows error measures as a function of the iteration number. The RMS error as well as the mean error between the current variance volume $V_i(x, y, z)$ and the variance target $V_{\text{target}}(x, y, z)$ are calculated within the field–of–view. The fastest convergence is observed for the constant variance prescription, while both FMpCT prescriptions A and B show a remaining RMS error that only reduces slowly in every iteration. The mean error quickly drops to zero within the first iterations. The relative change in RMS error for all prescriptions was below 1 % per iteration when they were stopped. Figure 8.3 (b) to (d) show $V_i(x, y, z)$ for prescription B at three different

iterations. At iteration 20, the high–variance region has reached the correct value, while in the low–variance region artifacts remain, but decrease up to the last iteration.

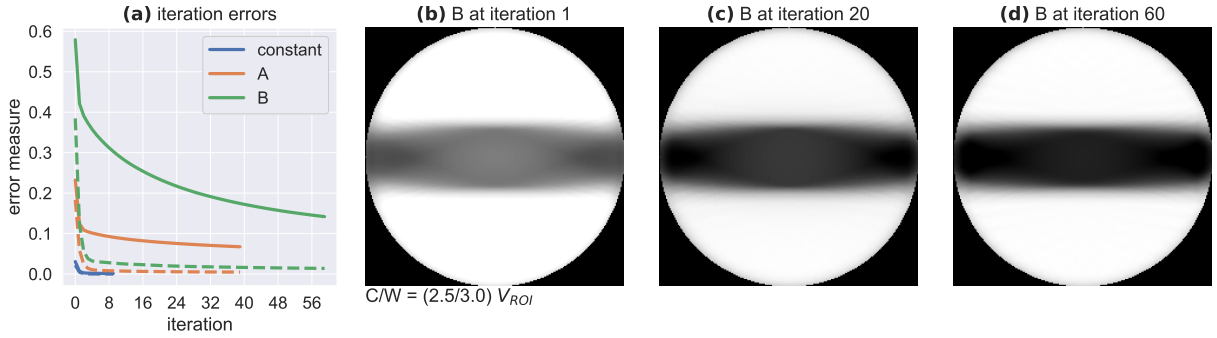

Figure 8.3: (a) Root–mean–square error (solid) and mean error (dashed) as a function of the iteration number of the three image variance targets. (b) – (d) Reconstructed variance volumes for prescription B for different iterations. The display center (C) and window (W) is noted below the figure.

Fluence optimization

To validate the use of the correction function $k(C)$, figure 8.4 shows $k(C)/C$ together with the relative increase of the image variance V_C at mean counts C. The relative increase is calculated as $(V_C \cdot C)/(V_{C_\infty} \cdot C_\infty)$ for $C_\infty = 310$ for simulated pCT data. Both curves agree, which shows that variance increases over–proportionally for low counts and that the correction function $k(C)$ describes this well.

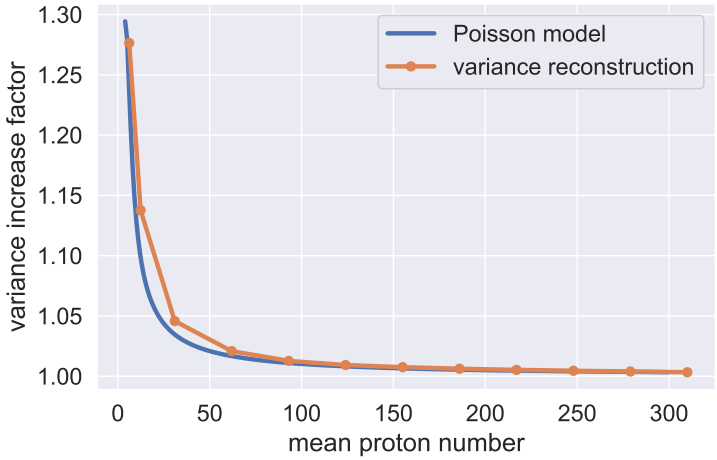

Figure 8.4: Overproportional increase of image variance with decreasing counts in a simulation with varying mean proton number C and agreement with the fluence correction function $k(C)/C$.

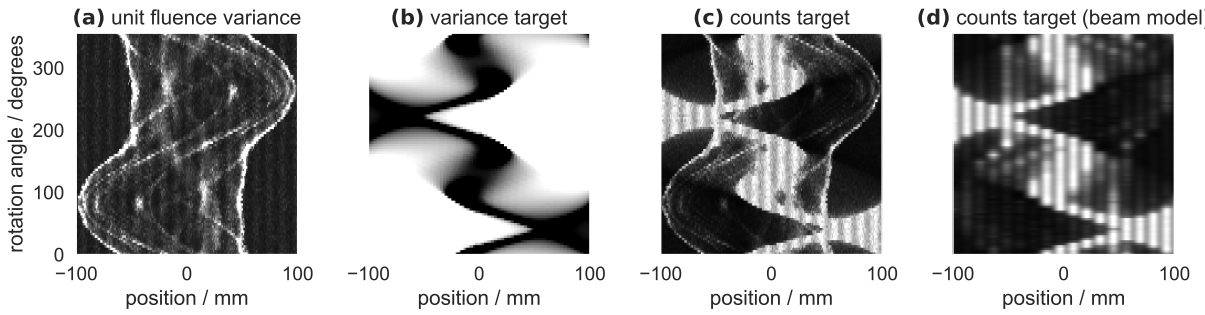

Figure 8.5: Intermediate results of the fluence optimization process for the pediatric head phantom and the orthogonal beams variance target: (a) unit fluence variance $V^{\alpha}_{\text{unit}}(u,0,0)$, (b) variance target $V^{\alpha}_{\text{target}}(u,0,0)$, (c) pixel–wise target counts $C^{\alpha}_{\text{target}}(u,0,0)$, and (d) target counts as fitted by the beam model. Data are shown as a function of the rotation angle α and the detector coordinate u. For display a center of $0.4\,\text{mm}^2$ and a window of $0.8\,\text{mm}^2$ has been applied for variances, and a center of 80 and a window of 160 for counts.

Figure 8.5 shows intermediate steps of the fluence optimization for the pediatric head phantom and variance prescription A. All projection data are shown as sinograms plotted as a function of the detector position in u direction and the rotation angle. Only data for $v = d = 0$ are shown. In figure 8.5 (a), variance at unit fluence $V^{\alpha}_{\text{unit}}(u,0,0)$ is shown (step 1 of the algorithm), which is high at the periphery of the object and around heterogeneities, as discussed in previous works (Dickmann et al. 2019). Figure 8.5 (b) shows the variance target $V^{\alpha}_{\text{target}}(u,0,0)$ as a result of the iterative optimization (step 2). Figure 8.5 (c) shows the pixel–wise counts target for fluence modulation $C^{\alpha}_{\text{target}}(u,0,0)$ (step 3) as given by equation 8.6. Parts of the variance target in (b) are assigned a value of 0, and receive the unit fluence in (c). In figure 8.5 (d), the counts target is fitted by the pencil beam model to get the weights required for fluence modulation (also step 3). This can be calculated as $\sum_b w^{\alpha}_b F_b(u,v,d)$. Some small features of (c) are not present in (d) if they are smaller than the extension of a pencil beam.

8.3.3 Simulation study

Figures 8.6 and 8.7 show simulated fluence modulations for all phantoms. RSP, variance and dose maps are shown together with the counts sinograms. For the water phantom imaged with unit fluence (figure 8.6 (a)), counts and dose were homogeneous throughout the phantom, variance was reduced in the center. This reduction was compensated in figure 8.6 (b) for the constant variance target, where instead counts and imaging dose were reduced in the center and variance was homogeneous across the phantom, except for a steep increase at the hull. The fluence modulations in figure 8.6 (c) and (e) for variance targets A and B can already be appreciated in the RSP maps. Variance levels followed the prescription with a sharp gradient. For prescription A some streaks of high variance were observed within the ROI. Using the reference approach in figure 8.6 (d) and (f), conformity

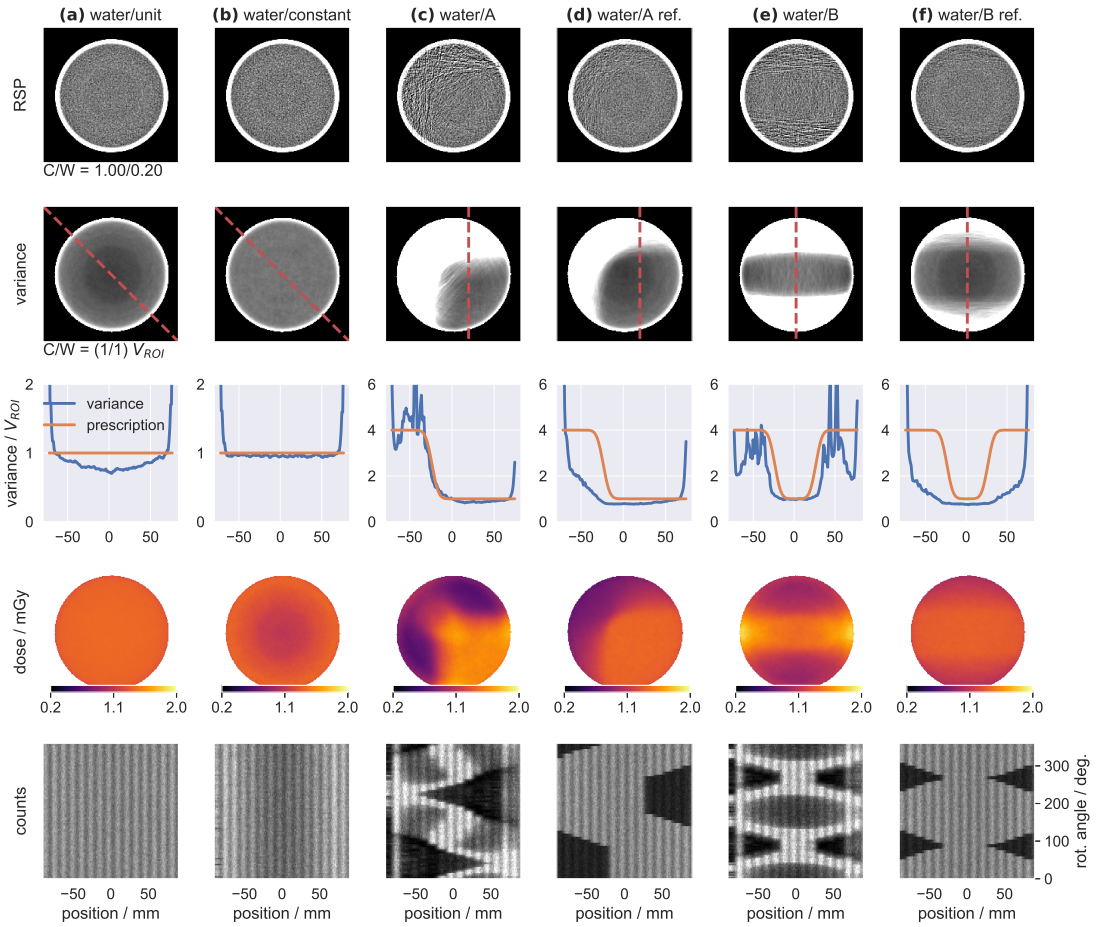

Figure 8.6: Simulation study for the water phantom and variance targets as indicated in the titles. Sinograms are shown for $v = d = 0$. Center (C) and window (W) settings for display of RSP and variance values are given.

of variance and dose maps with the ROI was degraded, in particular for prescription B, where variance and dose are at the same level as in the unit fluence scan for most of the phantom and the change in variance cannot be seen in the RSP maps. In the counts sinograms, regions of increased counts roughly agreed with those using the optimization, but were uniform, as required. Instead, using the optimization, a heterogeneous counts pattern was observed.

For the CTP404 phantom (figure 8.7 (a,b)) and the head phantom (figure 8.7 (c,d)), variance increased around heterogeneities both in unit fluence and fluence–modulated scans. For the head phantom in particular the palate exhibited locally elevated variance levels. The fluence modulation with prescription A was less conformal, compared to those of the water phantom. In particular for the CTP404 phantom variance contrast was impaired. Counts sinograms for prescription A in figure 8.6 (c) and figure 8.7 (b,d) are similar, but phantom–specific differences are noticeable.

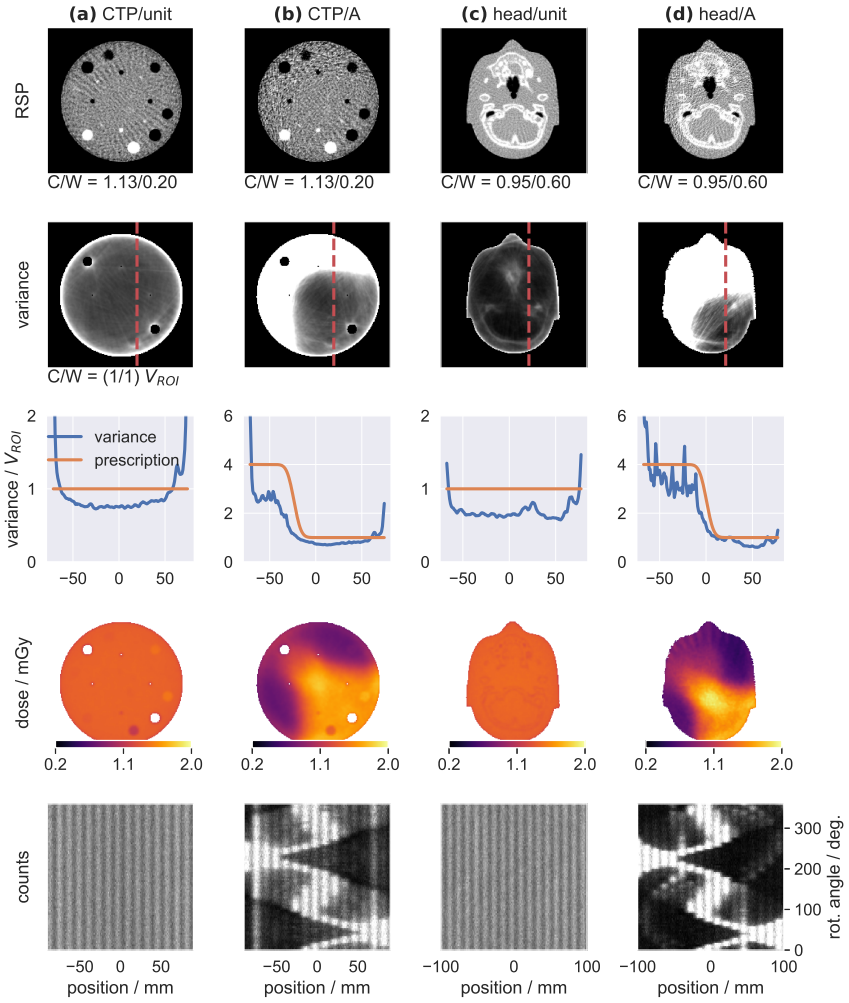

Figure 8.7: Simulation study for the CTP404 and the head phantom, and variance targets as indicated in the titles. Sinograms are shown for $v = d = 0$. Center (C) and window (W) settings for display of RSP and variance values are given.

Mean imaging doses are summarized in figure 8.8, where for fluence modulations the mean dose over the whole phantom as well as mean doses in the ROI region and the out–of–ROI region are reported. For the water phantom, prescribing constant variance resulted in a dose reduction of 8.9 % compared to the unit fluence dose. For the region–of–interest fluence modulations, dose saving outside the ROI was 40.5 % for prescription A and 25.7 % for prescription B. Using the simple reference approach, dose reductions were less pronounced and dropped to 29.2 % and 13.2 % respectively. For the FMpCT prescription A and the CTP404 phantom as well as the head phantom, dose savings outside the ROI were slightly lower compared to the 40.5 % of the water phantom (35.4 % and 38.9 % respectively). For all phantoms, fluence modulations A and B achieved a lower dose outside the ROI compared to the unit fluence, but after normalization with η required a higher

dose inside the ROI by 9.2 % to 19.2 %. Doses inside the ROI were approximately constant for the reference approach. Mean doses over the whole phantom were reduced by 7.2 % to 13.1 % using the reference approach and by 9.8 % to 18.6 % for the FMpCT optimizations.

For the CTP404 phantom, the two inserts and the body inside the ROI had an RSP value of 1.776, 0.881, and 1.143, compared to 1.776, 0.879, and 1.143 for the unit fluence case.

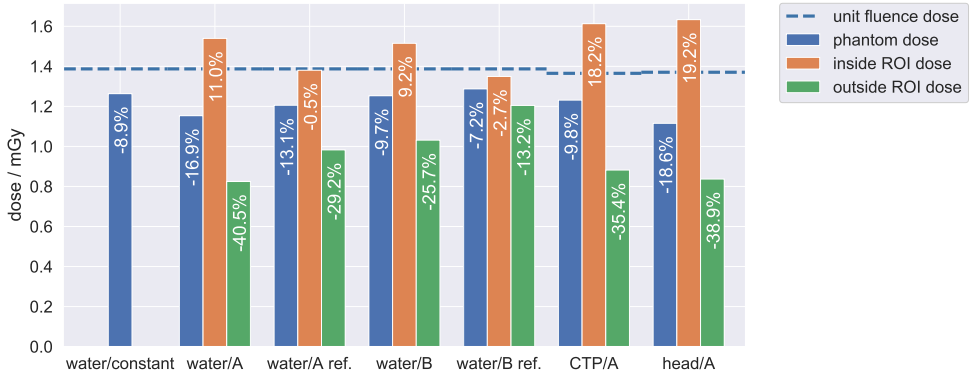

Figure 8.8: Mean imaging doses for the simulation study. The dashed line indicates the unit fluence dose while bars show the average phantom dose and doses inside and outside the ROI. The relative dose change compared to unit fluence dose is given inside the bars. Unit fluences were equal for all phantoms, but unit fluence doses differed slightly.

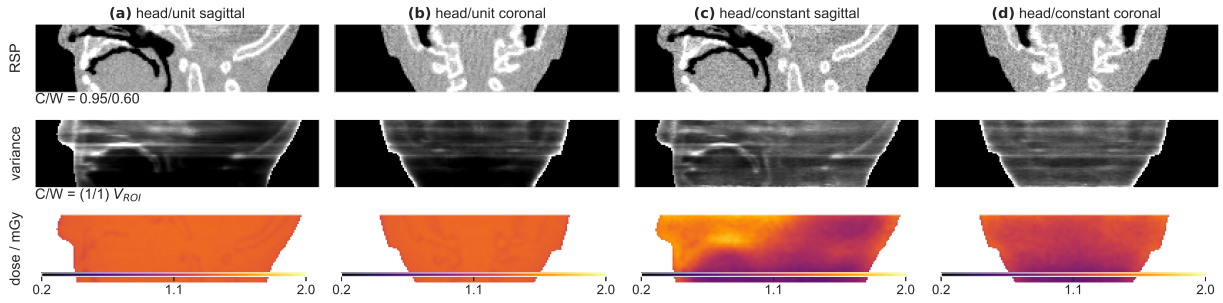

Figure 8.9: Simulation study for the pediatric head phantom with unit fluence (a), (b) and the constant variance target (c), (d). Row by row, the RSP, variance, and imaging dose are shown. Sagittal and coronal views are shown. Center (C) and window (W) settings for display of RSP and variance values are given.

Figure 8.9 shows the head phantom with unit fluence (a,b) and for the constant variance target (c,d) both in a sagittal view (a,c) and a coronal view (b,d). Dose is homogeneous for the unit fluence imaging, but the variance is notably lower in the back of the head and around the spinal cord compared to regions around the palate and the nasal cavities. These variations were compensated for in the fluence modulations achieving more homogeneous variance levels at reduced doses in regions where variance was low for unit fluence. Mean

dose over the whole phantom was 1.15 mGy compared to 1.37 mGy in the unit fluence case (16.0 % reduction). Around the palate and the nasal cavities, dose is increased in the fluence–modulated scan, which is not expected and may be due to the normalization by η.

8.4 Discussion

8.4.1 Gaussian pencil beam model

We found parameters of a Gaussian pencil beam model that allowed us to describe pencil beams at arbitrary fluences and positions. This is a key component of the fluence modulation scheme, as it allows to find a linear combination of a regular grid of pencil beams that achieves the required counts as calculated by our algorithm. Uncertainty bounds for fits in v direction were consistently larger than those in u direction, in particular for the divergence parameter δ. This was because the detector aperture is smaller in u direction and less data points were available. The beam spread σ_v was significantly larger than σ_u. While this anisotropy is not expected for clinical operation, it may have been caused by operating the beam line in research mode and modifying beam optics settings to keep proton fluence low and viable for the scanner. For future experimental studies, certain model parameters, such as the beam energy spread, may require adjustment to exactly match experimental variance levels.

8.4.2 Variance optimization

Iterative variance forward projection

Using an iterative approach, we calculated stacks of variance projections that yield a desired variance map in image space. Depending on the complexity of the variance prescription, this required a different amount of iterations and a non–zero RMS error remained. The easiest case (constant variance target), did not require negative variance values (a constant stack of variance projections would yield a constant image variance) and therefore converged quickly. The two inhomogeneous variance targets A and B did suffer from the positivity requirement and therefore only slowly converged towards a reduced RMS error. While with negative variance values, a (close to) zero RMS error would be possible, this was not the case when requiring physical variance values. The variance target stacks therefore already contained an inherent error, which impacted the achievable variance contrast. However, it did not impact fluence modulation in general, as the prescribed fluences could be rescaled, such that V_{ROI} was achieved inside the ROI.

Fluence optimization

We calculated the counts target according to equation 8.6, which could run into a lower and an upper limit. Firstly, to ensure that that data was available throughout the projection, we required at least $C_{\mathrm{min}} = 8$ protons in every pixel. Secondly, to avoid unreasonably

high imaging doses, only fluence modulation factors $m_b^\alpha \leq 1$ were allowed, even if the variance target from the previous step (iterative variance forward projection) was zero. This was relevant in particular at the hull of the object, which is also a limited area to be traversed by a therapeutic proton beam. Again, both limits impacted achievable variance contrast, but V_{ROI} could be achieved in the ROI by rescaling with η. Due to the limitation to pencil beams with a finite size, small variance features were averaged out, which may impact homogeneity of the achieved variance, in particular for phantoms with strong heterogeneities.

8.4.3 Simulation study

We simulated FMpCT scans for different phantoms and variance targets demonstrating two possible applications for dose reduction using fluence modulation: (1) for achieving constant variance throughout the object and (2) concentrating imaging dose in a high image quality ROI and reducing it elsewhere.

The dose reduction for constant variance with the homogeneous water phantom was 8.9 %, which already is considerable. As shown in previous investigations (Dickmann et al. 2019), variance for hetereogeneous phantoms is dominated by multiple Coulomb scattering, which depends on the local heterogeneity of the phantom. Therefore, variance maps of the head phantom in coronal and sagittal views were varying greatly. Assuming that good image quality is required in the complete field–of–view, a fluence–modulated scan can reduce the imaging dose by 16.0 % compared to a unit fluence scan, without any loss of diagnostic value. Equivalently, the signal–to–noise ratio could have been improved by 35 % at equal dose.

For all phantoms and two different image variance targets we could demonstrate considerable dose savings of 25.7 % to 40.5 % outside of the ROI. At the same time, the imaging dose inside the ROI increased compared to the unit fluence acquisition. Assuming that the ROI agrees with the treatment beam path and that treatment doses are several orders of magnitude higher than imaging doses, this increase is probably not relevant. At the same time, proton therapy allows for minimal doses outside the treatment beam path, requiring that this advantage is not compromised by frequent imaging. Mean imaging doses over the whole phantom were reduced for all combinations of phantoms and variance targets. Using a sensitometric phantom we showed that RSP accuracy is not compromised by fluence modulation. RSP errors were comparable for modulated and un–modulated scans, and all below 1 %, which is within the magnitude expected from literature (Schulte et al. 2005, Giacometti et al. 2017a, Dedes et al. 2019).

Imaging doses in fluence–modulated scans showed local increases and doses partially spilled out of the ROI. This may have impaired results in this study and could be caused by the fact that optimization was exclusively performed with a variance objective. Future studies should therefore include a dose objective outside of the ROI while keeping the variance objective inside the ROI, further developing ideas from studies for x–ray CT (Bartolac et al. 2011). Moreover, the optimal choice of the contrast in the image variance prescription should be studied in the future, but is out of scope for this work.

Using a simple intersection–based approach also showed dose savings compared to unit fluence acquisitions. However, dose savings were considerably less compared to the optimized FMpCT scans and conformity of variance with the prescription was degraded. By construction, a prescription of constant variance is not possible with this approach.

Future work should also address the impact of iterative image reconstruction, which is frequently used for pCT imaging (Penfold et al. 2009, Penfold et al. 2010, Hansen et al. 2014, Hansen et al. 2016, Schultze et al. 2019). In contrast to the direct filtered backprojection algorithm used in this study, iterative reconstruction employs a regularization method (typically total variation), which reduces noise and whose optimal weight depends on the object and the fluence level (Tian et al. 2011). While most fluence modulation studies for x–ray CT have been performed using filtered backprojection (Graham et al. 2007b, Bartolac et al. 2011), a first study (Gang et al. 2019) investigated a joint optimization of the fluence field and a spatially varying regularization parameter in the iterative reconstruction. For pCT, a comparison of iterative and direct reconstruction (Hansen et al. 2016) showed comparable image quality. Preliminary work of the authors using an iterative reconstruction algorithm (Hansen et al. 2014) and fluence modulation suggests feasibility of combining the two methods for pCT.

8.5 Conclusion

We developed a novel method for fluence–modulated proton computed tomography using pencil beam scanning and demonstrated its feasibility in a simulation study. Dose reductions achieved by prescribing uniform variance were considerable, in particular for an anthropomorphic head phantom. This suggests the need for employing non–uniform fluence patterns in future pCT studies, whenever dose efficiency is a key requirement. Furthermore, the proposed method allows us to prescribe arbitrary image variance targets, which were shown to further reduce imaging dose outside of a given region–of–interest. This can be of particular interest in the context of particle therapy and allow for daily imaging at a reduced imaging dose to healthy tissue outside of the treatment beam path.

Acknowledgments

This work was supported by the German Research Foundation (DFG) project #388731804 “Fluence modulated proton computed tomography: a new approach for low–dose image guidance in particle therapy” and the DFG’s Cluster of Excellence Munich–Centre for Advanced Photonics (MAP), by the Bavaria–California Technology Center (BaCaTeC) project Nr. 28 [2015-2] and by the Franco–Bavarian university cooperation center (BayFrance).

Chapter 9

Experimental realization of fluence optimization for proton CT

This chapter contains the work published in Physics for Medicine and Biology, volume 65, issue 19, page 195001, in September 2020, with the title *Experimental realization of dynamic fluence field optimization for proton computed tomography*, by Dickmann et al. (2020a) – (senior author: Dedes)

9.1 Introduction

Radiotherapy using protons allows to precisely deliver the planned dose to the tumor while minimizing radiation exposure of critical structures outside of the treatment beam path. Proton therapy, therefore, is believed to have better outcome with less severe side effects compared to conventional radiotherapy using x–rays for certain anatomical sites (Weber et al. 2012, Park et al. 2015, Nakajima et al. 2017). During treatment, protons stop inside the tumor, where, just before stopping, their energy loss spikes within the so–called Bragg peak. While this allows to achieve sharp dose gradients and to spare healthy tissues, proton therapy is sensitive to range uncertainties (Paganetti 2012) and thus requires frequent and accurate imaging (Engelsman et al. 2013, Landry & Hua 2018). The current clinical practice for acquiring spatial maps of the relative to water proton stopping power (RSP) of the patient is to employ x–ray CT images and to apply a calibration converting interaction of photons with matter (attenuation) to interaction of protons with matter (RSP). This yields range errors of up to 3 % (Yang et al. 2012), which need to be considered with additional dose margins and conservative beam directions, and necessarily increase the dose to healthy tissue.

To avoid conversion errors, RSP maps can directly be acquired using proton CT (pCT), as originally proposed by Cormack (1963) and later experimentally realized by Hanson et al. (1977). A pCT scanner measures the residual energy of protons after traversing the patient or an object (for a review of contemporary pCT scanners see Johnson (2017)). Since

protons gradually lose energy, and given that the initial proton energy is known, the energy loss is directly linked to a water equivalent path length (WEPL), that is the path length in water, along which a proton would lose the same amount of energy. For a proton crossing an object, the WEPL is a line integral through the RSP along the curved path of the proton in the same way that the logarithmic intensity is a line integral through the attenuation coefficient for x–ray CT. Unlike in x–ray CT, however, line integrals are not along straight lines and spatial resolution can suffer severely if not accounted for correctly (Krah et al. 2018). Therefore, pCT scanners typically also measure proton positions and directions prior and after the patient allowing to estimate the proton's trajectory and incorporating it in the reconstruction algorithm (Li et al. 2006, Rit et al. 2013, Hansen et al. 2016). A recent study by Dedes et al. (2019) comparing the accuracy of state–of–the–art dual–energy x–ray CT to the performance of a prototype pCT scanner concluded that errors are on par between the two modalities and suggested that artifact reduction may considerably improve the pCT performance.

Besides having good RSP accuracy, pCT images may potentially be acquired at very low imaging doses of only $1 - 2\,\mathrm{mGy}$ per tomography (Hansen et al. 2013, Meyer et al. 2019) based on an estimate using a homogeneous phantom (Schulte et al. 2005). Dickmann et al. (2019) showed increased noise levels due to the heterogeneity in an anthropomorphic phantom suggesting that imaging doses may range from $2\,\mathrm{mGy}$ to $6\,\mathrm{mGy}$ for the same noise level in heterogeneous phantoms. A low imaging dose is especially important for particle therapy, where the delivery of the treatment is divided in up to 30 fractions and changes in positioning of the patient or movement of internal organs must be tracked by imaging as frequently as possible to ensure a safe delivery of the therapeutic dose (Landry & Hua 2018). However, imaging for online adaptation only requires knowledge about the RSP within the limited sub–volume covered by the treatment beams. To exploit this, we recently developed an algorithm for dose reduction with FMpCT, an imaging technique adapted from Bartolac et al. (2011) to pCT by Dedes et al. (2017). The optimization method (Dickmann et al. 2020) tries to achieve an inhomogeneous image noise prescription using dynamically modulated fluence fields. Modulation is achieved by pCT acquisition using a grid of pencil beams. The output of the optimization algorithm are relative weights for pencil beams within every projection of the tomography. Using a Monte Carlo simulation, a corresponding "planned noise distribution" can be calculated, which is the noise map expected for the optimized set of pencil beam weights. The Monte Carlo code models the geometry of an existing prototype pCT scanner (Giacometti et al. 2017a) and realistic pencil beams at a clinical facility (Dickmann et al. 2020). It was also validated for fidelity in terms of image noise (Dickmann et al. 2019) and RSP accuracy (Giacometti et al. 2017a, Dedes et al. 2019). We prescribed low image noise (high image quality) within the volume covered by typical treatment beam configurations, which we label the region of interest (ROI), and high image noise (low image quality) elsewhere. This resulted in a considerable dose saving outside of the ROI by up to $40\,\%$, which depends on the phantom and shape of the ROI. The dose saving was shown to be superior to the simpler intersection approach used by Dedes et al. (2019). However, since the noise–planning study employed the same Monte Carlo code as was used to generate data input to the optimizer, an experimental

validation was indispensable.

The aim of this work, therefore, was to experimentally validate and reproduce the computational results obtained by Dickmann et al. (2020). We employed the same fluence patterns and used the same three phantoms that were modeled there. To employ modulated fluence patterns, we interfaced the control system of a pencil beam scanning (PBS) beamline at the Northwestern Medicine Chicago proton center and obtained tomographies using the preclinical phase II pCT scanner (Johnson et al. 2016, Bashkirov et al. 2016). We evaluated our ability to accurately control image noise by comparing experimental image noise maps to the corresponding planned noise as well as to the prescription. In the process, we identified image artifacts stemming from the pencil beam acquisitions and operation of the beamline in research mode and removed them successfully with a correction method. Finally, we tested the accuracy of RSP values by comparing mean values and distributions inside and outside of the ROI.

9.2 Materials and methods

9.2.1 Experimental setup

Data for this work was acquired using the phase II pCT scanner developed by Johnson et al. (2016) and Bashkirov et al. (2016). As depicted in figure 9.1, the scanner consists of two tracking modules, one prior and one after the phantom, and a scintillating energy detector. Protons emitted from the PBS beamline are individually registered by four silicon strip detectors in each of the two tracking modules, resulting in coordinate measurement pairs prior and after the phantom. The tracking system maintains a detection efficiency above 99% at trigger rates up to 1 MHz for the entire field–of–view and is capable of recovering missed hits if at least three of the four position measurements are available (Johnson et al. 2016). It has also been shown to safely sustain hit efficiency at 400 kHz under increased local count rates when operating with pencil beams (Dedes et al. 2018). The two coordinate measurement pairs allow for an estimation of the proton's direction vector before and after the object, which helps to determine a most likely path through the object (Schulte et al. 2008). Eventually, the proton's residual energy is measured using a five–stage scintillating energy detector. Because the mean initial energy of the protons is known and they gradually lose energy when traversing the object, the residual energy measurement can be used to calculate the WEPL along the proton's path. This uses a calibration curve generated with an object of known geometry and RSP (Piersimoni et al. 2017). To allow for tomographic data acquisition, the phantom is mounted on a rotation stage.

The scanner is located at the Northwestern Medicine Chicago proton center, a clinical facility for proton therapy equipped with a beamline (PROTEUS®PLUS, *Ion Beam Applications SA*, Louvain-la-Neuve, Belgium) with full PBS capability. Delivery of the proton fluence using pencil beams was a key requirement for this study, as it allowed employing individually modulated proton fluence fields. While the scanner is typically operated using

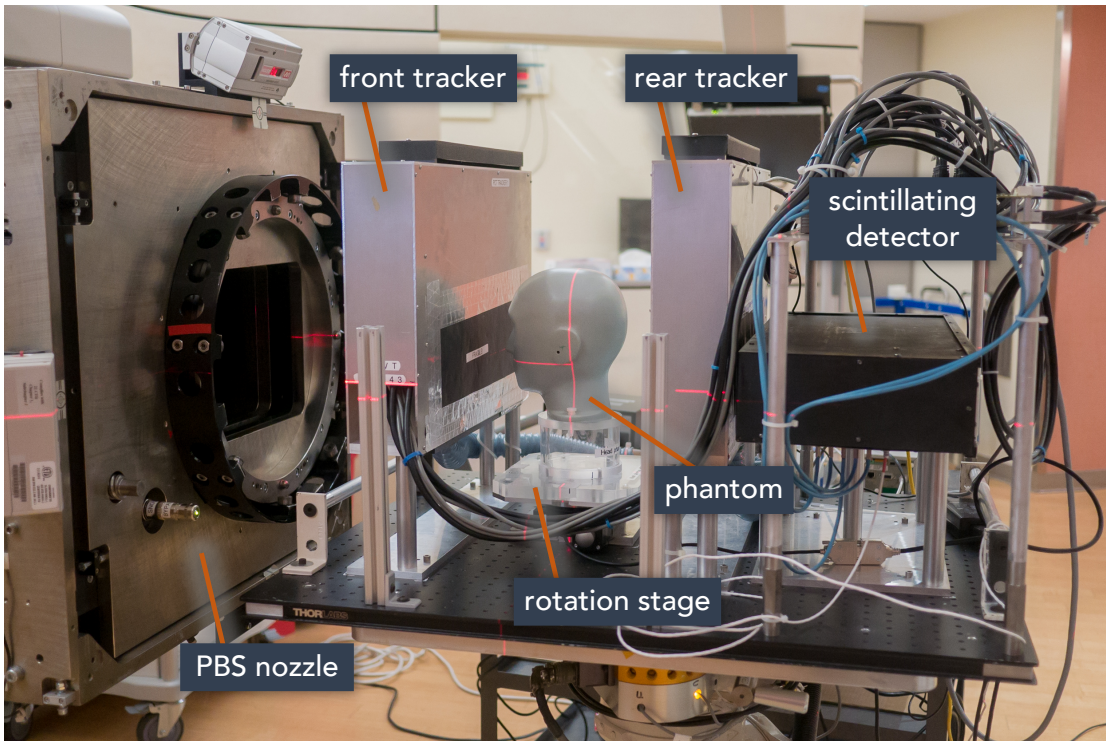

Figure 9.1: The phase II pCT scanner installed on a robotic arm at the Northwestern Medicine Chicago proton center. The important components of the scanner are indicated as well as the pencil beam scanning (PBS) nozzle.

a broad beam, data acquisition using PBS and static fluence fields has been shown to be feasible by Dedes et al. (2018) using protons and Volz et al. (2018) using helium ions (at another facility).

9.2.2 Image reconstruction and noise for proton CT

Data output from the scanner is in so–called "list–mode." Thus, for every registered proton, this list includes five quantities: position and direction vector prior to the object and subsequent to it as well as the WEPL inferred from the energy measurement. To obtain a three–dimensional map of the RSP from this data, WEPL information was binned to pixelated projections and averaged within each pixel. Prior to binning, data cuts were applied, removing protons with direction or WEPL information outside of a three standard deviation interval in each pixel of the projection, which is a common procedure in proton imaging (Schulte et al. 2005, Rit et al. 2013). The standard deviation in each pixel was calculated from the difference between the 30.85–th percentile value and the median, which corresponds to a 0.5–standard–deviation interval and is less sensitive to the non–Gaussian tails of the distribution. Data cuts eliminated about 25% of recorded protons in air and 15% inside a homogeneous phantom. Protons were predominantly rejected due

to overestimated WEPL information caused by spurious signals in the energy detector. A three–dimensional volume could be reconstructed using a special cone–beam filtered back-projection algorithm (Rit et al. 2013) making use of the most likely path information and thereby improving the spatial resolution in the image.

This study investigated the uncertainty of RSP values, i.e. image noise or image variance. This quantity could be calculated directly from the list–mode data by estimating the Gaussian variance of the WEPL values within each pixel of the projection and using variance reconstruction (Wunderlich & Noo 2008, Rädler et al. 2018). Projection noise and hence also image noise for pCT is predominantly affected by the incident proton fluence, i.e. the number of protons contributing to the mean WEPL value in one pixel of the projection. Therefore, modulation of the incident fluence was the principal lever for adjusting image noise. However, to carefully predict the absolute level of image noise in pCT for a given fluence level, various contributions had to be taken into account. A previous study by Dickmann et al. (2019) identified (1) the incident energy spread output from the accelerator, (2) energy straggling inside the object and (3) inside the detector, as well as (4) multiple Coulomb scattering, to be the main causes of image noise for pCT. In particular, multiple Coulomb scattering was found to strongly depend on the object's heterogeneity. Consequently, prior knowledge of the object's geometry is required to precisely predict the image noise level at a given incident fluence using a validated Monte Carlo simulation such as in Dickmann et al. (2019).

9.2.3 Fluence optimization algorithm

FMpCT may be performed by employing a regularly spaced set of pencil beams and individually modulating the fluence of each pencil beam resulting in inhomogeneous image noise distributions as originally shown by Dedes et al. (2017) and Dedes et al. (2018). To prescribe arbitrary image noise distributions, we employed the optimization algorithm of Dickmann et al. (2020), which requires an object model to calculate individual pencil beam weights. The algorithm consists of three steps, which are briefly summarized here and can be studied in detail in the original publication:

1. For a given phantom geometry, use a Monte Carlo simulation to predict variance levels in the projection for unit fluence (i.e., all relative weights set to unity).

2. Use an iterative method to calculate a stack of variance projections that yield the prescribed image noise map (thus transfer the problem from image space to projection space).

3. Calculate the required relative fluence as the ratio of the two variance quantities taking into account a non–linear inverse relationship between fluence and variance. Then, fit a pencil beam model to the required fluence to get relative pencil beam weights.

The algorithm thus outputs relative pencil beam weights resulting in a planned noise distribution, which can be calculated to check the agreement with the initial noise pre-

scription by employing the optimized pencil beam weights in a Monte Carlo simulation. Note that the planned noise was the best achievable noise distribution by our optimization algorithm and could differ from the prescription. Moreover, the optimizer tried to achieve peak noise equal to the prescription in the ROI and thus average noise may be below the prescription.

9.2.4 Phantoms

We used three phantoms in this study: a cylindrical water phantom, an anthropomorphic head phantom and a sensitometry phantom. The water phantom consisted of a cylindrical container made from polystyrene and was filled with distilled water. The head phantom (ATOM®, Model 715 HN, *CIRS Inc.*, Norfolk, USA) mimicks the head of a 5–year old child using tissue–equivalent materials. The sensitometry phantom was the cylindrical CTP404 module of the Catphan® 600 phantom (*Phantom Laboratory*, New York, USA) and features several inserts of different materials. To obtain the planned noise distributions and fluence patterns, phantom models were used in the Monte Carlo simulation by Dickmann et al. (2020), and details can be found therein. RSP values of the phantoms were characterized experimentally in other studies (Dedes et al. 2019, Giacometti et al. 2017b) and are also matched in the Monte Carlo engine by adjusting the mean excitation energy.

9.2.5 Noise prescriptions and pencil beam grids

We aimed at achieving the following image acquisitions: (1) unit fluence with all relative pencil beam weights set to one, which corresponded to the standard acquisition with PBS in pCT and served as a comparison, (2) a prescription of constant image noise V_{ROI} throughout the phantom, which was shown to reduce dose at equal peak–noise level, and (3) a ROI imaging task with a prescribed variance V_{ROI} inside one quadrant of the image (i.e. the ROI; corresponding to a potential two–beam treatment configuration in proton therapy), and a higher prescribed variance $4 \cdot V_{\mathrm{ROI}}$ outside the ROI, which allowed for dose reduction. In all evaluations we labeled the three cases (1) *unit*, (2) *constant*, and (3) *FMpCT*. The value of V_{ROI} was chosen to be the peak variance of the unit fluence acquisition and thus depended on the phantom (Dickmann et al. 2019). For the water phantom this was $V_{\mathrm{ROI}} = 4.61 \times 10^{-4}$, for the head phantom $V_{\mathrm{ROI}} = 11.96 \times 10^{-4}$ and for the sensitometry phantom $V_{\mathrm{ROI}} = 5.89 \times 10^{-4}$. A graphical representation of the FMpCT ROI is included in figures 9.3, 9.4 and 9.6. Note that the ROI corresponded to prescription *A* in Dickmann et al. (2020), which yielded imaging dose savings of 40% and 39% outside the ROI for the water and head phantoms, respectively.

The arrangement of the pencil beam grids was also chosen in agreement with the computational planning study. From an evaluation of experimental data, Dickmann et al. (2020) obtained an elliptical shape of the pencil beam with a full width at half maximum (FWHM) of $\mathrm{FWHM}_u = (9.5 \pm 0.2)\,\mathrm{mm}$ in the horizontal u–direction and $\mathrm{FWHM}_v = (12.3 \pm 0.2)\,\mathrm{mm}$ in the vertical v direction (the rotation axis was oriented along v). Pencil beams were interspaced by $\Delta_{\mathrm{PB},u} = 12\,\mathrm{mm}$ in u and by $\Delta_{\mathrm{PB},v} = 8\,\mathrm{mm}$ in v. This corresponded to

approximately $1.3 \cdot \mathrm{FWHM}_u$ in u and $0.65 \cdot \mathrm{FWHM}_v$ in v. In u the spacing was effectively also $0.65 \cdot \mathrm{FWHM}_u$, because we used a quarter pencil beam shift: the whole pencil beam grid was offset in u by $\Delta_{\mathrm{PB},u}/4 = 3\,\mathrm{mm}$. This compensated for the larger interspace in this direction, as pencil beam patterns at 0° and 180° were thus offset by $\Delta_{\mathrm{PB},u}/2$ with respect to each other, resulting in a smooth total fluence despite the larger interspace. This is analogous to the quarter detector pixel shift used to increase the sampling rate in x–ray CT. Note that such a quarter pencil beam shift could only be performed in the u direction, and that it depends on a precise alignment of the pencil beams with respect to the coordinate system of the scanner.

9.2.6 Interfacing the PBS beam line

We interfaced the control system of the PBS beam line at the proton center to deliver the optimized fluence plans, in order to experimentally validate that noise maps achieved with pencil beam weights output from the optimization algorithm yield the planned noise maps and are close to the prescribed image noise maps. This required transforming the pencil beam grid description (center coordinates and relative weights) to machine instructions. The fluence was modulated by changing the dwell time of each pencil beam spot, as in Dedes et al. (2018), and keeping the current constant at 1.3 nA. The instructions for the accelerator were thus set points for the scanning magnets as well as absolute dwell times. A program developed specifically for this study then converted those set points to currents applied to the scanning magnets, which required proprietary beam line information.

The maximum dwell time (the one corresponding to a relative weight of one) was chosen such that approximately four times (4×) the required number of protons was registered and the correct number of protons (1×) was selected in post–processing. In a first step, we acquired data at unit fluence without any phantom. From this data we later extracted the average number of hits within a central region of the front tracker. The same number was generated using data from the Monte Carlo simulation (with 1× fluence) and the acceptance ratio r_{accept} between experiment and simulation was calculated. For all subsequent scans, we then randomly accepted protons with a probability equal to r_{accept}, which was constant for all projections of all acquisitions, and rejected all other protons. Although the exact number of protons could have been delivered by fine tuning the maximum dwell time or the beam current on the spot, we opted for this conservative approach considering the limited available beam time.

To ensure that the pencil beam patterns were delivered in synchrony with the scanner rotation, the phantom was rotated at fixed time intervals large enough to allow for a manual initiation of each pencil beam pattern, which led to considerably longer scanning times than the 6 to 10 min that are typically needed for scans with this prototype pCT scanner when the beam is continuously on.

9.2.7 Image noise evaluation

We acquired unit fluence as well as FMpCT data for the water phantom and the head phantom. For the water phantom, a constant noise fluence modulation was additionally acquired. All acquisitions were performed on the same day, except for the unit fluence acquisition of the water phantom, which was acquired on a separate day. RSP and variance maps were calculated from the acquired data. All reconstruction parameters were chosen as in Dickmann et al. (2020). Image variance maps were compared to the planned noise distribution, as well as to the prescriptions. To illustrate the fluence patterns, we calculated so–called "counts sinograms" that display the number of protons (after data cuts) contributing to the calculation of the mean WEPL in each pixel of the detector and for each rotation angle.

9.2.8 RSP evaluation

To verify that the FMpCT acquisition does not compromise the RSP accuracy of pCT inside the ROI, we performed two scans with the sensitometry phantom: one with unit fluence and one with the FMpCT noise prescription. Inserts of the phantom as well as the phantom body were masked and distributions of RSP values inside the masks were analyzed and compared between the fluence–modulated and the uniform fluence scan. In agreement with Dedes et al. (2019), the mask of an insert was chosen to be a cylinder with a radius of 5 mm, which is half of the radius of the insert. The phantom's body is made from epoxy and inserts consist of polymethylpentene (PMP), low–density polyethylene (LDPE), polystyrene, acrylic, Delrin and Teflon, for which water–column RSP measurements can be found in Dedes et al. (2019). The ROI of the FMpCT scan encompassed the PMP and the Teflon inserts which have the lowest and highest RSP values respectively. We chose to investigate these inserts with the largest expected RSP error (Dedes et al. 2019) so that this evaluation can serve as a conservative upper limit estimate and because only two inserts could be accommodated in the ROI due to the geometry of the phantom.

9.2.9 Pencil beam WEPL correction

In retrospective analysis of the acquired data, RSP ringing artifacts were observed which we assumed to stem from a spatially varying incident energy distribution within each pencil beam of the grid, and which were more severe than the calibration artifacts discussed in Dedes et al. (2019). The distribution was assumed to be similar for all pencil beams. Since incident energy is not measured by the pCT scanner but instead assumed to be 200 MeV, a spatially varying distribution will distort the image.

To isolate this effect, we used a scan without phantom for which on average we expected a WEPL of 0.0 mm after data post processing. In a first step, we calculated the count rate over time and separated individual pencil beams at the drop in count rate between two spots. For each pencil beam p we then calculated the center of mass coordinate (u_p, v_p) at the second tracking plane of the front tracker module. For each proton, the coordinates

$(\tilde{u}, \tilde{v}) = (u - u_p, v - v_p)$ relative to the center of mass of the corresponding pencil beam p (again at the second plane of the front tracker) were determined. We then fitted a quadratic correction function

$$\Delta_{\mathrm{WEPL}}(\tilde{u}, \tilde{v}) = a + b_u\tilde{u} + b_v\tilde{v} + c_u\tilde{u}^2 + c_v\tilde{v}^2, \tag{9.1}$$

where a, b_u, b_v, c_u were c_v are free fitting parameters. The fitting parameters were found by minimizing the cost function

$$\mathrm{CF} = \sum_{n=1}^{N}(\Delta_{\mathrm{WEPL}}(\tilde{u}_n, \tilde{v}_n) - w_n)^2, \tag{9.2}$$

where the sum is over all N protons from all pencil beams, and $(\tilde{u}_n, \tilde{v}_n)$ is the coordinate of proton n relative to the center of mass of its pencil beam, and w_n its measured WEPL value.

To correct data of a subsequent phantom scan using equation 9.1, all pencil beams in the dataset needed to be separated and center of mass coordinates calculated for each of them. Then, a proton with WEPL w belonging to pencil beam p and with relative coordinates $(\tilde{u}, \tilde{v})$ was assumed to instead have a WEPL of $w' = w - \Delta_{\mathrm{WEPL}}(\tilde{u}, \tilde{v})$.

We performed reconstructions of the constant noise prescription for the water phantom both with and without pencil beam correction to show its effect on the data. To further illustrate the pencil beam spot and its energy distribution, we binned all protons that were used to find the correction function to a grid in relative coordinates $(\tilde{u}_n, \tilde{v}_n)$ and calculated the number of counts relative to the pencil beam center as well as their mean WEPL. From the binned counts distribution we also calculated the FWHM in both directions, which we refer to as $\mathrm{FWHM}_{u,\mathrm{exp}}$ and $\mathrm{FWHM}_{v,\mathrm{exp}}$.

9.2.10 Fluence sums

Since the acquisition with the quarter–shifted pencil beam grid described in section 9.2.5 depends strongly on a precise delivery of the fluence and misalignments could cause distortions of the resulting fluence patterns, we investigated if fluence was delivered as intended in the planning study by calculating "fluence sums": for every proton we estimated a most likely path $\vec{l}(d) = (u(d), v(d))$ based on the tracking information, where u and v are the lateral coordinates and d is the (signed) distance from the isocenter plane. For every projection n acquired at angle α_n we then calculated a three–dimensional counts map $C_n(u, v, d)$ which counts the number of most likely paths intersecting each voxel. The fluence sums $F(x, y, z)$ were then calculated as the sum over all C_n, where each C_n was rotated along the v–axis according to its rotation angle α_n and thus

$$F(x, y, z) = \sum_{n=1}^{N} C_n(x\cos\alpha_n + y\sin\alpha_n, -x\sin\alpha_n + y\cos\alpha_n, z), \tag{9.3}$$

where N is the number of projections. Fluence sums were calculated both for the experimental data and for the corresponding planning Monte Carlo data.

To investigate the robustness to changes of the spot size and the alignment of the fluence pattern, we performed an additional Monte Carlo planning study using the pencil beam $\mathrm{FWHM}_{u,\mathrm{exp}}$ and $\mathrm{FWHM}_{v,\mathrm{exp}}$ determined in section 9.2.9 instead of FWHM_u and FWHM_v used during optimization and noted in section 9.2.5. We added an artificial shift s_u in the u–direction to this data by calculating the fluence sums with the modified equation

$$F(x,y,z) = \sum_{n=1}^{N} C_n(x\cos\alpha_n + y\sin\alpha_n + s_u, -x\sin\alpha_n + y\cos\alpha_n, z). \tag{9.4}$$

Fluence sums were calculated for five misalignments between $s_u = -1.5\,\mathrm{mm}$ and $s_u = 1.5\,\mathrm{mm}$.

9.3 Results

9.3.1 Pencil beam WEPL correction

Figure 9.2 (a) shows the counts distribution summed over all pencil beams in coordinates relative to the pencil beam center. The expectation value (marked with a cross) is in the center as expected. The shape of the pencil beam is not isotropic: the FWHM, indicated by a dashed line, is elliptical and wider in the $\tilde{u}$–direction. The FWHM was $\mathrm{FWHM}_{u,\mathrm{exp}} = 8.6\,\mathrm{mm}$ in the $\tilde{u}$–direction and $\mathrm{FWHM}_{v,\mathrm{exp}} = 6.9\,\mathrm{mm}$ in the $\tilde{v}$–direction, which is slightly smaller than $\mathrm{FWHM}_u = 9.5\,\mathrm{mm}$ in the $\tilde{u}$–direction in the Monte Carlo planning study and considerably smaller than $\mathrm{FWHM}_v = 12.3\,\mathrm{mm}$ in the $\tilde{v}$–direction.

Figure 9.2 (b, c) displays the average WEPL and the fitted correction function Δ_{WEPL}. The correction function described the WEPL distribution well, especially within the indicated FWHM ellipse. There was a clear dependency of the WEPL distribution in the $\tilde{u}$–direction. The dependency in the $\tilde{v}$–direction was weaker, but still captured by the fit of the correction function. The parameters of the correction function were determined as $a = 0.956\,\mathrm{mm}$, $b_u = 9.63 \times 10^{-2}$, $b_v = -0.884 \times 10^{-2}$, $c_u = -2.08 \times 10^{-3}\,\mathrm{mm}^{-1}$, and $c_v = -1.52 \times 10^{-3}\,\mathrm{mm}^{-1}$.

In figure 9.2 (d) profiles through counts, average WEPL and the correction along $\tilde{u}$ and for $\tilde{v} = 0\,\mathrm{mm}$ are shown. In the center of the pencil beam a negative WEPL of around $-1.0\,\mathrm{mm}$ was observed. Again, there was a good agreement between the fit and the averaged WEPL, in particular for the center of the beam spot. The expectation value and the $\mathrm{FWHM}_{u,\mathrm{exp}}$ interval for the counts distribution are indicated by solid lines under the distribution, which was slightly left–skewed: the mode was shifted to the right but the expectation value was at $\tilde{u} = 0\,\mathrm{mm}$ as desired.

In figure 9.2 (e, f) two reconstructions of the water phantom for the constant noise prescription are displayed with and without the pencil beam correction. Rings in the image were efficiently removed by the correction. Minor rings and a dip in the center of the phantom remained. The average value over the whole water region (RSP = 1.0)

changed slightly from 0.990 ± 0.024 (without correction) to 0.993 ± 0.023 (with correction). All further reconstructions employed the pencil beam correction.

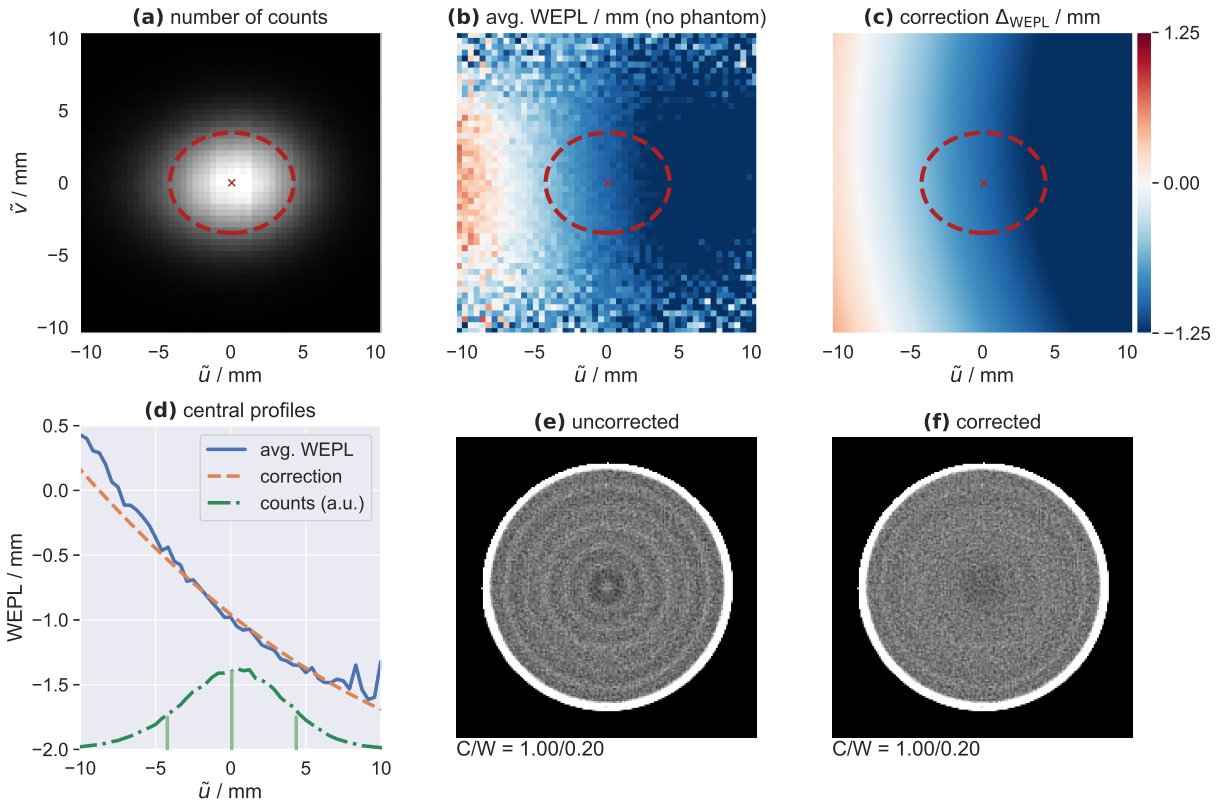

Figure 9.2: Illustration of the effect of the pencil beam WEPL correction: (a) number of counts over all registered pencil beams, (b) the average uncorrected WEPL, (c) the optimized correction function, (d) profile plots for (a) to (c) along $\tilde{u}$ for $\tilde{v} = 0\,\text{mm}$ (counts are in arbitrary units on the secondary axis), (e) an uncorrected RSP reconstruction and (f) the corrected reconstruction. (a) to (d) are given in coordinates $(\tilde{u}, \tilde{v})$ relative to the pencil beam center. In (a) to (d) the expectation value and the FWHM ellipse is indicated (for (d) as a projection to the $\tilde{u}$–axis). (b) and (c) share the same color scale.

9.3.2 Image noise evaluation

Before reconstruction, we randomly accepted $r_{\text{accept}} = 25.62\,\%$ of the measured protons after determining the actual delivered number of protons in a single run without phantom. This value is reasonable considering that we planned to acquire four times the required data (see section 9.2.6).

Figure 9.3 displays experimental results for the water phantom in (a) to (c) together with Monte Carlo results from the planning study in (d). The first row displays RSP maps for all acquisitions. The ROI for the FMpCT noise prescription is indicated by a dashed line in the RSP maps. For the constant noise prescription, this ROI was the entire phantom volume, except for the edges. In (c) and (d) a clear modulation of the noise pattern is already visible in the RSP image.

The variance map in the second row of figure 9.3 for unit fluence (a) showed the expected reduction in the center and an increase with increasing radius, peaking at the edges. In comparison to that the constant noise prescription in (b) was visibly more uniform, but rings distorted the flatness. Those rings could later be linked to the smaller spot size described before, together with a slight misalignment of the fluence pattern. The variance map in (c) confirmed what is already visible in the corresponding RSP map: noise for the FMpCT acquisition was low inside the ROI and high elsewhere. The planned noise in (d) was similar to the achieved noise in (c) inside the ROI, but slightly higher outside. Faint variance rings were also visible in the low variance area.

The third row displays line profiles through the variance maps as indicated by the dashed lines in the second row. Moreover, profiles through the prescription for (b) to (d) are displayed and the planned noise is added in (b) for comparison. Inside the ROI, the achieved experimental variance in (b) and (c) was in agreement with the planned noise. The prescription was achieved, considering that the optimizer tried to achieve peak variance or lower inside the ROI. In the center of (b) and outside of the ROI in (c) the experimental variance differed slightly from the prescription. However, also the planned noise in (d) differed from the prescription in this region.

The last row of figure 9.3 displays counts sinograms. Data is shown only for the same slice that is displayed in the first and second rows. Sinograms were not changing with the rotation angle for the unit fluence acquisition in (a). Nevertheless, the single pencil beam contributions were visible. The constant noise prescription in (b) featured reduced fluence in the center of the phantom which increased towards the edges. The counts sinograms for the FMpCT noise prescription in (c) and (d) agreed between simulation and experiment.

Figure 9.4 shows the same evaluations for the head phantom, albeit only for unit fluence and the FMpCT noise prescription. For unit fluence in (a) the variance map in the second row featured increased noise levels around strong heterogeneities (e.g. the nasal cavities) and the edges, as expected from previous studies. Additionally, variance rings were also noticeable, which were not expected from the simulation study (shown only as profile) and were caused by the smaller pencil beam size in combination with a misalignment as described before. The strongest variance peaks were also visible for the FMpCT noise prescription in (b), but in particular inside the ROI the variance distribution was flat as

intended. Variance profiles in the third row indicate agreement of the experimental data with the planned noise (c). Both the achieved noise and the planned noise were below the prescription as the optimizer tried to achieve peak variance inside the ROI.

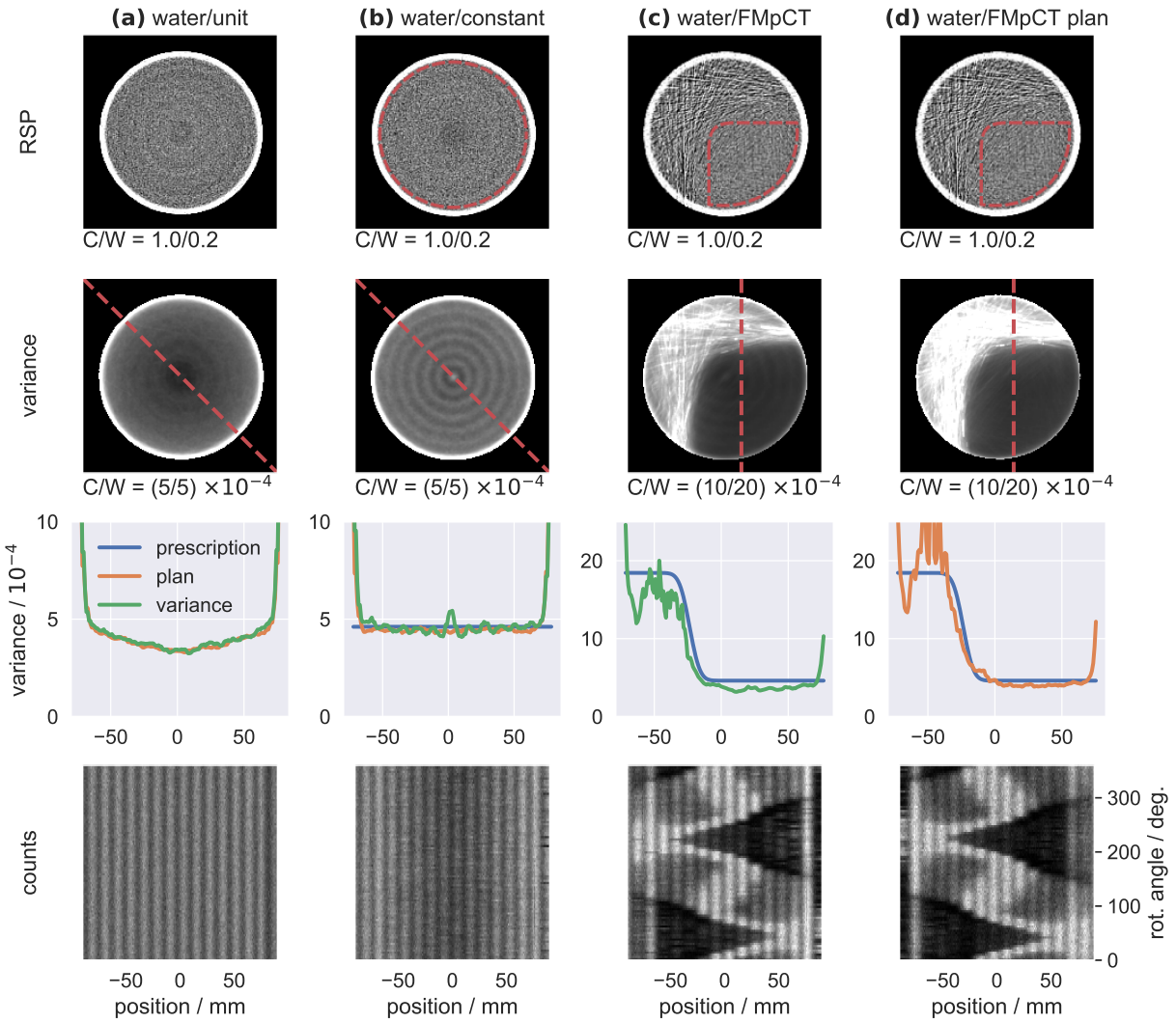

Figure 9.3: Proton CT acquisitions for the water phantom: (a) unit fluence scan, (b) constant noise prescriptions and (c) FMpCT scan together with (d) the planned noise distribution for the FMpCT scan. The four rows display: the reconstructed RSP maps, the corresponding variance maps, profiles through the variance maps, the prescription and the corresponding planned noise, and counts sinograms. The ROI for the noise prescriptions is indicated by a dashed line in the RSP maps. Center (C) and window (W) settings for displaying maps are given below each map. For the counts sinograms this was C/W = 140/280 protons.

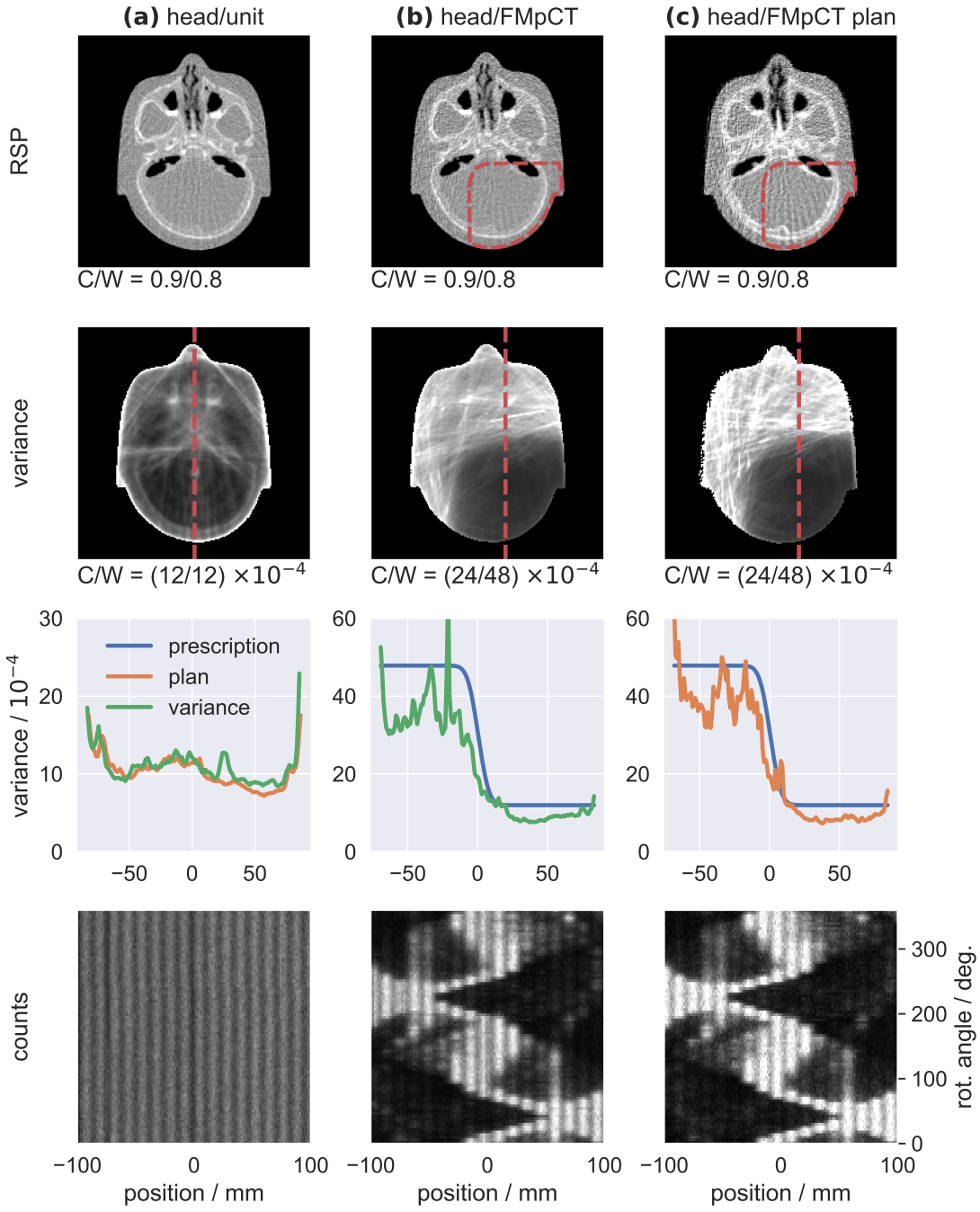

Figure 9.4: Proton CT acquisitions for the head phantom: (a) unit fluence scan and (b) FMpCT scan together with (c) the planned noise distribution for the FMpCT scan. The four rows display: the reconstructed RSP maps, the corresponding variance maps, profiles through the variance maps, the prescription and the corresponding planned noise, and counts sinograms. The ROI for the noise prescriptions is indicated by a dashed line in the RSP maps. Center (C) and window (W) settings for displaying maps are given below each map. For the counts sinograms this was C/W = 140/280 protons.

9.3.3 Fluence sums

In figure 9.5 (a, b) fluence sums are displayed (a) for all cases of the experimental study and (b) for the corresponding planning study. On the left of each column, one slice of the fluence sum, is shown and on the right of each column a perpendicular view is shown calculated along the dashed line indicated in the first panel. In general, the fluence sums of experiment and planning study showed similar modulations. Differences were caused by the smaller pencil beam size in the experimental data as well as by a slight misalignment of the grids. This was most noticeable in the perpendicular view where the single pencil beam contributions were visible in the experimental data, but were smeared out in the planning study which assumed a larger pencil beam spot size. This predominantly affected the direct vicinity of the rotation axis, which is in the center of the perpendicular view. The water/unit acquisition in (a) had constant fluence in the x'–direction and just a slight drop of the summed fluence close to the rotation axis. Modulations due to the pencil beam size were visible in the z–direction. The two other water phantom acquisitions and the head/unit data in (a) showed distinct pencil beam spots both in the x' and z–direction and in particular suffered from a distinct drop of fluence in the center, which hints that the fluence patterns were slightly misaligned in these scans and that the exact quarter pencil beam shift partially failed. In contrast to that, the head/FMpCT data in (a) also showed pencil beam patterns in both directions, but exhibited an increase of the summed fluence close to the rotation axis, which suggests that a misalignment happened in the opposite direction. All experimental acquisitions, except for water/unit consequently showed rings in the summed fluence in the xy–plane, and most prominently in the center of the volume.

The effects of the smaller pencil beam size ($\mathrm{FWHM}_{u,\mathrm{exp}} = 8.6\,\mathrm{mm}$, $\mathrm{FWHM}_{v,\mathrm{exp}} = 6.9\,\mathrm{mm}$) and a misalignment in the fluence pattern were further investigated in simulations in figure 9.5 (c). The simulated fluence sum with a shift of $s_u = 0\,\mathrm{mm}$ showed a very similar pattern as the water/unit data in (a). As the shift gets larger towards $s_u = -1.5\,\mathrm{mm}$, an interference pattern occurs in the x'–direction as well as a prominent drop of fluence in the center. With positive values of s_u the fluence increased in the center.

9.3.4 RSP evaluation

Figure 9.6 displays two scans of the sensitometry phantom: one with unit fluence (a) and one with the FMpCT noise prescription (b). As indicated in (c) the Delrin, acrylic, polystyrene and LDPE inserts were outside of the ROI and were imaged with high variance — the PMP and Teflon inserts were inside the ROI and were imaged with low variance.

In figure 9.6 (d) histograms are shown for the RSP distributions inside the inserts of the phantom as well as for the phantom body. For each insert, both the histogram of the unit fluence data and that of the fluence–modulated data are shown. Data were separated according to ROI–membership and the body was partially inside and partially outside the ROI. For the inserts outside of the ROI the distributions of the FMpCT scan were considerably broader compared to the unit fluence distributions. For the inserts inside the ROI, RSP distributions had a similar spread for the unit fluence scan and the FMpCT

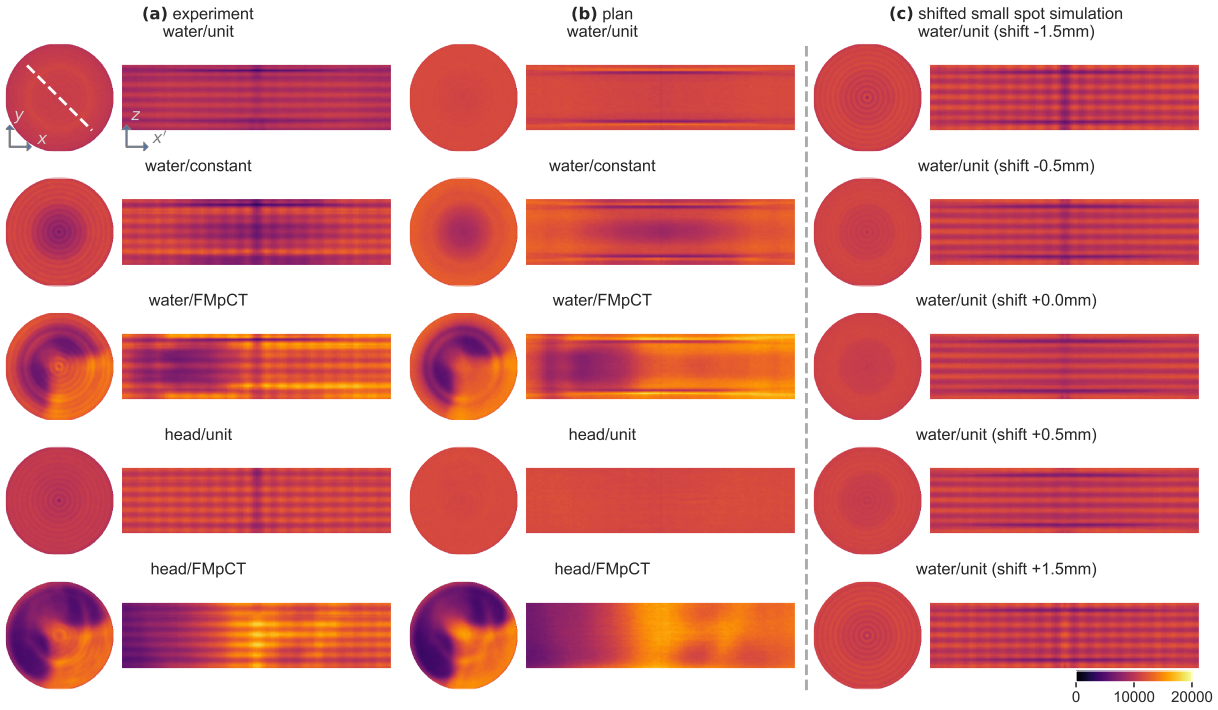

Figure 9.5: Fluence sums for (a) all experimental acquisitions of the water and head phantom and (b) for the corresponding Monte Carlo data. In (c) Monte Carlo data for the water phantom at unit fluence with the smaller spot size of the experimental data and different shifts s_u of the fluence pattern are shown. The three–dimensional fluence volume is displayed for one slice (showing the x and y coordinates) and as a perpendicular view along the indicated dashed line (showing the z coordinate together with a linear combination of the x and y coordinate, here indicated as x'). The color scale used for all displays is indicated with the last panel.

scan, which is expected as the variance prescription inside the ROI was equal to the peak variance in the unit fluence scan. The body exhibited only a slight broadening of its RSP distribution. There was no visible shift of the mean value of any of the distributions.

This was analyzed in detail in table 9.1 where for all inserts the water–column RSP value (Dedes et al. 2019) is given as well as relative errors for four datasets: data with and without pencil beam energy correction as well as for unit fluence and for the FMpCT noise prescription. Uncertainties are given as the error of the mean over all voxels inside the masked volume, thus yielding smaller uncertainties for the larger body mask. We calculated the mean absolute percentage error (MAPE) from the absolute value of the relative deviation of the measured and expected RSP for each of these datasets once averaging over all inserts (MAPE–ALL) and once averaging only over the two inserts inside the ROI of the FMpCT prescription (MAPE–ROI). Corresponding standard errors were calculated by error propagation.

In the unit fluence data, RSP errors improved for all inserts, except for the PMP

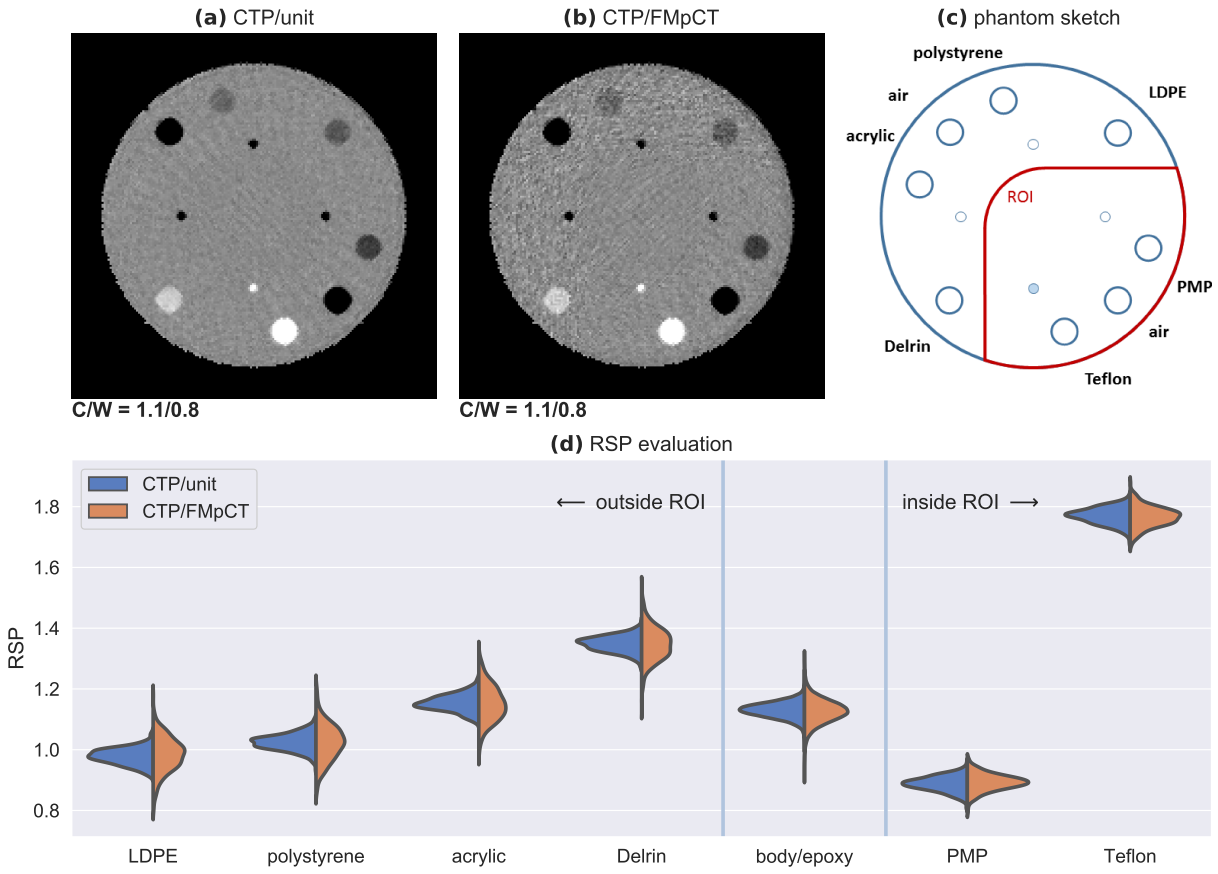

Figure 9.6: (a) Unit fluence scan and (b) FMpCT noise prescription for the sensitometry phantom together with (c) materials of the inserts and the ROI. (d) Histograms of the RSP value distributions inside the inserts and the body for both reconstructions. (PMP: polymethylpentene, LDPE: low–density polyethylene)

insert, when applying the pencil beam energy correction. The corresponding change of the MAPE–ALL, however, was less than its standard error. For the FMpCT acquisition, the MAPE–ROI of inserts inside the ROI was slightly increased by the correction, but the change was only a fraction of the corresponding standard error. Here in particular, the PMP (in the ROI) and LDPE (outside the ROI) inserts were degraded and all others improved. Changes in all inserts between corrected and uncorrected data were overlapping within the standard error of the mean, except for the body (compare distributions in figure 9.6 (d)). In the FMpCT scan using the correction, the MAPE–ROI was elevated compared to the corresponding unit fluence acquisition. Both inserts inside the ROI are slightly above 1% error, which is consistent with Dedes et al. (2019): the error there was 1.08% for PMP (corrected FMpCT: 1.06%) and –1.31% for Teflon (corrected FMpCT: –1.32%). For a further comparison, results from Dedes et al. (2019) are listed in the last column of table 9.1. In particular the MAPE–ALL and the MAPE–ROI agree between the two studies. Thus, the accuracy achieved in this study is comparable to Dedes et al.

Table 9.1: RSP values and errors for all inserts and the body of the sensitometry phantom for unit fluence and the FMpCT noise prescription. Scans with and without the pencil beam corrections were analyzed. For comparison, the last column shows results from Dedes et al. (2019) using the same phantom. The mean absolute percentage error is calculated from the absolute relative deviation of the measured and expected RSP values for the two inserts inside the ROI (MAPE–ROI) as well as for all inserts (MAPE–ALL). Uncertainties are given as the standard error of the mean over all voxels inside the masked volume and calculated by error propagation for MAPE–ALL and MAPE–ROI. (PMP: polymethylpentene, LDPE: low–density polyethylene)

		uncorrected error in %		corrected error in %		error in %
Insert	RSP	unit	FMpCT	unit	FMpCT	Dedes (2019)
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
polystyrene	1.024	-0.12 ± 0.30	-0.25 ± 0.66	0.06 ± 0.29	-0.11 ± 0.67	-0.04 ± 0.10
body/epoxy	1.144	-1.39 ± 0.02	-1.66 ± 0.03	-1.20 ± 0.02	-1.54 ± 0.03	—
acrylic	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

(2019) and consequently competitive to state–of–the–art clinical dual–energy x–ray CT.

9.4 Discussion

9.4.1 Pencil beam WEPL correction

In retrospective analysis of the acquired data we identified a possible incident spatial distribution of proton energy as the cause of RSP ringing artifacts in the reconstructed images seen in figure 9.2 (e). Such an energy distribution is not expected for the clinical treatment mode but may arise from the operation in research mode and reduction of the output fluence by closing slits in the beam line. Analysis of the proton data after transforming to coordinates relative to each pencil beam center allowed to isolate this energy spread (which manifested in a non–zero WEPL distribution in a scan without phantom). Fitting of this WEPL distribution and a subsequent subtraction allowed to flatten the WEPL distribution and to efficiently remove RSP rings in all acquisitions. Remaining rings and distortions visible in the water phantom scan are caused by inaccuracies in the detector calibration and known to distort pCT imaging accuracy (Dedes et al. 2019). While previous FMpCT publications (Dedes et al. 2018) also employed pencil beam scanning for

static fluence modulation, their absolute noise levels were higher compared to this study which partially masked the RSP ringing artifacts. Further analysis based on de–noising confirmed that similar distortions were also present in their data.

Analysis of the pencil beam spot revealed an elliptical shape, which was expected from previous investigations (Dickmann et al. 2020). However, the spot size was smaller than expected in both directions. In particular in the $\tilde{v}$–direction the spot size was almost halved. Operation of the scanner at very low proton fluence requires closing the momentum and divergence slits of the accelerator to values beyond what is commissioned for clinical use. Even though slit positions and magnet settings were equal as when estimating beam parameters in Dickmann et al. (2020), it is likely that maintenance of the beam line that was performed in between the acquisitions caused this change, although the ultimate cause could not be identified. For clinical operation such a change was not observed. While we investigate the impact of the distorted fluence delivery using Monte Carlo simulations in the next section, it was not possible to repeat the the experimental study with an amended optimization as this did not seem justified given the overall successful fluence modulation observed in this data. In general, if taken into account correctly, a smaller pencil beam size may be beneficial for the fluence optimization as more subtle modulations and smaller ROIs can be achieved. However, we re–optimized the fluence for the smaller pencil beam size and the impact on the large ROI chosen in this study was small.

Correcting for a shift in the incident energy distribution may have a minor impact on noise in the image. However, it is important to note that noise was originally distorted by the shifted WEPL values and the pencil beam correction brings it back the the original, expected value. Correction is therefore mandatory and we expect a better agreement with simulations (which do not contain the pencil beam effect) after correction.

9.4.2 Fluence sums

Analysis of the fluence sums and comparison of the experimental data to Monte Carlo data from the planning study revealed the effect of the smaller pencil beam spot size, in particular in the z–direction, which corresponds to the $\tilde{v}$–direction of the spot size. While the simulation study, which was performed with a larger spot size, employed smooth modulation patterns, the experimental scans exhibited the pencil beam spot structure. All scans suffered from the smaller spot size in the z direction (across the slices). Hence, in variance maps, we also expect an equivalent modulation of noise values across the slices.

One scan had a mostly smooth distribution within the slice (xy–plane). The other scans suffered from an additional misalignment of the quarter–shifted pencil beam pattern, which was characterized by a considerably reduced or, in one case, increased fluence close to the rotation axis, as well as a further unintended circular fluence pattern within the slice. This effect occurred in addition to the smaller spot size. The water/unit acquisition appears to have a better alignment than subsequent acquisitions. This is possible since it was performed on a different day. The alignment also may have changed within the magnitude of a millimeter while exchanging phantoms. We could show, using simulations, that misalignments of the order of $\pm 0.5\,\mathrm{mm}$ may introduce variance rings. Nevertheless,

the variance modulations of the FMpCT prescription are much larger than the distortions caused by the pencil beam spot size and the misalignment of the pattern. Therefore, the general FMpCT prescription was still achieved.

Dose savings were not investigated in this experimental study. However, the agreement of the fluence justifies the assumptions that experimental dose savings will be comparable to those evaluated in the planning study, which were up to 40% outside of the ROI for the FMpCT noise prescription. We therefore believe that this method will help to considerably reduce imaging dose for future pCT acquisitions.

9.4.3 Image noise and RSP evaluation

After correcting for the pencil beam WEPL distortion, the experimental scans using two phantoms and three fluence patterns (unit fluence, constant and FMpCT noise prescription) showed a very good agreement to what was expected from the planned noise distribution. Distortions from the prescription were limited to regions outside of the ROI where also the plan did not yield the prescription. This was because the optimizer tried to achieve a peak variance equal to the prescription and remaining spikes in the planned noise distribution forced the mean variance to be lower than the prescription. Consequently, also the out–of–ROI variance was below the prescription, as the variance contrast between high–noise and low–noise region is limited as discussed in Dickmann et al. (2020). In particular, for the constant noise prescription, which showed a very flat planned variance distribution, rings slightly distorted the variance map in the experimental data. These rings were not related to the pencil beam WEPL correction, but could be explained by the smaller pencil beam spot size and a slight misalignment of the patterns as described in the previous section. Nevertheless, agreement with the simulation study and with the prescription was satisfactory and the experimental realization of FMpCT was successful.

Evaluation of RSP values in a sensitometry phantom confirmed that the RSP accuracy was only slightly affected by the FMpCT acquisition and that the spread of the distribution of RSP values increased outside of the ROI but remained constant inside the ROI. While we observed slight degradations of RSP accuracy in some inserts, changes were mostly within the standard error of the mean (see table 9.1). RSP errors for unit fluence PBS scans agreed with values reported by Dedes et al. (2019) with the same phantom and a broad beam. We therefore conclude that the PBS acquisition is feasible and does not deteriorate the pCT accuracy if the pencil beam correction is applied. An investigation of the RSP accuracy in the FMpCT prescription for the two inserts inside the ROI with the most extreme RSP values led to a slight degradation due to FMpCT which was still within the standard error of the mean. It is important to note that the FMpCT ROI on purpose covered the two inserts which also showed the largest errors in Dedes et al. (2019), where the authors suggest that RSP errors are related to the design of this prototype scanner and could be avoided in a future scanner. Additionally, the FMpCT RSP errors for the inserts inside the ROI were very similar to those reported in Dedes et al. (2019). In the absence of artifacts using an idealized Monte Carlo simulation, Dedes et al. (2017) found no degradation of RSP accuracy with even stronger fluence modulations than those

used in this study. In the data presented in this study, all observed measurement errors were of the order of magnitude expected for a typical pCT scan. We, therefore, conclude that FMpCT can reduce dose (and increase noise) outside of a ROI while maintaining the required accuracy inside the ROI.

9.4.4 Lessons learned for future FMpCT acquisitions

Data in this study was acquired in a step–and–shoot mode, allowing enough time in between two projections for the scanner to rotate and for preparation of the next pattern. Consequently, the acquisition of one tomography took about 90 min, which would not be feasible for a patient scan. However, the beam–on time was 35.8 min for unit fluence tomographies and 20.3 min for FMpCT tomographies. This difference was caused due to the need of manually initiating the beam delivery for each projection, for which we generously prolonged the time the rotation stage rests at one angle to about three times the time needed to deliver the fluence. Considering that we also acquired four times the necessary data as discussed in section 9.2.6, the beam–on time required to record one of the datasets shown in this study would be only 9.2 min per tomography for unit fluence and 5.2 min for FMpCT scans, which is comparable to other acquisitions with this prototype scanner using a broad beam. If a synchronization between the scanner rotation and the beam delivery were possible (e.g. using a pulse generator) the phantom could potentially be rotated near–continuously and fluence be delivered with minimal beam–off time. Additionally, as discussed in Dedes et al. (2017), it may be possible to operate the scanner at a higher rate than 400 kHz without loss of efficiency.

Moreover, we observed inaccuracies in the fluence delivery, which may be avoided with automated software procedures to determine the size, the location and the proton count of the beam spot. This would allow to achieve the planned fluence delivery with a higher accuracy compared to this first experimental validation. It would also allow to directly deliver the correct proton fluence (and dose) without the need of a retrospective random rejection of protons with a constant rejection probability, as done in this study. A better automated alignment procedure may allow to avoid ringing in the summed fluence delivery. However, in conclusion of this study, it is advised that future FMpCT optimizations should employ a pencil beam grid with a considerably reduced spacing of pencil beams and requiring no quarter shifting of the pattern. This would prevent a constructive interference of the grids observed in this study and increase the robustness to a slight misalignment or change in spot size. Since the total number of protons per projection would be kept constant with a more dense pencil beam pattern, also the beam–on time and thus the total acquisition time would remain unchanged.

9.5 Conclusion

This study is the first experimental realization of using an optimization algorithm for dynamically modulated pCT to achieve task–specific image noise distributions. To this

end, we interfaced the control system of a PBS beam line to employ dynamically modulated fluence patterns for an experimental validation of FMpCT. We identified an incident beam energy dependence due to the pencil beam fluence delivery that caused ringing in the RSP maps. A proposed correction method captured the dependence and successfully removed rings. The resulting variance distributions showed a good agreement with the plans and the variance prescription and minor distortions of the fluence delivery were identified and can be avoided in the future.

Acknowledgments

This work was supported by the German Research Foundation (DFG) project #388731804 "Fluence modulated proton computed tomography: a new approach for low–dose image guidance in particle therapy" and the DFG's Cluster of Excellence Munich–Centre for Advanced Photonics (MAP), by the Bavaria–California Technology Center (BaCaTeC) project Nr. 28 [2015-2] and by the Franco–Bavarian university cooperation center (BayFrance). A technology transfer grant from the European Society for Radiotherapy (ESTRO) supported this project.

Chapter 10

Biological uncertainties in proton treatment planning

The investigation presented in this chapter was published in Physica Medica, volume 36, pages 91–102, in April 2017, with the title *Quantification of the uncertainties of a biological model and their impact on variable RBE proton treatment plan optimization*, by Resch et al. (2017) – (senior author: Dedes).

10.1 Introduction

By the end of 2015 more than 130.000 cancer patients have been treated with high energy proton beams (Jermann 2016). The main rationale for this technology is the proton depth dose profile, which is characterized by the Bragg peak with high dose deposition at the end of the proton range followed by a steep dose fall-off. This allows reducing the dose deposited in organs at risk (OARs). Furthermore, the integral dose to the patient can approximately be reduced by a factor of 2 by using protons instead of photon intensity modulated radiation therapy (Lomax et al. 1999). The relative biological effectiveness (RBE) of protons is considered to be equal to 1.1 in clinical practice (ICRU78 2007). However, in vitro experiments show that the RBE is increasing towards the end of the beam range (Wouters et al. 1996), where the linear energy transfer (LET) rises sharply (Sørensen et al. 2011, Jäkel 2008, Suit et al. 2010). As the RBE weighted dose (RWD) increases with dose averaged LET ($\mathrm{LET_d}$), the effective beam range increases by a few millimeters (Grün et al. 2013, Carabe et al. 2012). However, the RBE not only depends on $\mathrm{LET_d}$, but also on dose and tissue type (Tommasino & Durante 2015, Dasu & Toma-Dasu 2013).

For a certain cell type, the linear quadratic (LQ) model with the linear parameter α and the quadratic parameter β is commonly used to model in vitro cell survival experiments. The ratio of the two parameters, (α/β), is clinically used to distinguish late and early reacting tissues (corresponding to a low and high (α/β), respectively) (Tommasino & Durante 2015). RBE models based on the linear quadratic formulation of cell survival pro-

posed by Wilkens & Oelfke (2004), Carabe-Fernandez et al. (2007), Wedenberg et al. (2013) or McNamara et al. (2015) describe the RBE as a function of dose and $\mathrm{LET_d}$. The latter two models additionally include the $(\alpha/\beta)_x$ of photons to account for RBE dependence on tissue type. However, $(\alpha/\beta)_x$ values exhibit large error bars, which cause considerable RBE variations when applied in biological models (Carabe et al. 2013). In addition to the uncertainty of the $(\alpha/\beta)_x$, RBE modeling suffers from considerable uncertainties in experimental in vitro cell survival experiments (Paganetti 2014). Therefore, some authors suggest that the uncertainties in biological data are too high to apply RBE models in clinical practice (Paganetti et al. 2002), and suggest further investigations (Paganetti 2014). The clinical impact of the $(\alpha/\beta)_x$ uncertainty has been quantified in (Carabe et al. 2013) in the framework of the Carabe-Fernandez model (Carabe-Fernandez et al. 2007), but so far no study has reported on the uncertainty introduced by the fit to the experimental data. The purpose of this study was to investigate the uncertainty of the RBE weighted dose by considering both the uncertainty of the $(\alpha/\beta)_x$ and the intrinsic uncertainty of the biological model originating from the fit to experimental data. The Wedenberg et al. (2013) model was chosen as RBE model, since it was fitted to different cell types in contrast to the Carabe-Fernandez et al. (2007) or Wilkens & Oelfke (2004) model, which were fitted solely on V79 cell data. Furthermore, it was statistically tested to represent the fitted data well (Wedenberg et al. 2013). To ensure optimal dose and $\mathrm{LET_d}$ calculation, the treatment plans were created by a Monte Carlo simulation (Agostinelli et al. 2003) based treatment planning system (TPS) using clinically applied fractionation schemes (Dasu & Toma-Dasu 2013, Friedrich et al. 2014).

Additionally, we adapted the biological effect optimization method developed by Wilkens and Oelfke in order to yield a homogeneous RWD distribution when applying the Wedenberg et al. (2013) model (Wilkens & Oelfke 2005, Schell & Wilkens 2010, Friedrich et al. 2014). Tissue specific $(\alpha/\beta)_x$ values were used and assigned in a way that aims to reduce the consequences of potential $(\alpha/\beta)_x$ uncertainties.

10.2 Materials and methods

10.2.1 Biological model

The biological model used in this work was proposed by Wedenberg et al. (2013). Similar to other phenomenological models (Wilkens & Oelfke 2004, Carabe-Fernandez et al. 2007, Frese et al. 2011, McNamara et al. 2015), the Wedenberg model is based on the linear quadratic model (LQ) in which the cell survival fraction (S) is described as a function of the absorbed dose (D)

$$S = \exp(-\alpha D - \beta D^2). \tag{10.1}$$

The quadratic parameter for β protons (β_p) is assumed to be independent of $\mathrm{LET_d}$ and the linear parameter (α_p) is considered to increase linearly with $\mathrm{LET_d}$ (Wedenberg

et al. 2013):

$$\frac{\alpha_\mathrm{p}}{\alpha_x} := 1 + \frac{q}{(\alpha/\beta)_x} \cdot L; \quad \beta_\mathrm{p} := \beta_x \tag{10.2}$$

where α_x, β_x and $(\alpha/\beta)_x$ are the parameters of the reference photon radiation, L the $\mathrm{LET_d}$ and q a free parameter, which can be estimated by a fit to experimental cell survival data. Wedenberg et al. (2013) fitted equation 10.2 to a set of 24 data points of 10 different cell lines from 6 cell survival studies, which resulted in q equal to 0.435 [CI 95%(0.366, 0.513)] $\mathrm{Gy} \cdot \mu\mathrm{m} \cdot (\mathrm{keV})^{-1}$. Using the assumptions in equation 10.2 together with the definition of $RBE = D_x/D_\mathrm{p}$, RBE can be expressed as a function of $(\alpha/\beta)_x$, $\mathrm{LET_d}$ (L) and proton dose (D):

$$\mathrm{RBE}\left(\left(\frac{\alpha}{\beta}\right)_x, L, D\right) = -\frac{1}{2D}\left(\frac{\alpha}{\beta}\right)_x + \frac{1}{D} \times \sqrt{\frac{1}{4}\left(\frac{\alpha}{\beta}\right)_x^2 + \left(qL + \left(\frac{\alpha}{\beta}\right)_x\right) D + D^2} \tag{10.3}$$

The functional form of equation 10.3 is plotted in figure 10.1. The RBE in the Wedenberg model increases monotonically with $\mathrm{LET_d}$ and decreases with increasing $(\alpha/\beta)_x$. There is only a small dose dependence for high $(\alpha/\beta)_x$ and a considerable increase in RBE with decreasing dose for low $(\alpha/\beta)_x$.

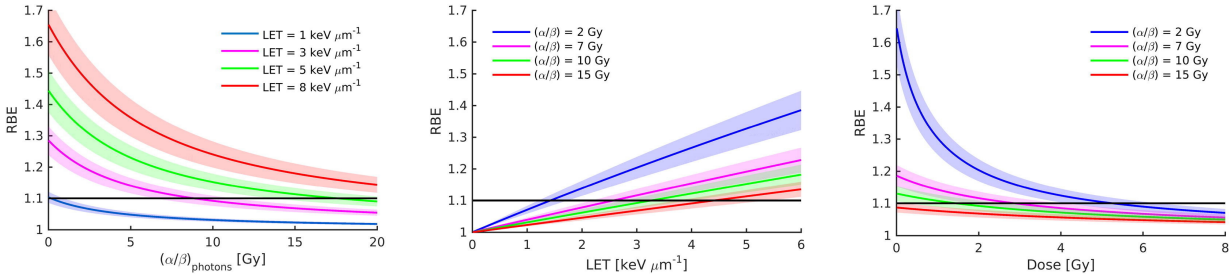

Figure 10.1: RBE dependence of the Wedenberg model on $(\alpha/\beta)_x$, $\mathrm{LET_d}$ and dose. In the left and the middle figure, dose is kept constant at 2 Gy and $\mathrm{LET_d}$ is kept constant at 3 keV/μm in the right figure. The uncertainty bands correspond to the 95% CI and originate in the fit. The horizontal black line represents the RBE = 1.1 approximation.

10.2.2 Dose and LET calculation

For this study, a Monte Carlo simulation–based treatment planning tool for clinical calculations in CT geometries was developed, following a similar concept as in (Mairani et al. 2013). A particle therapy extension of CERR ((Deasy et al. 2003, Schell & Wilkens 2010), Matlab/2014a) was used as TPS and GEANT4 (version 10.01) (Agostinelli et al. 2003) for dose and $\mathrm{LET_d}$ calculation.

Firstly, the lateral and axial pencil beam (PB) positions were determined in the TPS. The beam spots were placed on a 5 mm grid in lateral direction and with 3 mm spacing

in axial direction. In order to get the energies of the PBs, the Hounsfield units of the CT image were converted into stopping power ratios (RSPs). From the RSPs along the central ray of the PB, the water equivalent path length (WEPL) and hence the PB energies using an energy–range look–up table generated with GEANT4 were determined.

Secondly, the geometrical positions and corresponding energies of the PBs were exported to GEANT4 together with the patient CT, which was converted into density and elemental composition following Schneider's approach (Schneider et al. 2000) calibrated for the CT scanner (Toshiba Aquilion LB, Toshiba Medical Systems, the Netherlands) located at Klinikum Großhadern, and ionization potential according to ICRU report 49 (ICRU49 1993). Note that the conversion of CT numbers in the TPS and MC simulations need to yield the same RSP in order to prevent shifts between the range determination using the WEPL calibration in the TPS and intrinsic calculation in the GEANT4 simulations. From extensive testing, linear interpolation of the RSP sampling points used for GEANT4 input was found to yield the most consistent results and was hence used in the TPS. The dose and $\mathrm{LET_d}$ were converted on the fly following the dose–to–water concept (Paganetti 2009) and each PB was simulated in GEANT4 with 10^5 primary protons (events). This corresponds to a total number of approximately $1.5 \cdot 10^8$ and $5 \cdot 10^7$ simulated events (GEANT4), for the two studied patient cases, which are described in detail in section 10.2.4 (patient settings). Both the CT and the dose scoring voxel sizes were $1.074 \times 1.074 \times 3\,\mathrm{mm}^3$. $\mathrm{LET_d}$ scoring was successfully benchmarked against literature for different voxel sizes (Wilkens & Oelfke 2003, Cortés-Giraldo & Carabe 2015).

Finally, the three-dimensional dose and $\mathrm{LET_d}$ distributions of each PB were passed back to the TPS and used for TP optimization (see appendix C.0.1) and (Wilkens & Oelfke 2005, Schell & Wilkens 2010).

10.2.3 Biological effect optimization

Wilkens & Oelfke (2005) presented a biological effect optimization method based on the LQ model and a biological model which was fitted to Chinese hamster cell (V79) survival data (Wilkens & Oelfke 2004). This method was adapted in order to use the Wedenberg model instead of the original biological model and implemented into the TPS. Details about the implementation and the derivation can be found in the original paper (Wilkens & Oelfke 2005) and the appendix C.0.2.

10.2.4 Settings

Patient settings

According to equation 10.3, the RBE depends on $(\alpha/\beta)_x$. Therefore, two different tumor types were chosen for investigation, a nasopharyngeal cancer patient with a high $(\alpha/\beta)_x$ and a prostate cancer patient with a low $(\alpha/\beta)_x$. As the debate on the optimal fractionation scheme for prostate cancer is not settled, a standard and a hypofractionation TP with 35 fractions of 2 Gy(RBE) and 5 fractions of 7.6 Gy(RBE) respectively (Chan et al. 2016)

were investigated. The prescribed doses and treatment settings are listed in table 10.1. TP optimization objectives were: brainstem (center: $D_2 < 54\,\text{Gy(RBE)}$, surface: $D_2 < 63\,\text{Gy(RBE)}$)(Debus et al. 1997, Mayo et al. 2010), optic nerve ($D_2 < 56\,\text{Gy(RBE)}$)(Mayo et al. 2010a), rectum ($D_2 < 105\%$), bladder ($D_2 < 105\%$), where the index i of D_i is % volume. The necessary biological parameters α_x and β_x including the 95% CI are listed in table 10.2. In overlapping regions of OARs and the PTV, the $(\alpha/\beta)_x$ of the OARs were chosen for consistency with (Carabe et al. 2013).The structures were adopted from the intensity–modulated radiation therapy (IMRT) treatment scenario.

Table 10.1: Treatment settings used in this study. Prescribed dose (D_{pr}), number of fractions (n) and dose per fraction (D). Beam angles (ϕ) are reported on the IEC-scale. ⋆In this study we investigated the first phase of the treatment. In the second phase a 20 Gy boost was applied to a smaller volume.

Tumor	D_{pr} [Gy(RBE)]	n	D [Gy(RBE)]	ϕ [^{o}]
Nasopharyngeal	50*	25	2	0, 110, 260
Prostate stand.	70	35	2	90, 270
Prostate hypo.	38	5	7.6	90, 270

Investigating the uncertainties of the RWD

In order to investigate the uncertainties of the RWD using the Wedenberg model, the TPs for the nasopharyngeal and the prostate cancer patient were optimized using the RWD resulting from a constant RBE equal to 1.1. Dose weighted by a constant RBE will further be referred to as $\text{RWD}_{1.1}$. After the optimization procedure, the variable RWD according to the Wedenberg model using equation 10.3 was calculated, denoted as RWD_{w} in the following. The individual error bars were calculated separately using the lower and upper limits of the 95% CI interval of the fit parameter q and $(\alpha/\beta)_x$ ratio, respectively. The upper limit of the total uncertainty was estimated by using the upper limit of q and the lower limit of $(\alpha/\beta)_x$ and vice versa for the lower limit. The $(\alpha/\beta)_x$ were assigned for each tissue individually according to table 10.2. In this study, the $\text{RWD}_{1.1}$ is said to be significantly different from the RWD_{w} if it is not within the 95% CI originating in the uncertainty of either q, $(\alpha/\beta)_x$ or both depending on the origin of uncertainty we want to address.

Biological effect optimization

The biological effect optimization following Wilkens' approach was performed on the nasopharyngeal and prostate cancer patient. In order to take the uncertainty of the $(\alpha/\beta)_x$ ratio into account, a conservative strategy to assign $(\alpha/\beta)_x$ to each organ was chosen.

Table 10.2: The tissue parameters used in this study (Carabe et al. 2013, Schell & Wilkens 2010, Fowler et al. 2001, Joiner & van der Kogel 2009). Due to lack of literature values: *50% uncertainty assumed for the nasopharyngeal cancer and the eye, $^{**}\alpha_x$ and β_x of the rectum used in the bladder.

Tissue	α_x [Gy^{-1}]	β_x [Gy^{-2}]	$(\alpha/\beta)_x$ [Gy]	CI [Gy]
Prostate tumor	0.036	0.024	1.5	(1.2, 5.6)
Nasopharyngeal tumor	0.112	0.011	10	(5.0, 15.0)*
Rectum	0.040	0.010	4.0	(2.5, 5.0)
Bladder	0.040**	0.010**	4.0	(3.0, 7.0)
Brainstem	0.053	0.027	2.1	(1.5, 3.9)
Optical nerve	0.051	0.026	1.6	(0.5, 10.3)
Eye	0.040	0.020	2.0	(1.0, 3.0)*

All organs were separated into two types, targets and OARs. In a conservative strategy, the biological effect should be underestimated in a target and overestimated in an OAR by potential parameter errors. Therefore, the upper limit of the $(\alpha/\beta)_x$ CI as listed in table 10.2 was assigned to targets and the lower limit of $(\alpha/\beta)_x$ was assigned to OARs (inverse dependence of RBE to increasing $(\alpha/\beta)_x$ in equation 10.3). Due to the lack of literature values for α_x and β_x to the corresponding $(\alpha/\beta)_x$ ratio uncertainty, which were necessary for the biological effect optimization (equation C.7, appendix C.0.2), the α_x and β_x values had to be altered such that they matched the desired $(\alpha/\beta)_x$ ratio. This variation followed also a conservative approach. In the targets, the original β_x value was lowered in order to increase the $(\alpha/\beta)_x$ ratio, whereas in the OARs, β_x was increased in order to reduce the $(\alpha/\beta)_x$ ratio, while α_x was kept constant in both cases. Since α_x and β_x of the bladder were not available in literature, the values of the rectum were used. TPs optimized using the variable biological effect will be denoted as $\mathrm{TP_{BEO}}$, whereas TPs optimized using a constant RBE equal to 1.1 will be referred to as $\mathrm{TP_{1.1}}$.

10.3 Results

10.3.1 Uncertainties of the RBE weighted dose (RWD)

The $\mathrm{TP_{1.1}}$ in this section have been optimized using the constant RBE equal to 1.1, as described in section 10.2.4 (Investigating the uncertainties of the RWD) and according to the established clinical practice. The variable $\mathrm{RWD_w}$ using the Wedenberg model was calculated after the aforementioned optimization procedure.

Nasopharyngeal cancer patient

The RBE weighted dose volume histogram (DVH) for the nasopharyngeal cancer patient is shown in figure 10.2. The $RWD_{1.1}$ shows the desired steep dose fall–off in the CTV. Additionally, the RWD_w calculated by the Wedenberg formula is plotted together with the uncertainty band caused either by the uncertainty of the fit parameter q or the uncertainty of the $(\alpha/\beta)_x$. The major contribution to the uncertainty band of the RWD_w in the CTV, (left) optic nerve and the (left) eye was the uncertainty of $(\alpha/\beta)_x$ whereas the contributions of $(\alpha/\beta)_x$ and q in the brainstem were approximately equal. In the CTV, the fixed RBE value resulted in a consistently higher $RWD_{1.1}$ than the nominal RWD_w, calculated by the Wedenberg model including the uncertainty band originating in the uncertainty of q. However, the $RWD_{1.1}$ was inside the uncertainty band caused by the uncertainty of $(\alpha/\beta)_x$. In the case of the left optic nerve, the constant RBE was outside of the confidence interval of q and on the lower limit of the uncertainty band of $(\alpha/\beta)_x$. In the left eye and the brainstem, the fixed RBE was significantly lower than the variable RBE regardless of the origins of uncertainty.

In figure 10.2(a), the corresponding DVH including the total uncertainty band, i.e., the combined uncertainties of $(\alpha/\beta)_x$, is plotted. This figure demonstrates that the $RWD_{1.1}$ was inside the total uncertainty band of the RWD_w in the target, although the nominal RWD_w value was consistently lower. Also in the left optic nerve, the $RWD_{1.1}$ was within the total uncertainty band of the model, but further off the nominal value and in contrast to the target region, the RWD_w was higher than the $RWD_{1.1}$ due to a lower $(\alpha/\beta)_x$. In the left eye and the brainstem, the RWD was significantly underestimated by the constant RBE approximation. Several characteristic values of the DVH plotted in figure 10.2 are listed numerically in table 10.3.

Note that the RBE was lower in the CTV than in the left eye although the LET_d was mostly higher. This originates in the lower dose in the eye and the inverse dependency of RBE on dose (see figure 10.2). In the cumulative volume histogram, the LET_d appears highest in the brainstem. As a consequence, the RWD_w in the overlapping region of the brainstem and the PTV was up to 20% higher than $RWD_{1.1}$ (see figure 10.3). The optimization objective of decreasing $RWD_{1.1}$ in this structure caused an increase in LET_d and consequently RWD_w, as LET_d is a free parameter in $TP_{1.1}$ optimization. This LET_d effect was additionally amplified by the lower $(\alpha/\beta)_x$ the brainstem compared to the CTV. As expected, elevated LET_d values were observed towards the end of the beam range and the lateral penumbra (figure 10.3(c)) and correspondingly increasing RWD values with discontinuities at tissue boundaries.

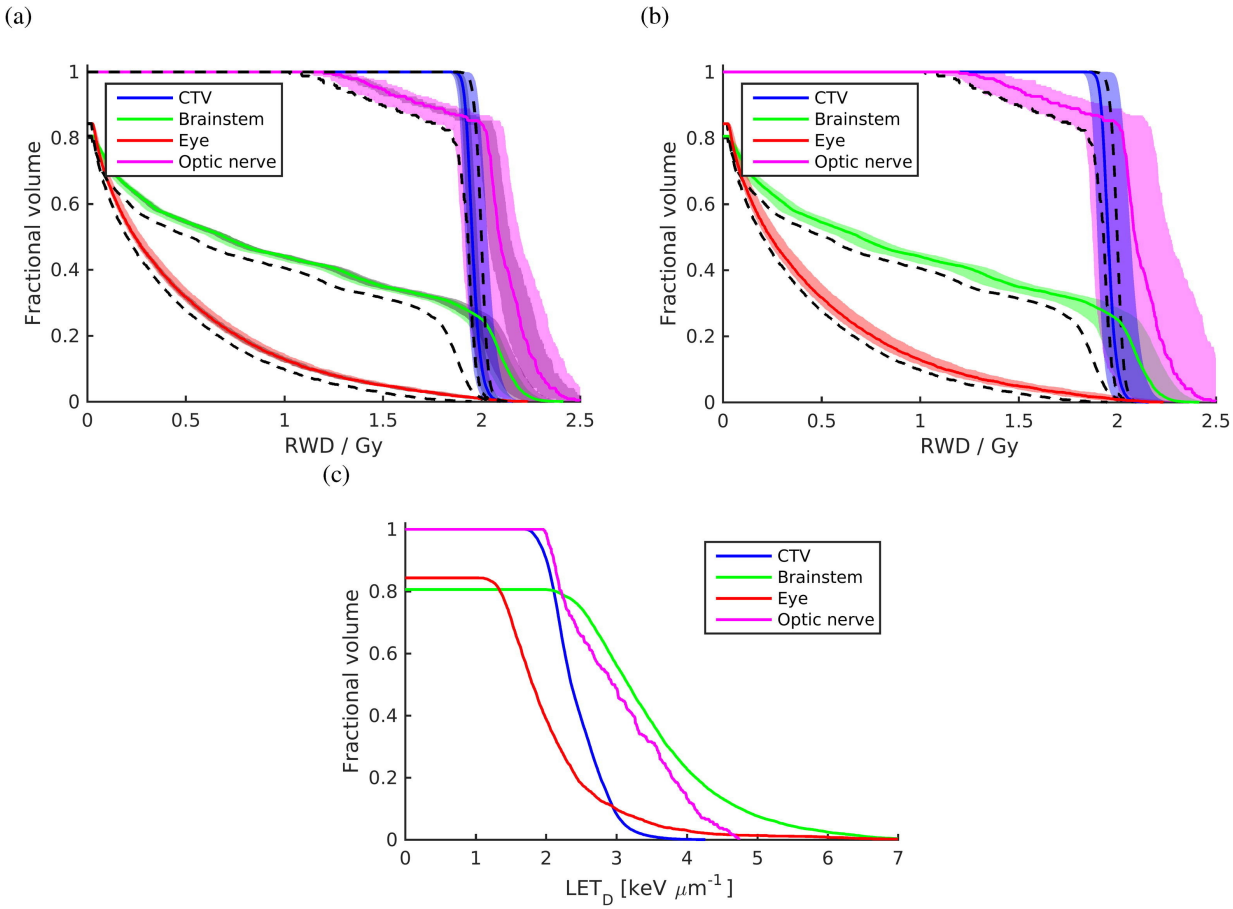

Figure 10.2: Graphs (a and b) show the cumulative DVH for the CTV, brainstem, left eye and left optic nerve of the $TP_{1.1}$ for the nasopharyngeal tumor patient. In (a), the uncertainties caused by q and $(\alpha/\beta)_x$ are plotted as dark and lightly shaded area, respectively. In (b), the total uncertainty band is plotted. The $RWD_{1.1}$ is shown as black dashed lines while the colored solid lines represent the nominal RWD_w values. The corresponding dose averaged LET volume histogram is plotted in (c). Cumulative volumes may not sum up to 1 (a and c), as only voxels receiving at least 1% of the dose prescribed to the target (i.e., 0.02 Gy(RBE)/fraction) were considered.

Table 10.3: DVH parameters (DP) for several regions of interests (ROIs) for the nasopharyngeal and prostate cancer patient corresponding to figures 10.2 and 10.4. The CI caused by either the fit parameter q ($\mathrm{CI_q}$) or the $(\alpha/\beta)_x$ ratio ($\mathrm{CI}_{(\alpha/\beta)_x}$) and the total $\mathrm{RWD_w}$ range ($\mathrm{CI_t}$), respectively, are listed. The index i of D_i indicates % volume.

ROI	DP	$\mathrm{RWD}_{1.1}$ [Gy(RBE)]	$\mathrm{RWD_w}$ [Gy(RBE)]	$\mathrm{CI_q}$ [Gy(RBE)]	$\mathrm{CI}_{(\alpha/\beta)_x}$ [Gy(RBE)]	$\mathrm{CI_t}$ [Gy(RBE)]
Nasopharyngeal cancer patient						
CTV	$\overline{\mathrm{D}}$	49.3	48.3	(47.5, 48.8)	(47.3, 50.3)	(46.7, 51.0)
	D_{98}	48.5	47.5	(46.8, 48.0)	(46.5, 49.3)	(46.0, 50.0)
	D_{50}	50.0	48.8	(48.3, 49.5)	(48.0, 50.8)	(47.5, 51.8)
	D_2	51.8	51.0	(50.3, 52.0)	(49.8, 53.5)	(49.2, 54.8)
Optic nerve	$\overline{\mathrm{D}}$	46.3	51.8	(50.0, 53.3)	(46.3, 54.5)	(45.5, 56.5)
	D_{50}	48.3	52.5	(50.8, 53.8)	(47.5, 55.0)	(46.7, 56.8)
	D_{10}	49.3	57.0	(55.0, 59.3)	(49.8, 60.8)	(48.7, 63.5)
	D_2	50.3	60.0	(55.0, 59.3)	(49.8, 60.8)	(48.7, 63.5)
Eye	D_{50}	5.0	6.0	(5.8, 6.3)	(5.5, 7.0)	(5.2, 7.5)
	D_2	39.0	45.8	(44.0, 47.5)	(44.0, 48.0)	(42.7, 50.0)
Brainstem	D_{50}	12.8	17.3	(16.3, 18.3)	(15.5, 18.3)	(14.7, 19.3)
	D_2	49.5	55.8	(53.8, 57.5)	(53.3, 56.8)	(51.5, 58.8)
Prostate cancer patient – standard fractionation TP						
CTV	$\overline{\mathrm{D}}$	70.4	79.1	(76.7, 81.6)	(72.8, 79.8)	(71.4, 82.6)
	D_{98}	69.0	77.0	(74.9, 79.5)	(71.4, 80.5)	(69.7, 81.6)
	D_{50}	70.4	78.8	(76.3, 81.2)	(72.8, 79.5)	(71.1, 82.3)
	D_2	72.1	83.3	(80.2, 86.1)	(76.0, 84.0)	(73.9, 87.2)
Rectum	D_2	70.7	75.3	(73.2, 77.0)	(73.9, 77.0)	(72.1, 79.5)
Bladder	D_2	49.7	54.6	(52.9, 56.0)	(51.8, 56.0)	(50.4, 57.8)
Prostate cancer patient – hypofractionation TP						
CTV	$\overline{\mathrm{D}}$	38.1	37.1	(36.7, 37.5)	(36.6, 37.1)	(36.3, 37.6)
	D_{98}	37.3	36.1	(35.8, 36.6)	(35.8, 36.2)	(35.4, 36.6)
	D_{50}	38.1	37.0	(36.6, 37.4)	(36.5, 37.0)	(36.7, 37.4)
	D_2	39.0	39.0	(38.3, 39.6)	(38.3, 39.1)	(37.7, 39.7)
Rectum	D_2	38.7	37.2	(36.9, 37.6)	(37.1, 37.4)	(36.8, 38.0)
Bladder	D_2	26.6	26.7	(26.0, 26.7)	(26.0, 26.7)	(25.7, 27.0)

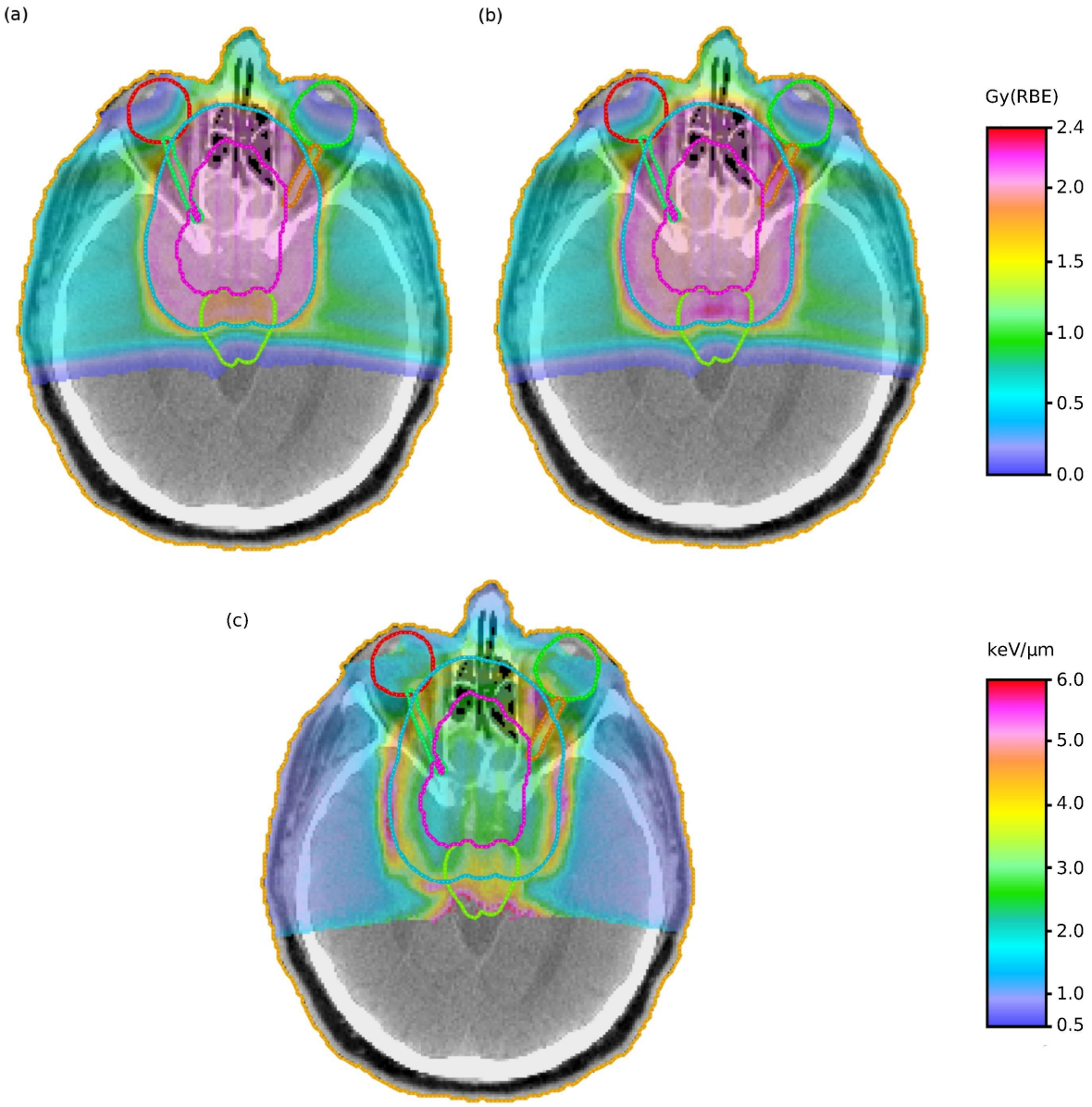

Figure 10.3: Representative RWD distribution of the TP for the nasopharyngeal cancer patient applying a constant RBE during $TP_{1.1}$ optimization. (a) $RWD_{1.1}$ [Gy(RBE)], (b) RWD_w [Gy(RBE)] and (c) LET_d [keV/μm]. The CTV (purple), PTV (blue), eyes (red and green), brainstem (green) and the optic nerves (green and orange) are delineated. No values are displayed in the nasal cavities with CT values equal to air.

Prostate cancer patient

Figure 10.4 shows the DVH of the standard fractionation (a and c) and hypofractionation TPs (b and d) for the prostate cancer patient. For both fractionation schemes and all ROIs, the magnitude of the uncertainty caused by either of the two sources of uncertainty was comparable. However, the RWD_w uncertainty band reflects the asymmetry of the CI of $(\alpha/\beta)_x$ around the nominal value for prostate cancer (see table 10.2). The uncertainty bands were distorted towards lower RWD_w values. The uncertainty bands caused by q were almost symmetrical (but not completely, due to the square root in the RBE function in equation 10.3). In figure 10.4 (c) and (d), the DVHs of the two TPs including the total uncertainty band are plotted. The $RWD_{1.1}$ was significantly underestimated in the standard fractionation TP. The hypofractionation TP exhibited a significant overestimation of $RWD_{1.1}$ with respect to the total uncertainty of the Wedenberg et al. model. The bladder did not receive a crucial dose in both TPs. The LET_d of the CTV in the standard fractionation TP ranged from 3 to 4 keV/μm. However, LET_d were mostly lower for the hypofractionation TP and ranged from 2 to 5 keV/μm. In the rectum, LET_d values were slightly higher than in the CTV ranging from 2.5 to 5 and 2 to 5 keV/μm for the standard fractionation and hypofractionation TP, respectively.

Figure 10.5 (e) and (f) show the LET_d distribution of a representative CT slice. High LET_d values could be found towards the end of the spread out Bragg peak and in the lateral penumbra, hence the rectum and bladder. The RWD_w distribution in (c) and (d) follows the LET_d distribution. The RBE increases towards the end of the beam range and caused an up to 25% higher RWD_w than the prescribed dose in the region between the CTV and the PTV, whereby the $RWD_{1.1}$ was distributed homogeneously at the prescribed dose level. Along tissue boundaries the RWD_w distribution showed discontinuities due to the discontinuity of $(\alpha/\beta)_x$.

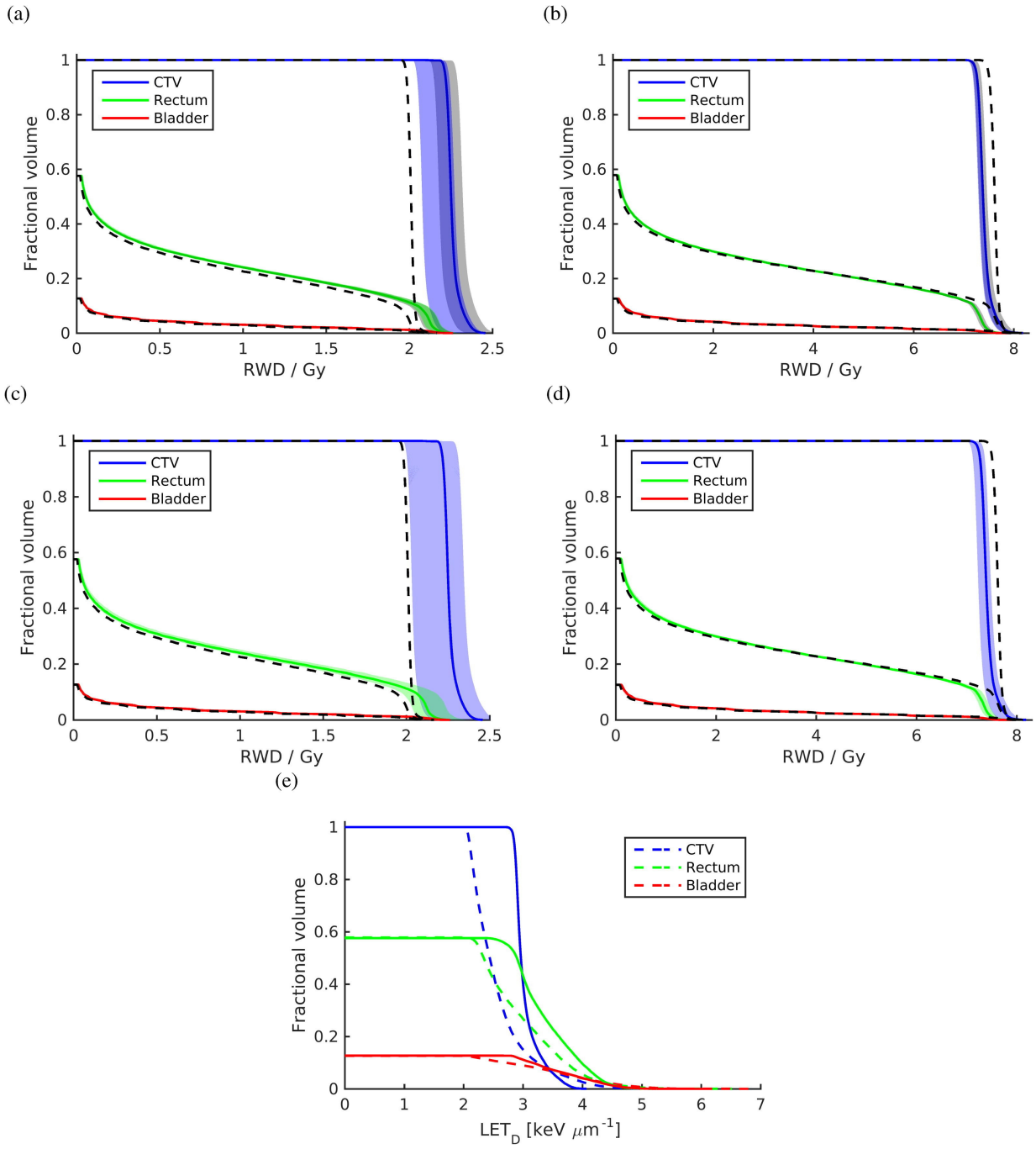

Figure 10.4: DVH of the standard fractionation (a and c) and hypofractionation (b and d) $TP_{1.1}$ for the prostate cancer patient. In (a) and (b), the uncertainty bands caused by q and $(\alpha/\beta)_x$ are plotted as dark shaded and colored shaded area, respectively. The total uncertainty band is shown in (c) and (d). The $RWD_{1.1}$ is shown by dashed black lines while the colored lines represent the nominal RWD_w values. In (e), the dose averaged LET cumulative volume histogram of the standard fractionation (solid lines) and hypofractionation (dashed lines) $TP_{1.1}$ is plotted. Volumes might not sum up to 1 as only doses higher than 1% of the prescribed dose were considered in all histograms.

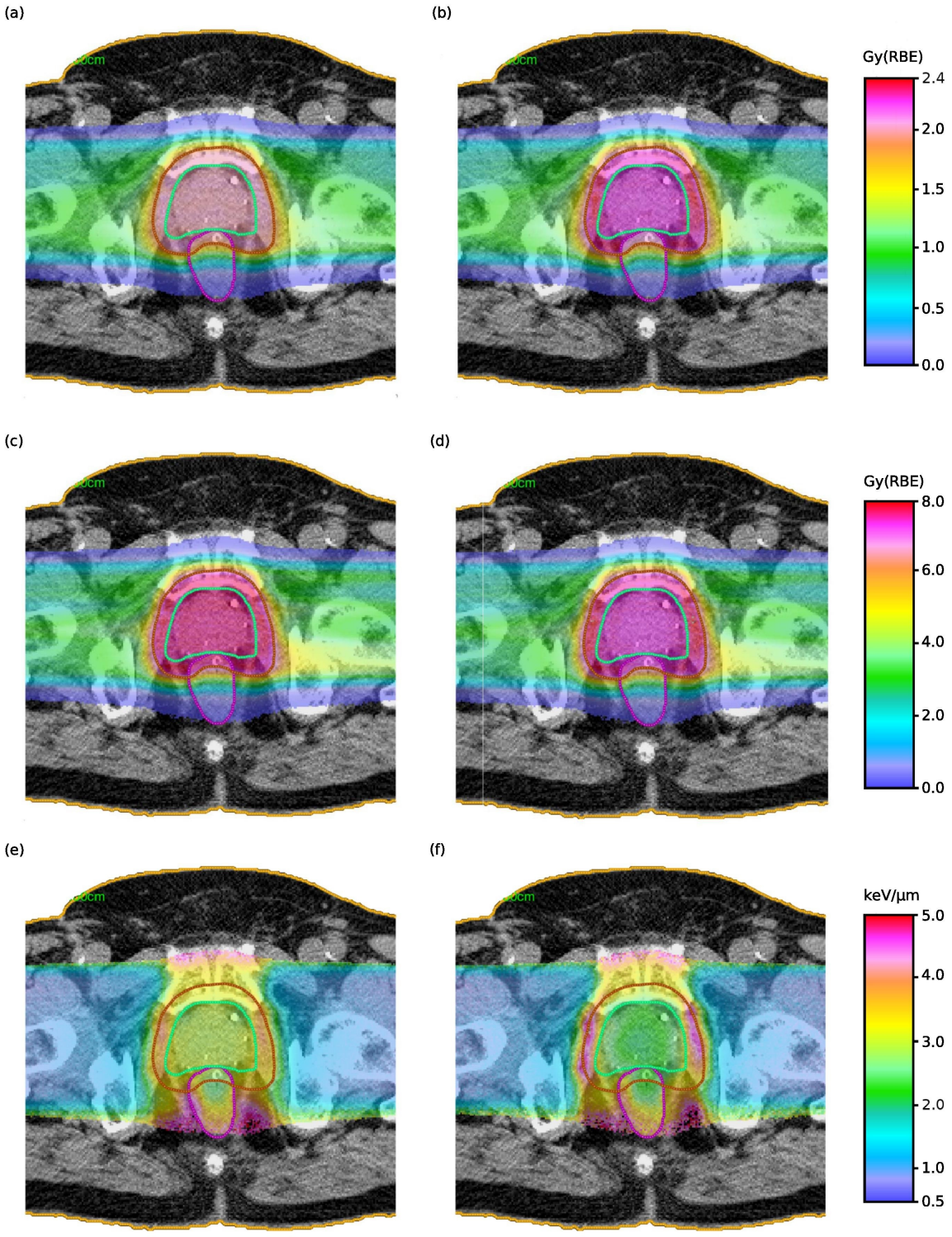

Figure 10.5: Representative RWD and LET_d distributions of the standard fractionation (a,b,e) and hypofractionation $TP_{1.1}$ (c,d,f) for the prostate cancer patient applying a constant RBE during $TP_{1.1}$ optimization. In (a and c) $RWD_{1.1}$ [Gy(RBE)], (b and d) RWD_w [Gy(RBE)] and (e and f) LET_d [keV/μm]. The CTV (green), PTV (red) and the rectum (pink) are delineated.

10.3.2 Biological effect optimization

Figures 10.6 and 10.7 show the results of the TPs for the two patient cases using the biological effect optimization (TP_{BEO}). The RBE weighted doses of the TP_{BEO} ($RWD_{w,BEO}$) and of the previous RBE = 1.1 optimized $TP_{1.1}$ (RWD_w) are shown.

Nasopharyngeal cancer patient

Figure 10.6 (a) shows the cumulative RWD volume histogram of TP_{BEO} and $TP_{1.1}$ for the nasopharyngeal cancer patient. The new optimization strategy (TP_{BEO}) yielded a $RWD_{w,BEO}$ fall-off in the CTV at the prescribed dose level equal to 2 Gy(RBE), as it was obtained for $RWD_{1.1}$ in the previous $TP_{1.1}$. In the optic nerve, the two objective variables in the optimization, namely the $RWD_{w,BEO}$ in the TP_{BEO} and $RWD_{1.1}$ in the $TP_{1.1}$, were also comparable. However, RWD_w was noticeably higher than $RWD_{w,BEO}$. Similarly to the optic nerve,the $RWD_{w,BEO}$ in the brainstem was noticeably lower than RWD_w. Note that the $RWD_{w,BEO}$ was calculated using the worst case $(\alpha/\beta)_x$, whereas the nominal $(\alpha/\beta)_x$ were used to calculate RWD_w. Consequently, comparing only the nominal RWD_w and $RWD_{w,BEO}$ would yield a bias in favor of RWD_w and therefore the uncertainty band needs to be taken into account. The worst case $(\alpha/\beta)_x$ assignment was reflected in the $RWD_{w,BEO}$ uncertainty band, which suggests a potential underestimation of the $RWD_{w,BEO}$ in the CTV, whereas $RWD_{w,BEO}$ was potentially overestimated in the OARs.On the other hand, the uncertainty band around the nominal RWD_w (see figure 10.2) was approximately symmetric.

Figure 10.6 (b) shows the corresponding $RWD_{w,BEO}$ distribution of are presentative CT slice. $RWD_{w,BEO}$ was homogeneously distributed inside the CTV at the prescribed dose level. In contrast to RWD_w, $RWD_{w,BEO}$ did not exhibit hot spots in the brainstem.

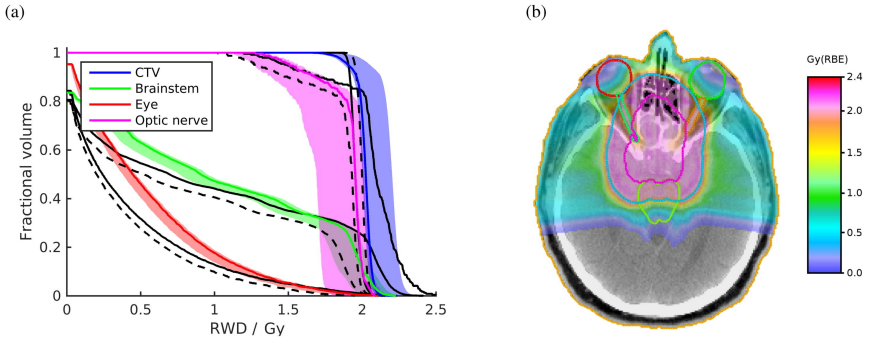

Figure 10.6: Biological effect optimized TP (TP_{BEO}) for the nasopharyngeal cancer patient. In (a) the DVH is shown, where the nominal $RWD_{w,BEO}$ is represented with a colored straight line and the total uncertainty band as shaded area. For comparison, the $TP_{1.1}$ optimized using the constant RBE equal to 1.1 is shown in black. The dashed black line represents $RWD_{1.1}$ and the solid black line represents nominal RWD_w. In (b) $RWD_{w,BEO}$ distribution for a representative CT slice.

Prostate cancer patient

Figures 10.7 shows the DVHs of the TP_{BEO} for the prostate cancer patient. For both fractionation regimens, the biological effect optimization yielded a $RWD_{w,BEO}$ fall–off at the prescribed dose levels(2 and 7.6 Gy(RBE)) in the CTV. The uncertainty band of the $RWD_{w,BEO}$ indicated potentially higher doses than prescribed in the CTV, and potentially lower doses in the OARs for both fractionation schemes (a and b). In contrast to the biological effect optimized TP_{BEO}, RWD_w was significantly over– and underestimated in the standard (a) and hypofractionation (b) RBE=1.1 optimized $TP_{1.1}$, respectively. Therefore, only the biological effect optimization yielded the prescribed RWD in the CTV when considering variable RBE.

In the rectum, the nominal RWD_w of the standard fractionation $TP_{1.1}$ was noticeably higher than $RWD_{w,BEO}$, even though the worst case $(\alpha/\beta)_x$ assignment disadvantages $RWD_{w,BEO}$. The uncertainty band of the $RWD_{w,BEO}$ indicated potentially lower RWD, whereas the uncertainty band of the RWD_w was approximately symmetric and did not favor lower RWD. In the hypofractionation regimen,the RWD_w was lower than the $RWD_{w,BEO}$ in the rectum. However, RWD_w was also lower than $RWD_{1.1}$ and the two optimization objectives of $TP_{1.1}(RWD_w)$ and $TPBEO(RWD_{w,BEO})$ were comparable.

Figures 10.7 (c and d) show the $RWD_{w,BEO}$ distribution. $RWD_{w,BEO}$ was homogeneously distributed inside the CTV in the standard fractionation and hypofractionation TPs.

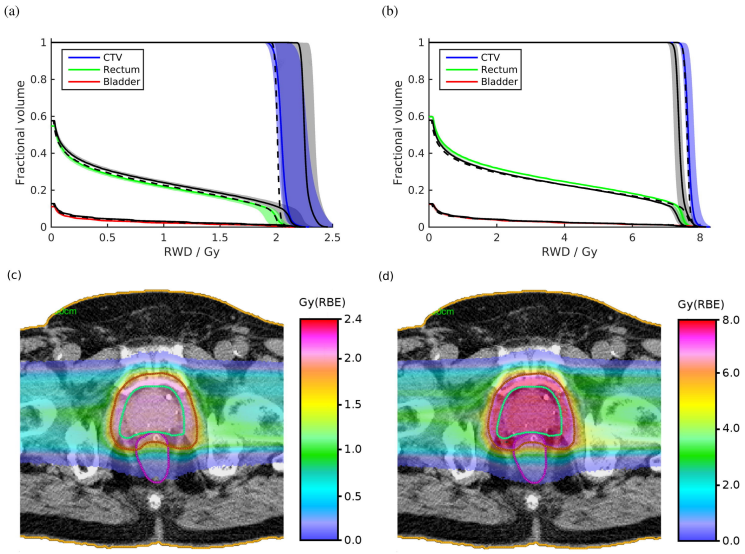

Figure 10.7: The biological effect optimized TP_{BEO} for the prostate cancer patient. The standard fractionation and hypofractionation TPs are plotted in the left (a and c) and right (b and d) column, respectively. In (a and b) the DVH is shown including the total uncertainty bands. The TP_{BEO} is drawn as colored line, whereas the RBE = 1.1 optimized TP is drawn in black for comparison. RWD_w is shown with a solid line, $RWD_{1.1}$ with a dashed line and the uncertainty band around RWD_w as dark shaded area. Graphs (c) and (d) show the $RWD_{w,BEO}$ distribution in a representative CT slice.

10.4 Discussion

The generic beam model used in this study assumed a constant initial beam size for all energies, although this is not achievable in all centers. However, LET_d dependence on spot size in a homogeneous field is not expected as the lateral LET_d dependence of a single PB is small (Wilkens 2004) and LET_d enters the RBE in equations 10.3 in the square root. Therefore, it is assumed that a constant spot size is a valid approximation for the purpose of this study.

The RBE in the Wedenberg formulation is a function of three variables, $(\alpha/\beta)_x$, LET_d and dose. The dependence on the first variable, $(\alpha/\beta)_x$, can be observed when comparing the nasopharyngeal cancer patient with the standard fractionation TP of the prostate cancer patient, where the dose levels (2 Gy(RBE)) matched. Although the LET_d in the former was slightly lower, the increase of RBE in the prostate cancer patient can be attributed to the lower $(\alpha/\beta)_x$.

The inverse dose dependence of RBE caused the RBE significantly greater than 1.1 in the standard fractionation TP1.1 and the RBE significantly lower than 1.1 in the hypofractionation TP1.1 for the prostate cancer patient. The LET_d in the standard fractionation plan was slightly higher than in the hypofractionation plan, as the TPs were optimized independently. LET_d is a free parameter in the RBE = 1.1 optimization and can therefore vary for similar dose distributions (Grassberger et al. 2011). However, LET_d enters the RBE formula in the square root and consequently small LET_d variations cause even smaller RBE variations.

The RBE higher and lower than 1.1 for the standard fractionation and hypofractionation regimen differs from the results in (Carabe et al. 2013), but agrees with the observations in (Dasu & Toma-Dasu 2013). The differences come from the different RBE dependence of the Carabe–Fernandez model on dose compared to the Wedenberg et al. model (see figure 3 in (Giovannini et al. 2016)). The study by Carabe et al. (2013) reported the RWD uncertainties caused by $(\alpha/\beta)_x$ as an average over 5 patients, whereas the present study presented only 1 patient for each cancer type. For the standard fractionation prostate scheme, the D_{90} RWD CI in (Carabe et al. 2013) was (+6, +55)% to the RWD1.1, whereas in this study the corresponding CI was (+4, +16)% (only considering $(\alpha/\beta)_x$). As only averaged values were reported in (Carabe et al. 2013), it cannot be determined whether the discrepancies originate in the different models or the averaging over different patients. Nevertheless, in accordance to (Carabe et al. 2013), relative RBE uncertainties in this study decreased with increasing dose per fraction. Here, the relative RWDw CI of D_{98} was +/–2% in the hypofractionation TP1.1 for the prostate cancer patient, whereas it was (–6, +10)% in the standard fractionation TP1.1.

All these results are based on the assumption of the applicability of the Wedenberg model, which is based on the LQ model to describe cell survival to irradiation, but complications in OARs are only partially correlated to cell death (Joiner & van der Kogel 2009). For OARs, approaches taking different biological endpoints into account, might be more suitable than the concept of $(\alpha/\beta)_x$. Although the LQ model is well established, a recent study proposed a modified LQ model, which results in a slightly reduced RBE (Kuperman 2016).

However, this RBE reduction becomes more pronounced with increasing LET and most $\mathrm{LET_d}$ values in this study were rather low (mostly 2–4 $\mathrm{keV}/\mu\mathrm{m}$, see figures 10.2 and 10.4). Furthermore, the RBE model is fitted to in–vitro cell survival experiments and cells might react differently to ionizing radiation in–vivo. There are several models which try to describe the RBE and there is no proof that one model yields better agreement to the available data than the others. The disagreement between these models by a few percent (McNamara et al. 2015, Giovannini et al. 2016) may be interpreted as additional systematic uncertainty, but is partially included in the uncertainty of the fit parameter. The Wedenberg et al. model was chosen for this study on the grounds that there is a statistical analysis which demonstrates that the model represents the fitted data well. However, it has been fitted to a considerably smaller data set than the McNamara et al. (2015) model. The Carabe–Fernandez and the Wilkens and Oelfke model have only been fitted to V79 cells whereas the model used in this study was fitted to different cell lines. Since all RBE models are fitted to experimental data, it is expected that all models exhibit uncertainties originating in the fit. The Carabe–Fernandez and McNamara et al. model have two free parameters unlike the Wedenberg et al. or Wilkens and Oelfke model, which have only one.

Overlapping ROIs with different $(\alpha/\beta)_x$ values are problematic and it is not clear which value should be assigned. In this study and in (Carabe et al. 2013), the $(\alpha/\beta)_x$ of the OAR was chosen for the intersections. Consequently, the RBE and RWD at the transitions are discontinuous and problematic in the biological effect optimization. This explains the reduced RWD in the PTV along the boundary to the OARs. The conservative way of assigning the $(\alpha/\beta)_x$ ratios as investigated in this study is a possibility to account for the $(\alpha/\beta)_x$ uncertainties during the optimization procedure. But the concept suffers from the lack of α_x and β_x values to the corresponding $(\alpha/\beta)_x$ ratio and although we used a reasonable approximation to derive these values, they are still artificial. The LQ parameters α_x and β_x can only be determined from in–vitro cell survival experiments, whereas an effective, cohort–averaged $(\alpha/\beta)_x$ ratio can also be obtained from clinical studies (Stavrev et al. 2015).

Instead of optimizing a cost function of the biological effect, the RWD (i.e., RBE×Dose) can be used. The advantage of the RWD method is that the separate knowledge of α_x and β_x is not needed. The disadvantage is an increased computational effort (see equation 10.3). Therefore, the gradient of the RBE is typically neglected in the optimization, justified with its assumed small contribution to the total gradient (Mairani et al. 2013). However, optimizing the biological effect allows us to take the full gradient in the optimization procedure into account and therefore allows for a fast optimization. The RBE depends on the $\mathrm{LET_d}$ and the dose, thus minimizing the $\mathrm{LET_d}$ or the dose in an OAR separately does not necessarily result in a lower biological effect. Hence, the biological effect optimization is a possibility to directly account for that and additionally for the sensitivities of different tissue types.

Additionally to the uncertainties in RWD, proton therapy suffers from setup and range uncertainties (Lomax 2008). These uncertainties are typically accounted for in the derivation of the PTV margins (Park et al. 2012). In order to minimize the clinical conse-

quences of these uncertainties, robust treatment planning techniques are currently investigated (Casiraghi et al. 2013, McGowan et al. 2015, Ammazzalorso et al. 2014, Hopfgartner et al. 2012). However, robust treatment planning could also improve by taking biological uncertainties in account.

This study showed only 2 exemplary patient cases. For generalization of the results, further studies investigating more patients are required.

10.5 Conclusion

In this study the uncertainties of the RWD estimation in the framework of the Wedenberg model have been investigated for exemplary cases of nasopharyngeal and prostate cancer tumors. For the nasopharyngeal cancer patient, the uncertainty of $(\alpha/\beta)_x$ dominated. Nevertheless, the uncertainty originating in q was not negligible. For both fractionation schemes in the prostate cancer patient, the uncertainty of either of the two sources of uncertainty was of comparable magnitude. Therefore, neglecting the uncertainty originating in the fit to experimental data would lead to a considerable underestimation of the total uncertainty. Biological modeling can hence benefit by a reduction of both q and $(\alpha/\beta)_x$ uncertainties.

In the CTV, the RWD1.1 was higher than the RWDw in the nasopharyngeal cancer patient, but within the total uncertainty band. Although dose and $\mathrm{LET_d}$ were comparable in the standard fractionation TP of the prostate cancer patient (low $(\alpha/\beta)_x$), the RWD was significantly underestimated by the RBE = 1.1 approximation. However, in the hypofractionation plan for the prostate cancer patient, RWD was significantly overestimated by the RBE = 1.1 approximation. In contrast to the CTVs, the $\mathrm{RWD_{1.1}}$ significantly underestimated the $\mathrm{RWD_w}$ in all OARs apart from the optic nerve with respect to the total uncertainties.

In the conservative optimization strategy presented in this paper, the worst case tissue specific $(\alpha/\beta)_x$ was used in the biological effect optimization. The $\mathrm{RWD_w}$ to OARs could be reduced while yielding a homogeneous $\mathrm{RWD_w}$ distribution in the target. Since it is unlikely that the accuracy of the $(\alpha/\beta)_x$ ratio increases in a short term, a possibility to go beyond RBE = 1.1 could be to consider the uncertainties in a robust optimization strategy.

Acknowledgment

This work was supported by the Federal Ministry of Education and Research of Germany (BMBF), Grant No. 01IB13001 (SPARTA), and by the German Research Foundation (DFG) Cluster of Excellence Munich-Centre for Advanced Photonics (MAP). This work was additionally supported by the Bavaria California Technology Center (BaCaTeC). We further acknowledge support by Dr. Kerstin Hofmann (Department of Radiation Oncology, Technical University of Munich) in the usage of CERR and Dr. Barbara Knäusl (Division Medical Radiation Physics, Department of Radiation Oncology, Medical University of Vienna/AKH Wien) for helpful discussions.

Chapter 11

GEANT4 hadronic models for prompt–gamma emission

The material contained in this chapter was published in Physics for Medicine and Biology, in volume 59, in issue 7, in April 2014, with the title *Assessment and improvements of GEANT4 hadronic models in the context of prompt–gamma hadrontherapy monitoring*, by Dedes et al. (2014).

11.1 Introduction

The utilization of ions for radiotherapy purposes provides significant advantages in comparison to conventional x–ray radiotherapy. Massive charged particles have a maximum energy deposit close to the end of their range. This, combined with an energy dependent finite range, results in a sharp energy deposit called the Bragg peak. The characteristic dose–depth profile of ions allows for a high conformation of the dose to the tumor volume. Consequently, the main advantage of hadrontherapy is the potential to significantly reduce dose to healthy tissues, while delivering the prescribed dose to the malignant tissue. Furthermore, in the case of carbon ions, the increased biological effectiveness in the Bragg peak region is advantageous when treating radioresistant tumors.

The sharpness of the dose deposit which characterizes hadrontherapy, makes the online monitoring of the dose a major challenge. Several methods have been proposed for monitoring of the dose and/or the ion path in the patient. Positron emission tomography is the only modality used in clinical practice so far (Pawelke et al. 1997, Parodi et al. 2007) that provides control of the delivered dose after treatment. It takes advantage of the positron emission and annihilation, originated from β^+ radioactive nuclei that are produced by nuclear reactions in the patient's body. Methods which exploit other types of radiation are currently under study and development. Efforts are mainly invested in the usage of secondary proton and prompt–gamma emission, induced by nuclear reactions in the body of the patient during irradiation. Both for prompt–gammas (Min et al. 2006, Testa

et al. 2008, Testa et al. 2009) and for protons (Henriquet et al. 2012, Gwosch et al. 2013) correlation between the ion path and the emission profiles has been observed, which makes them promising candidate modalities for real–time hadrontherapy monitoring.

Monte Carlo simulation tools play an important role in the study and development of real–time monitoring systems for hadrontherapy. They are thoroughly used in feasibility studies, for detector optimization and finally they will be used during monitoring in order to identify deviations from the treatment plan. In the present work we use the GEANT4 toolkit (Agostinelli et al. 2003, Allison et al. 2006) in order to assess its performance and accuracy in the description of prompt-gamma emission from carbon ion irradiation. Similar studies, focusing on proton irradiation, have been done in the past. In Jarlskog & Paganetti (2008), physics settings for GEANT4 that best reproduce various experimental results related to proton therapy have been proposed. In Verburg et al. (2012), a detailed comparison between different Monte Carlo codes, dedicated nuclear reaction codes and experimental data of prompt–gamma emitted depth profiles and specific prompt–gamma spectral line cross sections is presented.

Having identified discrepancies between simulated and experimental data, we focus on the quantum molecular dynamics (QMD) model (Niita et al. 1995, Koi 2010) in order to improve GEANT4 predictions in hadrontherapy relevant observables. The QMD model describes the dynamic part of ion–ion collision and, interfaced to statistical de–excitation models, it produces the final states of the nuclear interaction. We use physical free parameters of the model, which are known to have a major impact on the nuclear reaction dynamics and products, in order to achieve better description of prompt–gamma yields obtained with various beam energies. As proposed by past theoretical studies (Hartnack et al. 1998, Maruyama et al. 1998) the main initial constraints of the parameters come from the optimization of the QMD description of well established nuclear properties such as nuclear density and nuclear binding energy. In our work, we perform this optimization mainly considering nuclei anticipated in a hadrontherapy application, rather than much heavier nuclei which were used in order to obtain the default values of the parameters used in the current GEANT4 implementation. We present the impact of our optimization compared to the available prompt–gamma experimental data. Furthermore, as we intend to improve the performance of GEANT4 concerning several typical hadrontherapy monitoring observables rather than focusing only on prompt–gamma emission, we also investigate the impact of our optimization on the prediction of charged particle emission.

11.2 Materials and methods

11.2.1 Experiments

The feasibility of hadrontherapy monitoring by means of prompt–gamma detection has been demonstrated and confirmed by several experimental studies (Min et al. 2006, Testa et al. 2008). Some of those experiments have been performed by our group at various facilities. Experiments using carbon ion beams at the Grand Accélérateur National d'Ions Lourds (GANIL, Caen, France), the GSI Helmholtzzentrum für Schwerionenforschung (GSI, Darmstadt, Germany), the Heidelberger Ionenstrahl–Therapiezentrum (HIT, Heidelberg, Germany), and using proton beams at the Westdeutsche Protonentherapiezentrum Essen (WPE, Essen, Germany) and at the HIT.

The main common characteristic of all the aforementioned experimental setups is the detection of prompt–gammas emitted from the target, at an angle of 90° with respect to the beam axis. For this purpose we have used scintillating detectors behind a collimator in order to scan the target along the direction of the impinging beam. Measurements have been also performed at 60° and 120° (Testa et al. 2009). Those experiments have shown that the signal to background ratio could be improved when avoiding forward angles, due to the fact that prompt–gammas are emitted isotropically while massive particles are mainly oriented towards the forward region. However, the geometrical simplicity of sampling at 90° with respect to the beam axis when aiming to obtain a depth profile, made the perpendicular geometrical layout the preferred selection.

CATANIA 80 MeV/u ^{12}**C**

This experiment was performed at the Laboratori Nazionali del Sud (LNS) Instituto Nazionale di Fisica Nucleare (INFN) di Catania, by a collaboration of Italian groups from Rome, Frascati and Catania (Agodi et al. 2012). An array of four LYSO detectors of $15 \times 15 \times 15\,\mathrm{mm}^3$ was placed at 90° with respect to the beam line, at 740 mm from a $40 \times 40 \times 40\,\mathrm{mm}^3$ PMMA target (approximate chemical composition as atomic ratio: 8% H, 60% C, 32% O).

GANIL 95 MeV/u ^{12}**C**

A barium fluoride (BaF_2) detector, of hexagonal shape with 50 mm edge and 140 mm length, was positioned along with some additional lead shielding at 605 mm from the beam axis. A 200 mm thick collimator with a 2 mm slit also made of lead was placed between the target and the detector. The target consisted of 27 PMMA slices $50 \times 50 \times 2.2\,\mathrm{mm}^3$ each. In order to suppress most of the neutron background, an energy threshold of 2 MeV on the collected energy in the detector and a time window of 2.7 ns in the time–of–flight (TOF) spectrum (around the prompt–gamma peak) were used. Finally, a background subtraction method based on the TOF measurement has been developed to retrieve net gamma yields. The data were first presented in Testa et al. (2010) and the results of the newest analysis used here were published later in (Pinto et al. 2015).

GSI 310 MeV/u ^{12}**C**

A BaF_2 detector, similar to the one described above was positioned at a distance of 1345 mm from the beam axis with a 200 mm thick lead collimator between the target and the detector. The target consisted of five plastic flasks filled with water. Each flask had dimensions of $120 \times 250 \times 40\,mm^3$. Special care was taken to prevent open spaces between flasks. The collimator had a slit of 4 mm. Some lead shielding was added to the setup and several water containers were placed after the collimator and before the detector in order to slow down neutrons, and thus to better identify prompt–gammas by TOF. An energy threshold of 2 MeV on the collected energy in the detectors and a time window of 3 ns were applied. The same background subtraction as for the GANIL experiment was used. The data were first presented in Testa et al. (2010) and the results of the newest analysis used here were published later in (Pinto et al. 2015).

WPE 160 MeV **p**

This experiment has been performed at the Westdeutsche Protonentherapiezentrum Essen (WPE) by IBA and ULB (Université Libre de Bruxelles) (Smeets et al. 2012). A cylindrical PMMA target of 75 mm radius and 200 mm length has been irradiated with a 160 MeV proton beam. At 90° with respect to the beam direction and the PMMA cylinder axis, a cylindrical NaI crystal was placed at a distance of 500 mm. The dimensions of the crystal were 76.2 mm in length and diameter and the cylinder was oriented with its axis being perpendicular to the beam direction. A 100 mm thick lead wall was placed between the target and the detector, at a distance of 200 mm from the target axis. In the lead wall a $52 \times 50\,mm^2$ hole was either open or blocked, so as to register the two respective energy spectra. Finally, a net energy spectrum was calculated by the subtracting the open/closed wall spectra.

In the next sections we compare data from the aforementioned experiments with GEANT4 Monte Carlo simulations. In order to perform this comparison in a reliable manner, we have modeled the experimental setup in the following way: rather than an ideal particle detection in the sense of counting particles entering a volume that simply matches the geometrical characteristics of the actual detector, therefore taking into account only the geometrical acceptance, we have modeled the detector material and all the interactions of every particle that reaches the simulated detector. That means that for every simulated particle that enters the detector volume, all interactions are simulated and both the initial impinging particle and all the created secondaries are tracked until they get absorbed or escape the detector volume. What is registered is the total energy deposit in the detector and the TOF of the initial particle. The analysis of the simulations follows closely what was done for measured data. The selection of an event in the detector is done by applying the same criteria as for the measurements, on the total deposited energy and TOF, without any other a priori identification of the impinging particle. The results simulated are then compared to the experimental data (again hits in the detector represented by an energy

deposit and a registered TOF), which are calibrated and corrected for the dead time.

The scope of this study is to assess the performance of the hadronic models available in the GEANT4 toolkit in the context of hadrontherapy applications and furthermore to propose improvements.

11.2.2 GEANT4 hadronic models

General description

The GEANT4 simulation toolkit offers a large variety of physics models for the simulation of several processes which can occur during the transportation of particles through matter. In addition to that, there is quite often more than one model describing the same process (either complementary in terms of energy regime validity or valid for the same energy range). Therefore the user can select not only among the physical processes that are essential for his/her study, but also choose among alternative description/models for a given physical process. In the current work we will mainly discuss the influence of the hadronic nuclear models on the predictive performance of GEANT4 in the context of the prompt–gamma hadrontherapy monitoring. In table 11.1 we present the theoretical models used in this study for the simulation of a low energy nuclear collision, grouped in three phases which describe such an interaction: the collision, the equilibration and the de–excitation phase.

In the simulation of the proton beam experiment we have used the binary cascade (BIC) model. In the simulation of our carbon beam experiments we have used two of the available nuclear models, the QMD model (Niita et al. 1995, Koi 2010) and the binary light ion (BLI) reaction model[1]. We should note here that the QMD model, contrary to the BLI model, does not make use of the precompound model (see footnote 1) as a transition to the de–excitation phase. For both cases the de–excitation phase is the same. A detailed table of the hadronic physics lists used can be found in appendix D.0.1. In the next subsection we give a brief description of each of the aforementioned models.

Table 11.1: Different phases of a nuclear reaction in GEANT4 and the corresponding theoretical models used in the current study. 'Ions' refers to all the charged particles heavier than protons in the framework of this study.

Particle	Collision phase	Equilibration phase	De–excitation phase
Ions	Quantum molecular dynamics	∅	Statistical multi–fragmentation Fermi break–up, Evaporation
	Binary light ion reaction	Precompound	
Nucleons	Binary cascade		

[1]http://geant4.web.cern.ch/geant4/UserDocumentation/UsersGuides/PhysicsReferenceManual/BackupVersions/V9.4/fo/PhysicsReferenceManual.pdf

Quantum molecular dynamics (QMD) model

The collision process which is handled by QMD is described in a collective/dynamic approach. All the nucleons in the projectile and target nuclei are considered as participants in the process and are tracked simultaneously during the collision. Each nucleon state is represented by a Gaussian wave function $\phi(\boldsymbol{r})$ of width $\sqrt{L}$,

$$\phi(\boldsymbol{r}) \equiv \frac{1}{(2\pi L)^{3/4}} \cdot exp\left(-\frac{(\boldsymbol{r}-\boldsymbol{r}_i)^2}{4L} + \frac{\mathrm{i}}{\hbar}\boldsymbol{r}\cdot\boldsymbol{p}_i\right), \tag{11.1}$$

with i being the imaginary unit, and $\boldsymbol{r}_i$ and $\boldsymbol{p}_i$ annotating the centers of position and momentum of the i-th nucleon. The total wave function of the system is considered to be the direct product of all the nucleon wave functions participating to it. During a collision, the equation of motion of each nucleon is described by Newtonian equations and the stochastic nucleon–nucleon collision term. The Newtonian equations are derived by the Hamiltonian (H) derivatives on momentum and position. The stochastic nucleon–nucleon collision term is the free body collision between two nucleons, under the condition that it is not Pauli blocked.

It is important here to provide some additional information about the terms of the QMD Hamiltonian. The Hamiltonian $H = T + V$, consists of the kinetic energy term T and the potential energy term V. The potential energy term is further decomposed to a Skyrme type potential (Skyrme 1958), a Coulomb potential related to electrostatic interactions and a symmetry term which accounts for the instability of nuclei due to the difference between proton and neutron numbers. The Skyrme interaction is characterized by an effective potential which reduces the description of N–N body interactions to a two-body contact interaction (closely related to that of the free nucleon–nucleon scattering) and a three–body repulsive contact interaction, which describes the effect of the presence of additional nucleons on the two–body interactions. Averaged over all nucleons, the net result of Skyrme type interactions is equivalent to a two–body nuclear density dependent interaction.

For the calculation of a collision event, QMD implemented in GEANT4 follows the subsequent steps:

- The projectile and the target nuclei are initialized in their ground states.
- The constituents of the system are propagated under the influence of their mutual potential, taking into account the hard scattering interactions:
  - calculation of interaction densities, forces and the Hamiltonian
  - propagation of all particles according to Hamilton's equation of motion
  - perform all hard scatterings within the hardcoded time limit
  - decision for each hard scattering whether its final state is Pauli blocked.

- Clustering of all the participant nucleons to excited fragments, according to their momenta and distances. Excited fragments created after clustering are forwarded to the de–excitation part.

The clustering distance ΔR^2 is given by the following equation:

$$\Delta R^2 = (\boldsymbol{r}_i - \boldsymbol{r}_j)^2 \left(1 + \gamma_{ij}^2 \beta_{ij}^2\right) \leq R^2 \tag{11.2}$$

where γ_{ij} is the Lorentz factor, β_{ij} the ratio of the speed of the j-th particle to the speed of light with respect to the i-th particle in the cluster and R^2 a distance cut–off value. In the end of this iterative algorithm some of the nucleons are identified as clusters (heavy or light nuclei) and some are left isolated (secondary protons and neutrons).

Binary cascade (BIC) model

BIC (Folger et al. 2004) treats the collision of nucleons with nuclei. It is an intranuclear cascade approach in which the primary projectile nucleon and the secondary particles are propagated in the nucleus. Interactions take place exclusively between a primary or a secondary projectile–like particle and an individual nucleon of the nucleus, contrary to the collective and simultaneous approach of QMD where the nucleons of the target nucleus can interact with each other, through the Skyrme type potential.

The projectile particle is transported towards the target nucleus. Once inside the nucleus, the projectile nucleon has its energy corrected taking into account Coulomb effects. Instead of using a self–generating potential as in QMD, in the case of BIC the collective nuclear effect of an homogeneous nucleus upon primary and secondary particles is approximated by an effective time independent nuclear potential. Particles are propagated inside this nuclear field by numerically solving the equation of motion. The basis of the description of the reactive part of the scattering amplitude is the two–particle binary collision. The resulting secondaries are checked for the Pauli exclusion principle. If a secondary particle has a momentum lower than the Fermi momentum, the interaction is suppressed and the original primary particle is tracked to the next collision. If the interaction is allowed, the secondaries are treated as new primaries, interacting in the target nucleus. The cascade finally terminates when the average and maximum energy of secondaries are below a given threshold. The remaining fragment is treated by the precompound and de–excitation models. More details can be found in the GEANT4 Physics Manual, chapter 26.

Binary light ion (BLI) reaction model

This is an extension of the BIC for light ion reactions. It is a hybrid model between a classical cascade model and the QMD description. In analogy to the BIC, the initial cascade is prepared in the form of two detailed three–dimensional nuclei. The lightest of the collision partners is selected to be the projectile. The nucleons of the projectile are then entered, using their positions and momenta, into the initial stage of the cascade. The

participant nucleons that are transported into the target nucleus have again exclusively binary scattering with the target's internal nucleons.

It bears a similarity to QMD in the sense that here also, each participant is described by a Gaussian wave packet and the total wave function is assumed to be the direct product of each individual wave function. Nevertheless, in the BLI the Hamiltonian from which the equations of motion are derived is calculated from a simple time independent optical potential. One additional and essential difference compared to QMD is that a participant particle in BLI is either a primary particle, i.e. a nucleon in the projectile/target nucleus, or a secondary particle generated in the collision process, which is also propagated in the target nucleus. But scattering between participant particles of the same system (either of the target or the projectile) is not taken into account. As in the case of BIC, secondary products of the cascade are checked by Pauli's exclusion principle before stored.

Precompound model

The precompound model can provide either a possibility to extend at low energy range the hadron kinetic model for nucleon–nucleus inelastic collisions, or to act as a smooth transition phase between the kinetic stage of the reaction (i.e. BLI) and the equilibrium stage where the equilibrium de–excitation models take over. In the case where the precompound model is used as a low energy dynamic part of the collision, the recommended maximum single nucleon projectile energy is 170 MeV. When using BIC or BLI models, the passage to the precompound model as an equilibration phase of the nuclear reaction is decided as follows: as long as there are still particles above the kinetic energy threshold of 75 MeV, cascade will continue. If this is not the case and when the mean kinetic energy of the participants is lower than 15 MeV, the cascade stops and the precompound model handles the reaction.

The initial information required for the calculation of the precompound nuclear stage is the mass number A, the atomic number Z, the four–dimensional momentum–energy vector $\boldsymbol{P}_0$, the excitation energy U and the numbers of excitons, holes and charged excitons. The exciton number is $n = p + h$, where p is the number of particles and h the number of holes. All cascade nucleons with kinetic energies above the Fermi energy are called particles. The holes are results of cascade interactions which occupy states below the Fermi energy and occur when particles strike nucleons from a nucleus. A more comprehensive description on the simulation of the precompound reaction and the modeling of pre–equilibrium emissions of n, p, d, t, ^{3}He and α particles can be found in the GEANT4 Physics Manual.

De–excitation

After the collision and equilibration phase (wherever available), nuclear fragments have reached a state of statistical equilibrium with a non–zero excitation energy. The next step is to de–excite those fragments according to the different processes that depend on A, Z and excitation energy. This phase is handled in GEANT4 by the Excitation Handler class. It is responsible for dispatching individual fragments to the models that perform the actual

de–excitation of the compound nucleus, based on various parameter ranges of applicability of the models (depending on the atomic and mass numbers, excitation energy etc). In the end of the de–excitation process a list of secondary single particles (i.e. nucleons), and nuclear fragments in a final state of zero excitation energy are produced. The Excitation Handler manages five de–excitation models: evaporation (of light particles), fission, Fermi break–up (FBU), multi–fragmentation and photon evaporation.

The handling of all available de–excitation models in GEANT4 version 9.4 is the following: each product of the nuclear interaction is forwarded to the handler which takes the decision of the model route to be followed. If the fragment is a single nucleon (no de–excitation can take place) it is stored directly as a final product of the collision. In case it is not a nucleon, several criteria are checked: isotope abundance, excitation energy per nucleon, A and Z of the fragment. Depending on those criteria, the fragment follows a two–step approach. The first is a hierarchical linear course applied only once: if the fragment has an excitation energy lower than a given threshold (1 keV) and it is a stable isotope, it is considered a final product. Otherwise, different mechanisms are tried in the following sequence: FBU, then statistical multi–fragmentation (SMF) and finally evaporation. The fragment that is handled by any of those three processes breaks/creates one or more de–excitation products which are stored in a temporary list.

This list is then used for the second step. The only difference in this stage compared to the first one, is that SMF is not available and that the whole process is applied in an iterative approach until only final state products are obtained (no possible further de–excitation). This second step is: unless the temporary product is a single nucleon then first FBU is tried and if it is not applicable, then evaporation. It is constantly checked if a fragment other than a single nucleon has energy lower than a given threshold (1 keV) and if it is a stable isotope (to be considered again as a final stable product). In the end of those two steps, the first applied only once and the second iterative, there are only final stable products that are passed to the transportation algorithm.

It is useful to describe briefly each de–excitation model that was mentioned above and also state the criteria applied to the candidate fragment in order to qualify for each of those de–excitation models.

The FBU model predicts final states as results of an excited nucleus with mass number $A < 17$ and atomic number $Z < 9$. The model describes the situation of light nuclei where the excitation energy per nucleon is often comparable with nucleon binding energy. Therefore the light nucleus can break up in two or more fragments whose types and properties depend on the available phase space. For the break–up of such a nucleus to be possible, there should be a sufficient excess of energy in the system, which will be attributed to all fragments of a given break–up configuration as kinetic energy. The calculation of the available energy (E_{kin}) takes into account the mass and the excitation energy of the initial excited nucleus, the excitation energies and the masses of the potentially produced fragments and finally, the Coulomb barrier. This energy is derived by the following equation:

$$E_{\mathrm{kin}} = U + M(A,Z) - E_{\mathrm{Coulomb}} - \sum_{b=1}^{n} (m_b + \epsilon_b) \tag{11.3}$$

where the Coulomb energy is approximated by:

$$E_{\mathrm{Coulomb}} = \frac{3}{5}\frac{e^2}{r_o}\left(1 + \frac{V}{V_o}^{-1/3}\right)\left(\frac{Z^2}{A^{1/3}} - \sum_{b=1}^{n}\frac{Z^2}{A_b^{1/3}}\right) \tag{11.4}$$

U and $M(A,Z)$ are the excitation energy and the mass of the initial excited nucleus, n is the number of potential fragments, m_b and ϵ_b are the masses and the excitation energies of the potential fragments, V_o is the volume of the system corresponding to the normal nuclear density and $\kappa = V/V_o$ is a free parameter. The default value of this parameter is 6 and the physical meaning of V is the freeze–out volume of the expanding excited nucleus. The probability (statistical weight) for a nucleus to break up into n components (nucleons, deuterons, tritons, alphas etc) is given by:

$$W_{E_{\mathrm{kin}},n} = S_n G_n \frac{V^{n-1}}{\Omega}\left(\frac{1}{\sum_{b=1}^{n} m_b}\prod_{b=1}^{n} m_b\right)^{3/2} \cdot \frac{(2\pi)^{3(n-1)/2}}{\Gamma(3(n-1)/2)} E_{\mathrm{kin}}^{\frac{3n-5}{2}} \tag{11.5}$$

where S_n is the spin factor (gives the number of states with different spin orientations), G_n is the permutation factor (takes into account the identity of components in the final state), $\Omega = (2\pi\hbar)^3$ is the normalization volume, and Γ is the gamma function.

The SMF model is used to predict final states resulting from a highly excited nucleus. The basic criterion to be fulfilled by the excited nucleus is that its excitation energy should exceed 3 MeV/u. At those excitation energies the mechanism of a complete break of the nucleus into fragments becomes dominant. The excited primary fragments propagate independently in the mutual Coulomb field and undergo further de–excitation. Further details concerning SMF are omitted from the current report due to the fact that this de–excitation channel was not available in the GEANT4 version used for this study (9.4). Comprehensive information on the model can be found in GEANT4 Physics Reference Manual.

The evaporation model handles an excited nucleus which is characterized by its mass number A, atomic number Z and excitation energy U. Depending on those characteristics the nucleus can eject single nucleons (p, n), light fragments (d, t, ^{3}He, α) and photons. The emission of particles by an excited compound nucleus has been successfully described by comparing with the evaporation of molecules from a fluid. The first statistical treatment of compound nuclear decay is due to Weisskopf & Ewing (1940). In that approach the probability to go from a state i to another state d and vice versa is related to the density of states of the two systems (detailed balance principle):

$$\rho_{\mathrm{i}} P_{\mathrm{i}\to\mathrm{d}} = \rho_{\mathrm{d}} P_{\mathrm{d}\to\mathrm{i}} \tag{11.6}$$

Therefore the probability that a parent nucleus i with and excitation energy U emits a fragment j in its ground state with kinetic energy ϵ is:

$$P_{\mathrm{j}}(\epsilon)\mathrm{d}\epsilon = g_{\mathrm{j}}\sigma_{\mathrm{inv}}(\epsilon)\frac{\rho_{\mathrm{d}}(E_{\mathrm{max}}-\epsilon)}{\rho_{\mathrm{i}}(U)}\epsilon\mathrm{d}\epsilon \tag{11.7}$$

where $\rho_{\mathrm{i}}(U)$ is the level density of the evaporating nucleus, $\rho_{\mathrm{d}}(E_{\mathrm{max}})$ is that of the residual nucleus after emission of a fragment j and E max is the maximum energy carried by the ejectile. g_{j} is a coefficient taking into account the spin and the mass of the ejectile, and σ_{inv} is the inverse reaction cross section. The inverse cross section reaction is expressed by means of an empirical equation, fixing parameters which give a good fit to the continuum theory. In addition, for charged ejectiles, the Coulomb barrier is taken into account. An alternative model is the generalized evaporation model (GEM), which also considers the emission of heavier nuclei than α particles and uses a more accurate level density function for the total decay width. In the current study, the GEM was used. More information can be found in GEANT4 Physics Reference Manual.

The two remaining de–excitation channels, fission and photon evaporation, are incorporated (nested calls) in the evaporation channel, although they can be also called separately. Those channels are competitive to particle/light fragment emission. More specifically, fission takes place for excited nuclei with mass number $A > 65$, therefore it is not of great importance for our current studies. On the contrary, photon evaporation is an essential de–excitation mechanism for the hadrontherapy energy and projectile–target combination domain, and especially for the prompt–gamma monitoring application. During the calculation of each channel's evaporation probability, photon emission is one of the candidates taken into account. Despite the very low probability for relatively high excitation energies, this channel can still in principle compete with particle emission. In case evaporation is selected, the new nucleus is iteratively processed by the excitation handler in order to be further de–excited, unless it is already in its ground state. For each iteration, probabilities are recalculated and a new evaporation channel is selected. Nevertheless, the impact of photon evaporation is much more important when the excited nucleus has no other de–excitation channel available than photon emission. In that case, instead of emitting only one photon and recalculating all the channel probabilities, a cascade of photon emissions takes place.

There are two different mechanisms in GEANT4 that describe photon emission. The first is the continuous photon emission. This mechanism is in competition with the particle emission channels and the term continuous characterizes the emitted gamma energies (not quantized). Only giant dipole resonances (GDR, mass–charge center separation) E1-transitions are considered as the main source of emission from highly excited nuclei. The probability of such an emission is related to the inverse reaction cross section (photoabsorption) which is given by the expression:

$$\sigma_{\gamma}(\epsilon_{\gamma}) = \frac{\sigma_o\epsilon_{\gamma}^2\Gamma_{\mathrm{R}}^2}{\left(\epsilon_{\gamma}^2 - E_{\mathrm{GDR}}^2\right)^2 + \Gamma_{\mathrm{R}}^2\epsilon_{\gamma}^2} \tag{11.8}$$

where $\sigma_o = 2.5A$ mb, $\Gamma_{\mathrm{R}} = 0.3E_{\mathrm{GDR}}$ and $E_{\mathrm{GDR}} = 40.3A^{-1/5}$ MeV are empirical parameters of the GDR. The second type of (photon) evaporation process consists of discrete photon emission. Excited fragments whose excitation energy is below the highest level of a discrete gamma transition, de–excite via a cascade of photon emissions of discrete energies. Those photon emissions follow tabulated nuclear levels and E1, M1 and E2 transitions. At this step, competition between discrete energy photons, GDR photons and particle evaporation is neglected. For the energy range and nuclei type of hadrontherapy the main source of photons from nuclear de–excitation comes from the discrete emission mechanism (approximately an order of magnitude higher than from the continuous one). It is important to note that, usually, excited fragments within the discrete nuclear levels energy domain do not initially have exactly the energy of one of the tabulated nuclear levels. In the GEANT4 implementation of discrete transitions this is solved by selecting the nearest nuclear level as a starting excitation energy and then performing one of the allowed discrete transitions. For this reason, an energy tolerance is used as a criterion applied when seeking the nearest level. This tolerance is by default equal to 10 TeV, which means that it is always possible to find an initial nuclear level and then proceed to the allowed transitions according to the tabulated data. The discrepancy between experimental prompt–gamma yields and GEANT4 simulated yields has been also addressed in the past (Lestand et al. 2012), using the value of the tolerance as a free parameter. In this approach, although the desired reduction in prompt–gamma emission is achieved when significantly reducing the energy tolerance value (100 keV), it leads to the problem of residual excited fragments. Those fragment are in that method artificially considered as being in the ground state, although they are effectively left infinitely into an undefined excited state.

11.2.3 Parameters for QMD model optimization

The calculation of an ion–ion collision described by the QMD model and its implementation in the GEANT4 toolkit involve a large number of parameters. Those parameters can be divided into three classes, as described in Hartnack et al. (1998): reaction parameters, which define the kinematics of a single event (projectile and target masses, energy etc). Physics parameters that correspond to a detailed description of the interactions (interaction range, potential parameters etc). Those physical parameters may be changed within a reasonable range and their deduction is a particular goal of the comparison with experiments. Technical parameters (time step size, total reaction time etc), that are used for increased computing time performance of the calculations and should not affect the observables. After a thorough study of the model as implemented in the GEANT4 toolkit, we have identified two key parameters that have a considerable effect on the outcome of a simulated collision and therefore to the final state products occurring after the de–excitation phase. Those two quantities are the square of the width of the Gaussian wave packet L describing each nucleon state (see equation 11.1), and the maximum distance criterion R (see equation 11.2) used to assemble the excited nuclei after the collision phase, which will be forwarded to the de–excitation models. The user can access those parameters only from the GEANT4 source code. L is defined in

/source/processes/hadronic/models/qmd/src/G4QMDParameters.cc, as the wl variable, with the default value of $2\,\mathrm{fm}^2$ (described in the code as having units of length, instead of the correct squared length, as in the original publication (Niita et al. 1995)). R is defined in /source/processes/hadronic/models/qmd/src/G4QMDMeanField.cc as the rclds variable, described as the distance for cluster judgment with a default value of 4 fm.

The wave packet width $\sqrt{L}$ in fm should, in approximation, be in the order of the range of the strong interaction, namely in the order of the fm. Actually the strong force is repulsive at distances lower than 0.8 fm (i.e. strong force between quarks in a nucleon), attractive at larger distances (residual nuclear force between nucleons) up to about 3 fm where it becomes negligible, reaching a maximum strength approximately at $0.8-1$ fm. The parameter L determines the effective interaction range of the nuclear potential, thus influencing the final states of the collision. As further explained in (Niita et al. 1995, Hartnack et al. 1998, Maruyama et al. 1998), the value of L is a free parameter of the QMD model, and as a consequence it can be optimized for certain configurations of projectile/target nucleus species. Apart from the description of the collision, the interaction range has a deep impact on the description of nuclear properties of a single nucleus, such as the surface properties as well as the binding energy of the modeled nucleus. Those properties, alongside with experimental data describing particle fluences, can be used for the determination of the most appropriate value of the parameter. For heavy ion collisions relatively high values of L are used, while for Ca–Ca and lighter systems values down to $1\,\mathrm{fm}^2$ have been used to various QMD model flavors. Note that there is a factor of 4 difference in the definition of values of L between the mathematical formalism of Hartnack et al. (1998) and the formalism used both in the present paper and in the GEANT4 QMD implementation. Therefore a value of $4\,\mathrm{fm}^2$ in Hartnack et al. (1998) is equivalent to a value of $1\,\mathrm{fm}^2$ in the GEANT4 formalism. Throughout this work we will be using the GEANT4 notation, mathematically consistent with the form of equation 11.1.

Our adaptation of QMD for better performance in the domain of hadrontherapy is based on three types of observables:

- In regard to prompt–gamma emission we examine two types of yields. The first one concerns the prompt–gamma depth profiles, compared to the ones obtained in the experiments described in section 11.2.1. The main aim of this work is to obtain the best possible agreement with the experimental prompt–gamma profiles. A second type of observable is the number of emitted prompt–gammas in the target, accessible only by Monte Carlo simulations. Although this quantity cannot be directly compared to measurements, it provides an indication of the influence of the adaptation of QMD on the total number of produced prompt–gammas.

- The modeling of an ion–ion collision starts with a basic description of the projectile and target nuclei. The parameters of the model, apart from the final states of the collision, also affect the description of the initial nuclei. Therefore, we utilize some of well established nuclear properties in order to obtain physical limits for the free parameters and additional justification of the selection of those parameters that provide the best agreement with the experimental data. In the current study the nuclear

properties considered are the binding energy per nucleon and the nucleon density of the nucleus.

- In order to obtain a more complete overview of the adaptation of QMD towards a better description of prompt–gamma emission, we also study the effect of the proposed changes on the emission of charged secondary particles. For this purpose we use experimental data that are relevant to hadrontherapy applications. Such measurements have been performed at GANIL, using a 95 MeV/u ^{12}C beam impinging on PMMA targets of various sizes (Braunn 2010, Braunn et al. 2011). We focus on the angular distributions of charged particles.

11.3 Results

11.3.1 GEANT4 performance

For the comparison between experimental data (section 11.2.1) and simulations, GEANT4 version 9.4 was used. In all cases, unless otherwise mentioned, both the particle transportation step size and the particles production cut (in range) used were the default 1 mm. For the hadronic interactions of ions with matter the QMD model was used (unless mentioned otherwise, in those cases the BLI model was alternatively tested). Proton and neutron hadronic interactions with matter were simulated by the BIC model. Especially for neutrons with energy lower than 20 MeV, the high precision models available in GEANT4 were used (`G4NeutronHPElastic`, `G4NeutronHPInelastic`, `G4NeutronHPCapture`, `G4NeutronHPFission`, see appendix D.0.1). Those models use tabulated cross sections for the elastic and inelastic interactions, fission and capture of neutron by different nuclei. Finally, for the electromagnetic interactions of particles with matter we have used the so–called `G4EmStandardPhysics_option3` GEANT4 models package.

At the time of the study, the publicly available GEANT4 version was the 9.4. Until the completion and submission of it, versions 9.5 and 9.6 were released. Although all the results presented in the current document were obtained with version 9.4, their validity in versions 9.5 and 9.6 was verified. In figure 11.1, the prompt–gamma emission in the cases of 95 and 310 MeV/u ^{12}C, impinging on PMMA or water target is shown (beam and target characteristics were selected so as to match those of the GANIL and GSI experiments, described in section 11.2.1). The emission profiles in versions 9.5 and 9.6 are compatible to those obtained with 9.4, within 10%. Therefore, the content and conclusions of this study, are relevant also for GEANT4 9.5 and 9.6.

In order to assess the performance of GEANT4 in the context of hadrontherapy monitoring by means of prompt–gamma detection, we have compared the energy spectrum obtained with an 80 MeV/u ^{12}C beam on PMMA target and the absolute yields of the depth profiles for ^{12}C beams of 95 MeV/u and 310 MeV/u using PMMA and water target respectively, with the predictions of GEANT4. For the case of proton beam, we have used the prompt–gamma energy spectrum obtained by irradiating a PMMA target with 160 MeV protons.

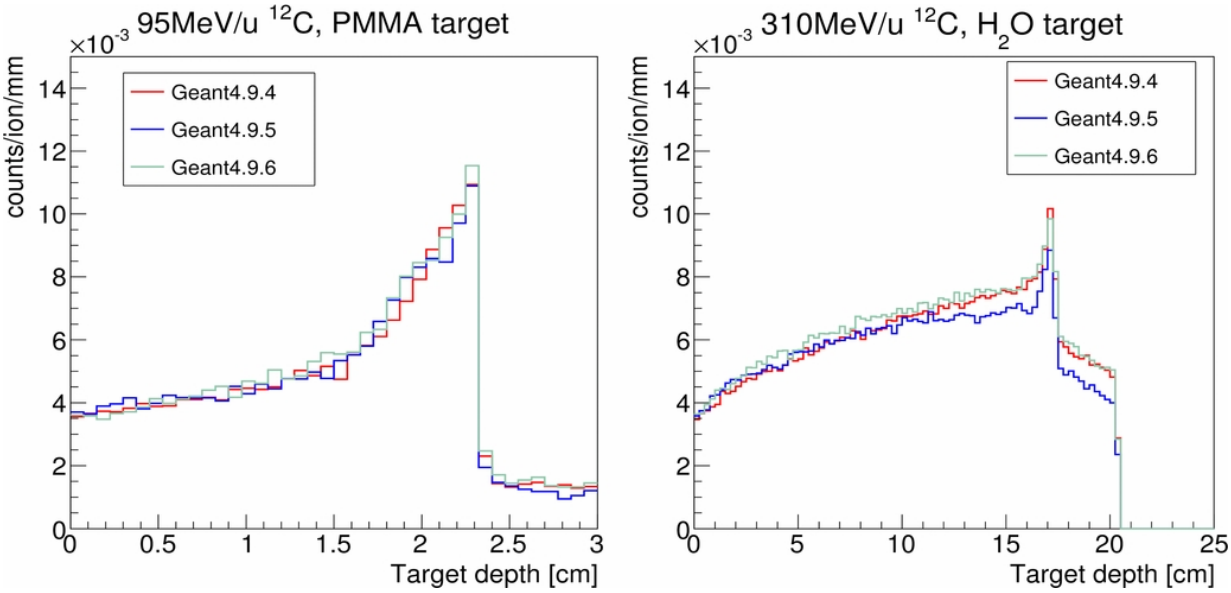

Figure 11.1: Comparison of the simulated prompt–gamma emission for three different GEANT4 releases (9.4, 9.5 and 9.6). Left, from a 95 MeV/u ^{12}C beam in a PMMA target. Right, from a 310 MeV/u ^{12}C beam in a water target (the steep drop after 20 cm is due to the geometrical limit of the target volume).

Throughout all our comparisons, as observed in figure 11.2, GEANT4 simulations consistently overestimate the prompt–gamma yields by a factor of about 1.8 to 2.8, over an energy range from 80 to 310 MeV/u for the case of ^{12}C. For the case of 160 MeV protons, the integral of the open/closed wall difference prompt–gamma energy spectrum from 1 to 10 MeV is overestimated by a factor of 1.7. From the presented results it is obvious that the GEANT4 nuclear models in their current status, cannot describe quantitatively the experimental profiles with sufficient accuracy.

In the next sections we proceed a step further towards the improvement of the existing models, in order to obtain a more realistic quantitative description of the experimental data. For this study we focus on the QMD ion–ion reaction model and therefore on the prediction of prompt–gamma yields from ^{12}C beams. This decision was based on the fact that QMD yields slightly better results than BLI (see figure 11.2 upper right), but most importantly because the QMD model is expected to provide a more realistic and complete description of the N–N particle system interaction, as described in section 11.2.2.

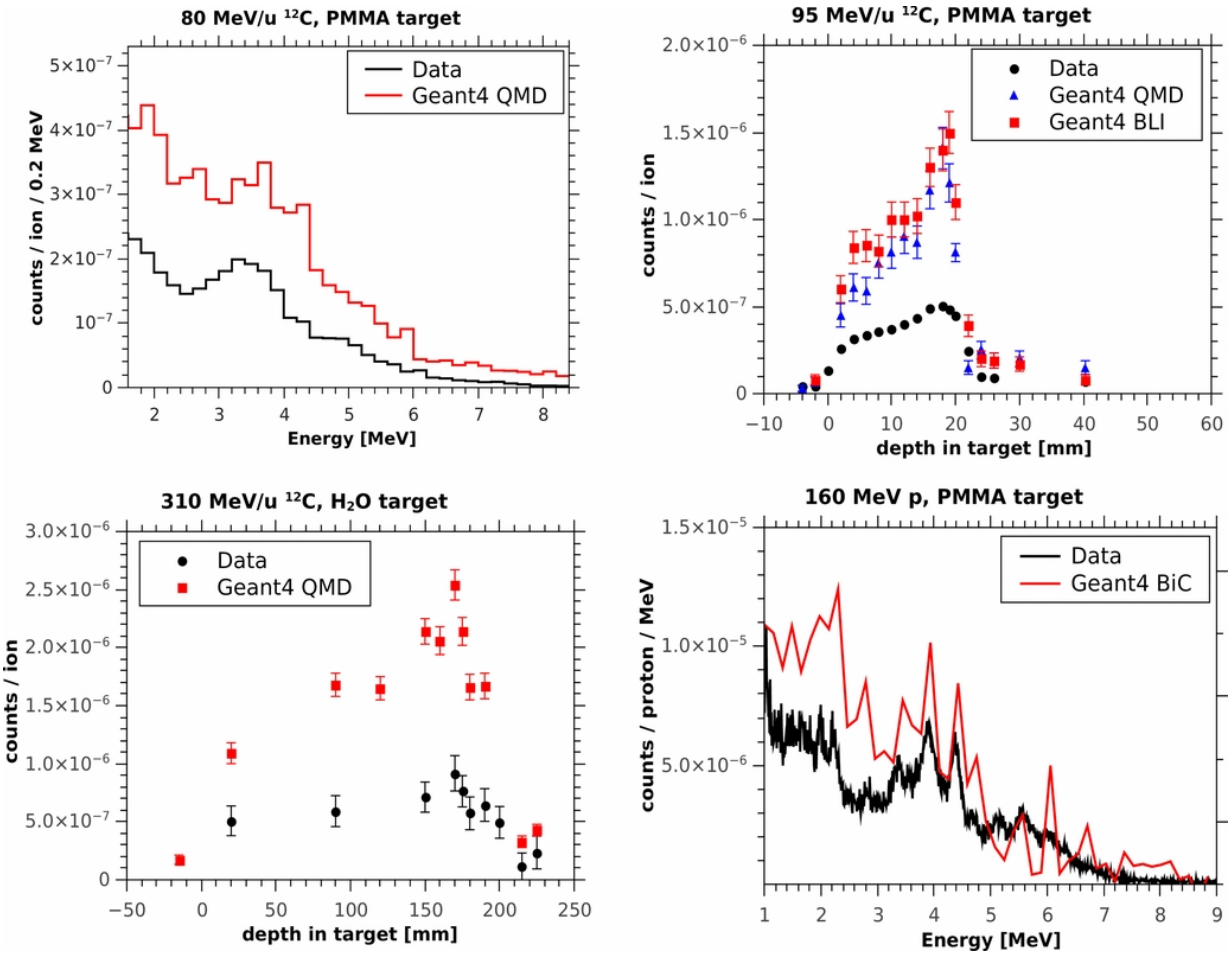

Figure 11.2: GEANT4 comparison with the experiments described in section 11.2.1. Upper left, the prompt–gamma energy spectrum obtained in Catania (Agodi et al. 2012) by irradiating a PMMA target with a 80 MeV/u ^{12}C beam. Upper right, the prompt–gamma depth profile obtained at GANIL by the irradiation of a PMMA target by a 95 MeV/u ^{12}C beam (systematic uncertainty of approximately 20%). Lower left, the depth profile obtained from a water target with a 310 MeV/u ^{12}C beam (systematic uncertainty of approximately 2%). Lower right, the open/closed wall difference prompt–gamma energy spectrum obtained by IBA at WPE (Smeets et al. 2012), by the irradiation of a PMMA target with 160 MeV protons. The error bars denote statistical error of one standard deviation.

11.3.2 Prompt–gamma emission

As a first step, we have studied the impact of L on the number of photons, protons, neutrons and alpha particles produced/emitted in a $50 \times 50 \times 50\,\mathrm{mm}^3$ PMMA target irradiated with 95 MeV/u ^{12}C beam (this was the target of the actual GANIL experiment). The results are presented in figure 11.3 in terms of secondary particles emitted in the target. The default value of the parameter in GEANT4 QMD is $L = 2\,\mathrm{fm}^2$. As seen in the results, using the default parameter value gives the maximum of the emitted gamma distribution. For the range scan of the L values shown in figure 11.3, no physical limitation was taken into consideration (some of the values depicted there be will later on excluded as unphysical), as at this stage it is crucial to stress the sensitivity of photon emission to the value of the parameter L. Two conclusions can be drawn from those results. First, that as the default value yields a maximum of gamma emissions in the target, any selection of either lower or higher values will reduce gamma emission. The second conclusion regards proton, neutron and alpha yields (chosen here as a representative of light nuclei emission). Proton and neutron yields seem almost unchanged when using $0.7\,\mathrm{fm}^2 < L < 2\,\mathrm{fm}^2$ while they increase dramatically when using $L < 0.5\,\mathrm{fm}^2$ or $L > 2\,\mathrm{fm}^2$. Alpha particle yields seem to be affected in a similar pattern as photons.

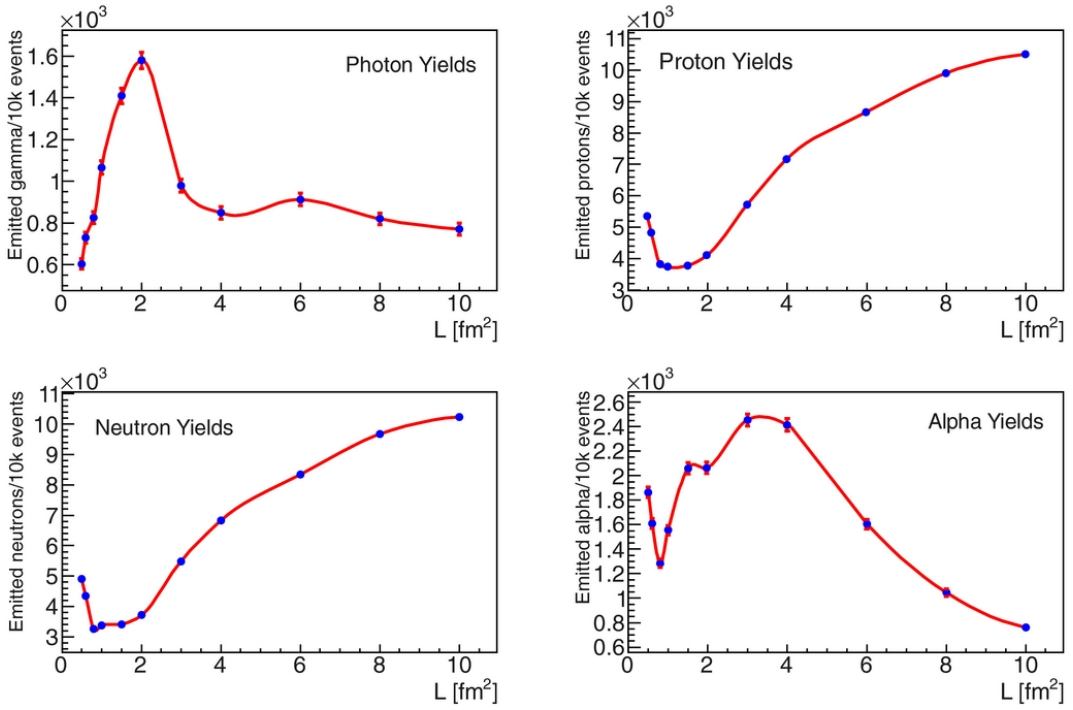

Figure 11.3: Emitted particles per 10^4 ^{12}C ions of 95 MeV/u impinging on a $50\times50\times50\,\mathrm{mm}^3$ PMMA target as a function of the nucleon wave packet width squared L. The error bars denote statistical error of one standard deviation. The absence of error bars signifies that one standard deviation would lead to bars smaller than the size of the point markers in the graph.

The variation of the L parameter influences the QMD reaction dynamics and consequently,the produced excited final states. The impact of the variation of L on the emitted prompt–gammas is further explained in figure 11.4 which describes several observables of 95 MeV/u ^{12}C and their secondaries' induced collisions in PMMA. On the left, the species of excited nuclear fragments created exactly at the end of the dynamic part of the reaction, just before the de–excitation process, are shown. In the middle, the distribution of excitation energies of all excited fragments produced by the dynamic (QMD) part and on the right the energy spectra of prompt–gammas produced in the target are shown. For lower L values, QMD produces lighter and less excited nuclear fragments that will be forwarded to the de–excitation phase. Consequently, those fragments require fewer de–excitations (usually gamma emissions) until they reach their ground state.

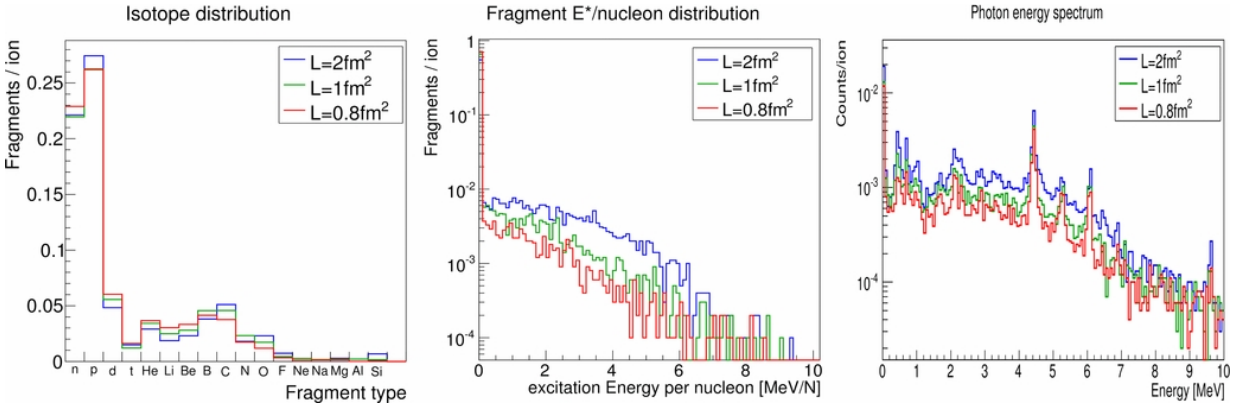

Figure 11.4: Left, the species of excited nuclear fragments created exactly at the end of the dynamic part of the reaction (QMD). Middle, the distribution of excitation energies of all excited fragments produced by the dynamic part (QMD). Right, the energy spectra of prompt-gamma produced in the target. All distributions are shown for different values of L.

The second parameter of QMD under study is the maximum distance R used in clustering of the nucleons after the collision. The process is applied after the lapse of 100 fm/c which is the time that each collision lasts. At this point the collision process freezes and all nucleons in the system (former projectile and target) are accessed and clustered according to their momenta and distances. All nucleons that lay at distances ΔR^2 smaller than R^2 and whose momentum difference is not exceeding a certain threshold are considered to belong to the same heavier fragment. The clustering distance ΔR^2 is described by equation 11.2. In figure 11.5 we present the distribution of ΔR^2 between nucleons, which occurs for collisions of 95 MeV/u ^{12}C in PMMA target. For indicative reasons we mention that the GEANT4 QMD default clustering maximum distance is 4 fm (corresponding to a $R^2 = 16\,\mathrm{fm}^2$). Figure 11.6 shows the impact of different R values on the number of emitted gammas in $50 \times 50 \times 50\,\mathrm{mm}^3$ PMMA target irradiated with a 95 MeV/u ^{12}C beam.

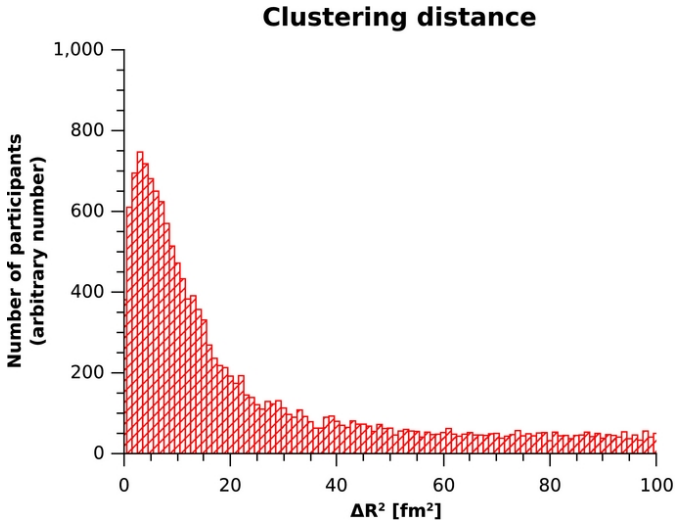

Figure 11.5: Distribution of the square distance between nucleons (equation 11.2), used for the assembly of excited fragment at the end of the dynamic part of the collision (QMD). By default nucleons at a distance lower than $R^2 = 16\,\mathrm{fm}^2$ are considered to belong to the same cluster/fragment.

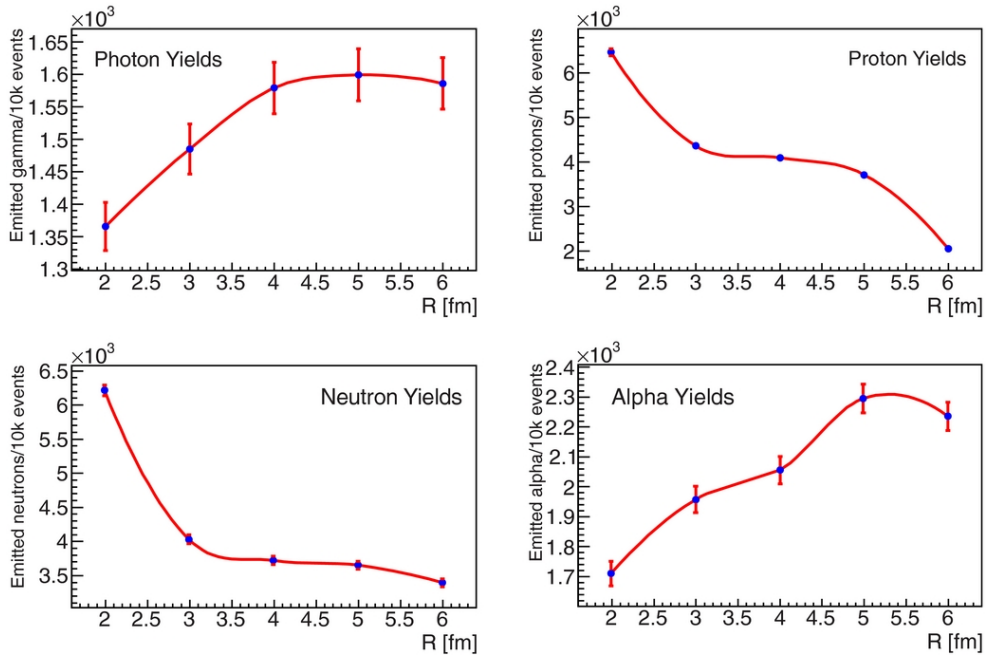

Figure 11.6: Emitted particles per 10^4 ^{12}C ions of 95 MeV/u impinging on a $50\times50\times50\,\mathrm{mm}^3$ PMMA target as a function of the maximum clustering distance R. The error bars denote statistical error of one standard deviation. The absence of error bars signifies that one standard deviation would lead to bars smaller than the size of the point markers in the graph.

Similarly to what has been shown for L, figure 11.7 shows that for smaller values of R the QMD collision products are lighter and less excited fragments, resulting in fewer gamma emissions during de–excitation.

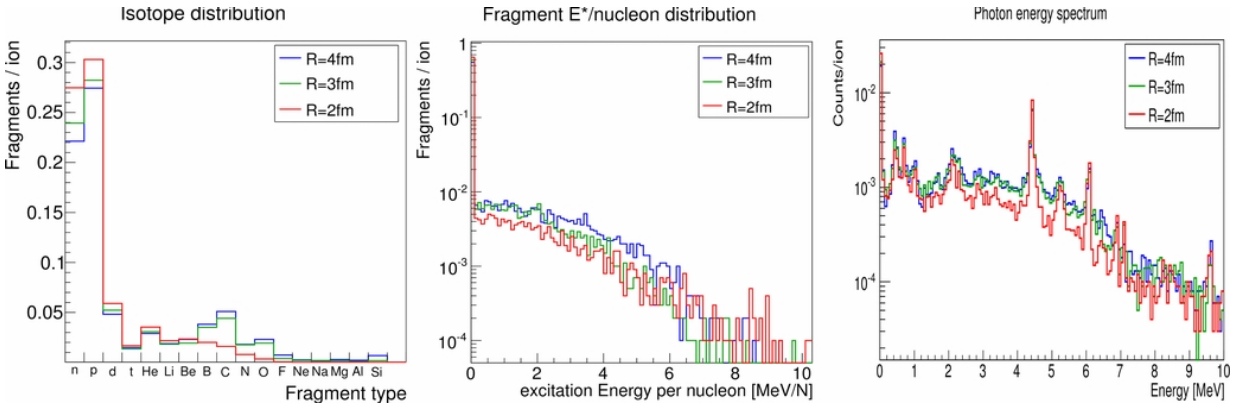

Figure 11.7: As in figure 11.4, but for different values of R and for $L = 2\,\mathrm{fm}^2$.

Both R and L variations affect prompt–gamma emission yields, without changing the characteristic emission lines of the emitted spectra. Nevertheless, the impact of L on the prompt–gamma emission is more significant than that of R, as one can see in figures 11.3 and 11.6. Furthermore, the parameter L influences not only the dynamics of the reaction but also the creation of the projectile and target systems, which can be benchmarked using nuclear properties.

11.3.3 Nuclear properties

The L parameter values for which one achieves a large reduction of emitted photons is for values of $L \leq 1\,\mathrm{fm}^2$ or $L \geq 3\,\mathrm{fm}^2$. As there are two possible parameter ranges, that differ by such a large factor, the selection of L has to be constrained by additional properties of the collision system. In that way we will be able to define an L–value range that will reduce gamma emission significantly, but at the same time yield a reasonable prediction of basic nuclear properties. As such properties we considered the binding energy per nucleon and the nucleon/charge density of the nucleus, which are independent of the subsequent de–excitation phase.

In figure 11.8 we present the deviation of the GEANT4 QMD calculated binding energy per nucleon for five different nucleus species, from the values obtained from the National Nuclear Data Center (NUDAT2) of the Brookhaven National Laboratory. We selected ^{12}C,^{16}O and ^{40}Ca nuclei based on their abundance in the human body, as well as in the targets used in our experiments. In addition, we extended our study to ^{4}He, as an example of very light nuclei, but also included ^{127}I as a token of much heavier nuclei. Such heavy elements, though less abundant in a patient, have often been used for the selection of the default value of the wavepacket width. The L range studied was from 0.5 to $3\,\mathrm{fm}^2$. For $L \leq 2\,\mathrm{fm}^2$, for all studied nuclei, the binding energy per nucleon calculated by QMD

is consistent with the NUDAT2 values within a fraction of a per cent, dominated by statistical fluctuations. Therefore a selection of a value lower than the default will decrease the emission of prompt–gammas (as shown in section 11.3.2), while at the same time it will reproduce the expected binding energies accurately. On the contrary, for values of $L \geq 2\,\mathrm{fm}^2$, the calculated binding energy per nucleon diverges up to 70%. For the heaviest of the studied nuclei (^{40}Ca and ^{127}I), deviations remain small throughout the whole parameter range, while the lighter the nucleus the more significant they become. Therefore, for a typical hadrontherapy application, where the majority of the nuclei will be light, the preferable selection for a new value of L would be lower than the default of $2\,\mathrm{fm}^2$, achieving lower prompt–gamma emission and better binding energy calculation for a wide range of nuclei.

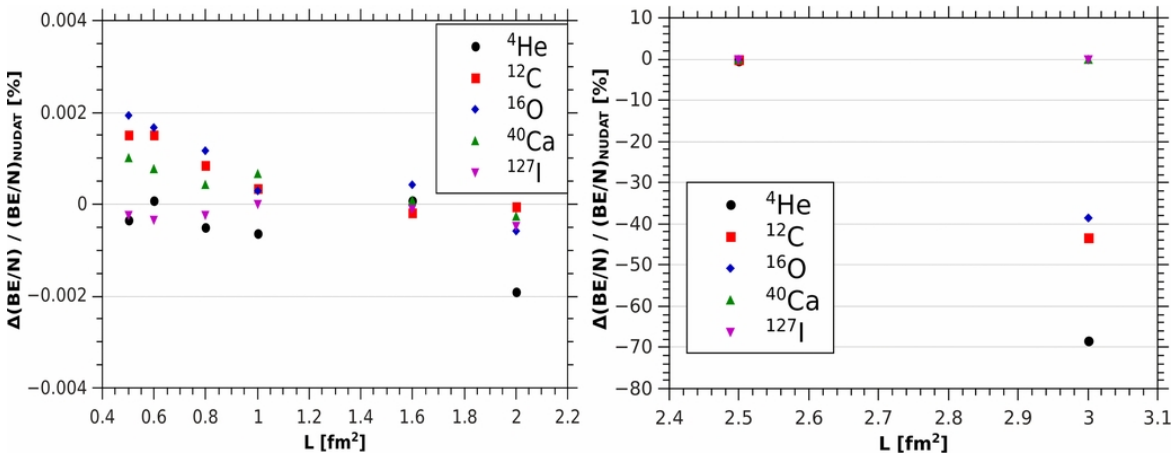

Figure 11.8: Difference in the binding energy per nucleon (BE/N) between GEANT4 QMD calculations and NUDAT2 values for ^{4}He, ^{12}C, ^{16}O, ^{40}Ca and ^{127}I ground state nuclei.

Subsequently we study the dependence of the nuclear density on the parameter L. In figure 11.9, the nuclear density of ^{12}C nuclei as a function of the radial distance from the center of the nucleus, calculated for different values of the parameter L (ranging from 0.8 to $4\,\mathrm{fm}^2$) is shown. The nuclear density is calculated as the number of nucleons within a radius $r + \Delta r$ divided by the spherical shell volume $V(r + \Delta r) - V(r)$ and compared with data and theoretical calculations found in Gasques et al. (2002), where their experimental data have been analyzed using the assumption of either Fermi or harmonic oscillator shapes of the ^{12}C nucleon density. Their theoretical calculations are based either on Dirac–Hartree–Bogoliubov (DHB) or Fermi (2pF) distribution. The results show general agreement of QMD with the data taken from Gasques et al. (2002) for the regions close to the center of the nucleus ($r \leq 2\,\mathrm{fm}$), while it seems that QMD fails to reproduce the nuclear halo (tail for $r \geq 2\,\mathrm{fm}$). In specific, for $r \leq 2\,\mathrm{fm}$, $0.8\,\mathrm{fm}^2 \geq L \leq 2\,\mathrm{fm}^2$ values seem to yield results close to the available data points and the theoretical DHB calculation, which describes the central nuclear region data better than the Fermi distribution. For the same region, $L \geq 2\,\mathrm{fm}^2$ for which also the simulated binding energies deviate from the experimental values, nuclear density is still comparable to that of data and DHB calculation. For $r \geq 2\,\mathrm{fm}$, where QMD

deviates both from the theoretical calculations and the experimental data for the whole L parameter range, low L values ($L \leq 2\,\mathrm{fm}^2$) seem to affect the shape of the tail towards the right direction (extending nuclear halo), without achieving though a very good agreement with data.

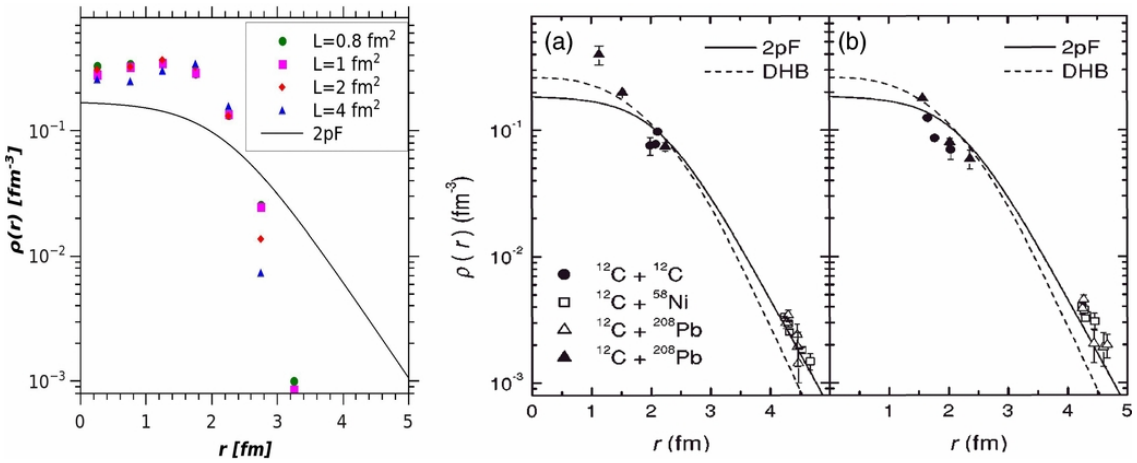

Figure 11.9: ^{12}C nucleon density. Left, GEANT4 calculations for different values of L. For indicative reasons, the theoretical Fermi distribution (2pF) is also shown. Right, experimental nucleon density values (Gasques et al. 2002) obtained by using in the analysis the assumption of (a) Fermi or (b) harmonic oscillator shapes of the ^{12}C nucleon density. In the right figure, the points represent the experimental values and the lines the theoretical Dirac–Hartree–Bogoliubov (DHB) calculation or the Fermi distribution (2pF).

11.3.4 Charged particle emission

Although the main goal of this work is the improved description of prompt–gamma emission, a reasonable prediction of charged particle emission has to be maintained. In order to test the proposed optimization of the QMD model in terms of charged particle emission, we compared GEANT4 predictions with the data taken from Braunn (2010) and Braunn et al. (2011). Contrary to the approach for the simulation of the prompt–gamma experimental setups, we have applied an ideal detection approach for the simulation of this experiment. Nevertheless, the absence of a realistic detector simulation is expected to have a very small impact on the presented results. This was verified in (Braunn 2010), pages 84–85, where a comparison between the results obtained with a full detector simulation and with an ideal detection approach is presented. A good agreement was found between the two approaches for detection angles above 10°. Therefore the detection efficiency could be considered close to 100%, permitting the usage of an ideal detection approach. This is not valid for detection angles lower than 10° and especially for charged fragments with $Z = 1$ and $Z = 2$, where differences up to 35% and 15% were respectively reported. Finally, in addition to the ideal detection based only on geometrical acceptance, realistic particle detection energy thresholds as applied to the measurements were used (Braunn et al. 2011).

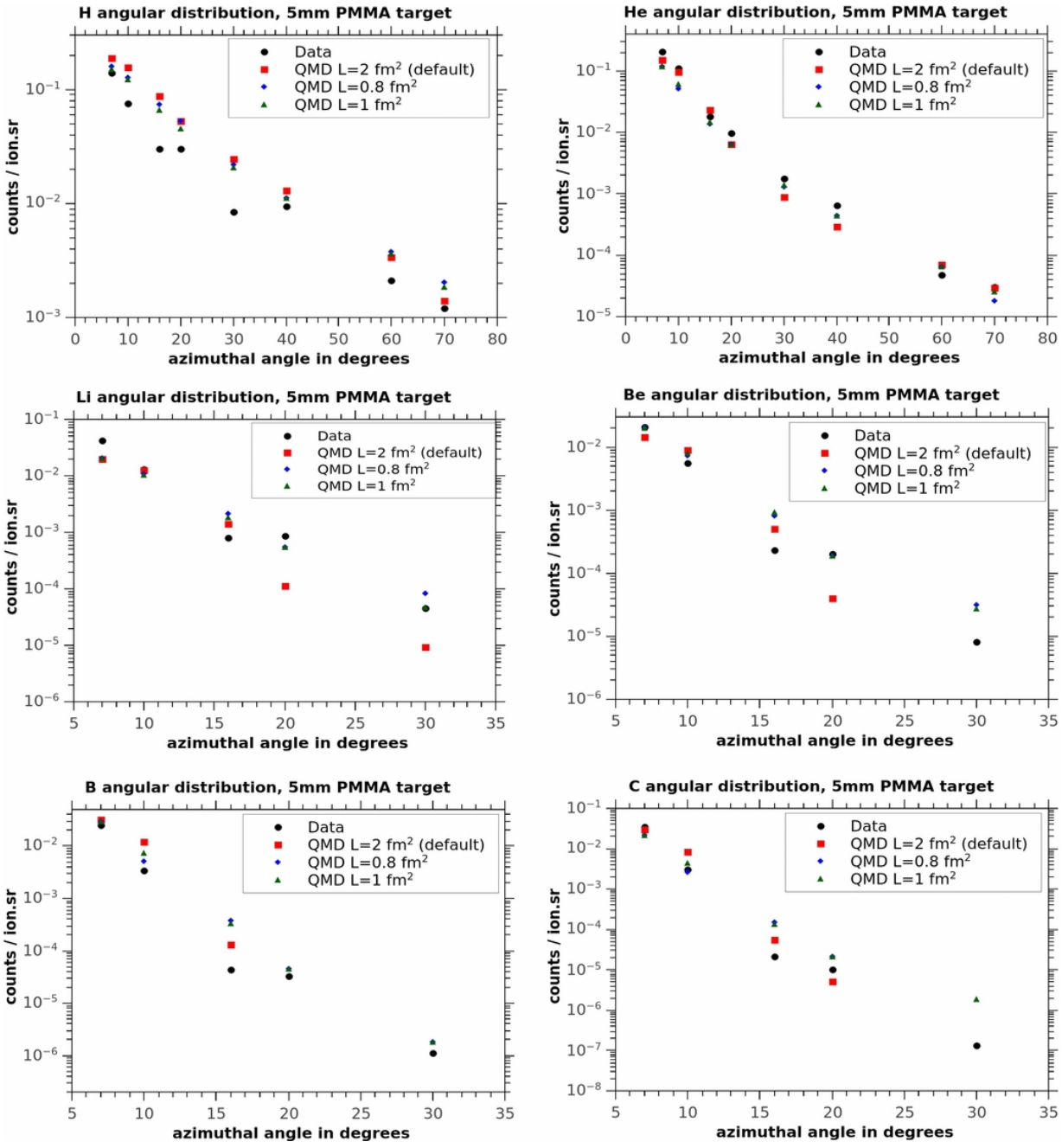

Figure 11.10: Comparison of angular distributions of charged particles with GEANT4 predictions. Charged fragments of $Z = 1$ to 6 (H to C), produced by collisions of 95 MeV/u ^{12}C on a 5 mm thick PMMA target. Data taken from Braunn (2010).

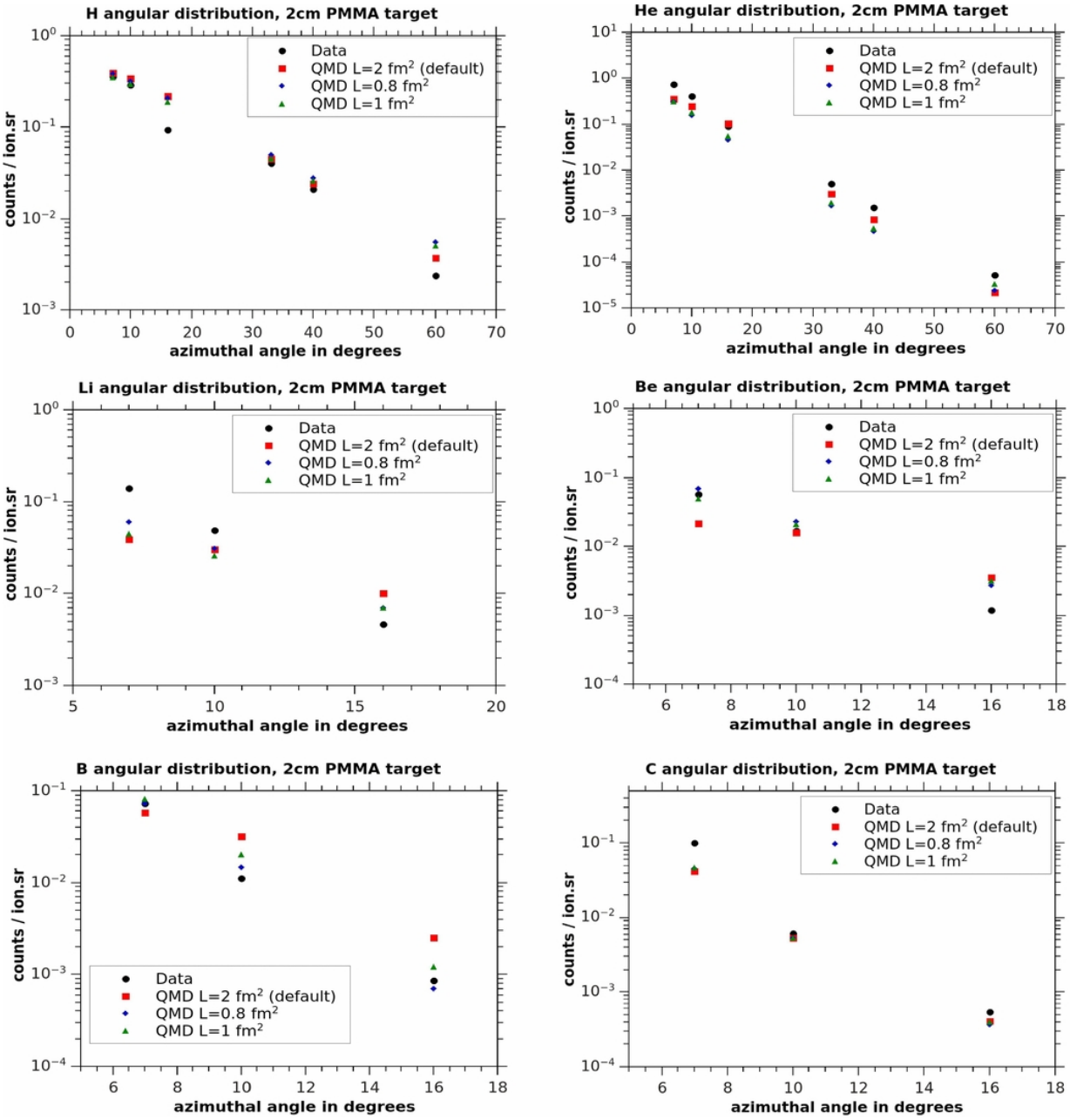

Figure 11.11: Comparison of angular distributions of charged particles with GEANT4 predictions. Charged fragments of $Z = 1$ to 6 (H to C), produced by collisions of 95 MeV/u ^{12}C on a 2 cm thick PMMA target. Data taken from Braunn (2010).

This experiment has been performed at GANIL, using a 95 MeV/u ^{12}C beam and PMMA targets with thicknesses varying from 5 to 40 mm. In figures 11.10 and 11.11 we present experimental data for the angular distributions of charged fragments with $Z = 1$ to $Z = 6$ in counts per incident ion per steradian (count/ion/sr). Further information with the corresponding numbers of figures 11.10 and 11.11 can be found in the tables D.2 and D.3

in appendix D.0.2. For the 5 and 20 mm thick targets, we compare the measurements with the GEANT4 results for three values of L: default value 2, 1 and 0.8 fm^2 (two values in the parameter space constrained by the studies presented in the previous sections). The results obtained with $L = 1\,\text{fm}^2$ and $L = 0.8\,\text{fm}^2$ are in all cases comparable with the predictions of the default QMD ($L = 2\,\text{fm}^2$). Furthermore, using values of L lower than 2 fm^2 improves the predictions of GEANT4 in comparison to the experimental data, in the majority of the angles and fragment charges for both target thicknesses tested. The only occasion where the optimized QMD provides consistently a larger discrepancy with the experimental data than the default QMD ($L = 2\,\text{fm}^2$), is for $Z = 2$ (He isotopes) and especially in the case of the 20 mm thick target.

11.3.5 Prompt–gamma yields – Optimized QMD

In figure 11.12 we compare the predictions of GEANT4 to the measurements obtained from the experiments described in section 11.2.1. In all three cases presented (upper left: 80 MeV/u ^{12}C beam at PMMA – prompt–gamma energy spectrum, upper right: 95 MeV/u ^{12}C beam at PMMA – prompt–gamma depth profile, bottom: 310 MeV/u ^{12}C beam at PMMA – prompt–gamma depth profile), experimental data are better described when using $L \leq 1\,\text{fm}^2$.

For the energy spectrum, the initial (QMD $L = 2\,\text{fm}^2$) overestimation of the integral by 100% improves to about 30% for $L = 1\,\text{fm}^2$ (or –2% for $L = 0.8\,\text{fm}^2$). For the 95 MeV/u ^{12}C profile, the initial discrepancy of 86% at the entrance of the target and of 165% at the maximum of the profile decreases to 37% at the entrance and 70% at the maximum of the profile for $L = 1\,\text{fm}^2$, or –4% at the entrance and 23% at the maximum of the profile for $L = 0.8\,\text{fm}^2$. Finally, for the 310 MeV/u ^{12}C profile, the initial overestimation was 115% at the entrance of the target and 178% at the maximum of the profile. This discrepancy has changed to –4% at the entrance and 72% at the maximum for $L = 0.8\,\text{fm}^2$.

In both 95 MeV/u ^{12}C and 310 MeV/u ^{12}C profiles, some discrepancy, although much smaller than the initial one, is still observed even when using the proposed values of L. Two important characteristics of this difference between measurements and experiments are that it is more pronounced close to the end of the ion range, and furthermore that it increases at higher beam energy (23% for 95 MeV/u ^{12}C and 72% for 310 MeV/u ^{12}C in the case of $L = 0.8\,\text{fm}^2$). Those features can be explained by figures 11.13 and 11.14, where we present the simulated depth profile of emitted gamma–rays in the target as well as the distinction between the gamma–rays emitted by ion induced nuclear reactions (QMD) and p/n induced reactions (BIC). As there is an increasing number of p/n induced reactions along the ion path, the effect of the QMD optimization gradually becomes less important at larger depths in the target, where the description of the dynamic part of nuclear reaction is handled more often by the BIC model. In the extreme case of 310 MeV/u ^{12}C, in terms of prompt–gamma emission, QMD does not even play the major role beyond the first few cm of the ion path. At this energy, as expected, we also observe the largest discrepancy between experimental and simulated profiles. Despite that, even at 310 MeV/u ^{12}C, for the first few cm of the ion path where QMD plays the major role as shown in figure 11.14, we

observe that the proposed lower values of L reproduce accurately the experimental point (figure 11.12 bottom, point at 20 cm).

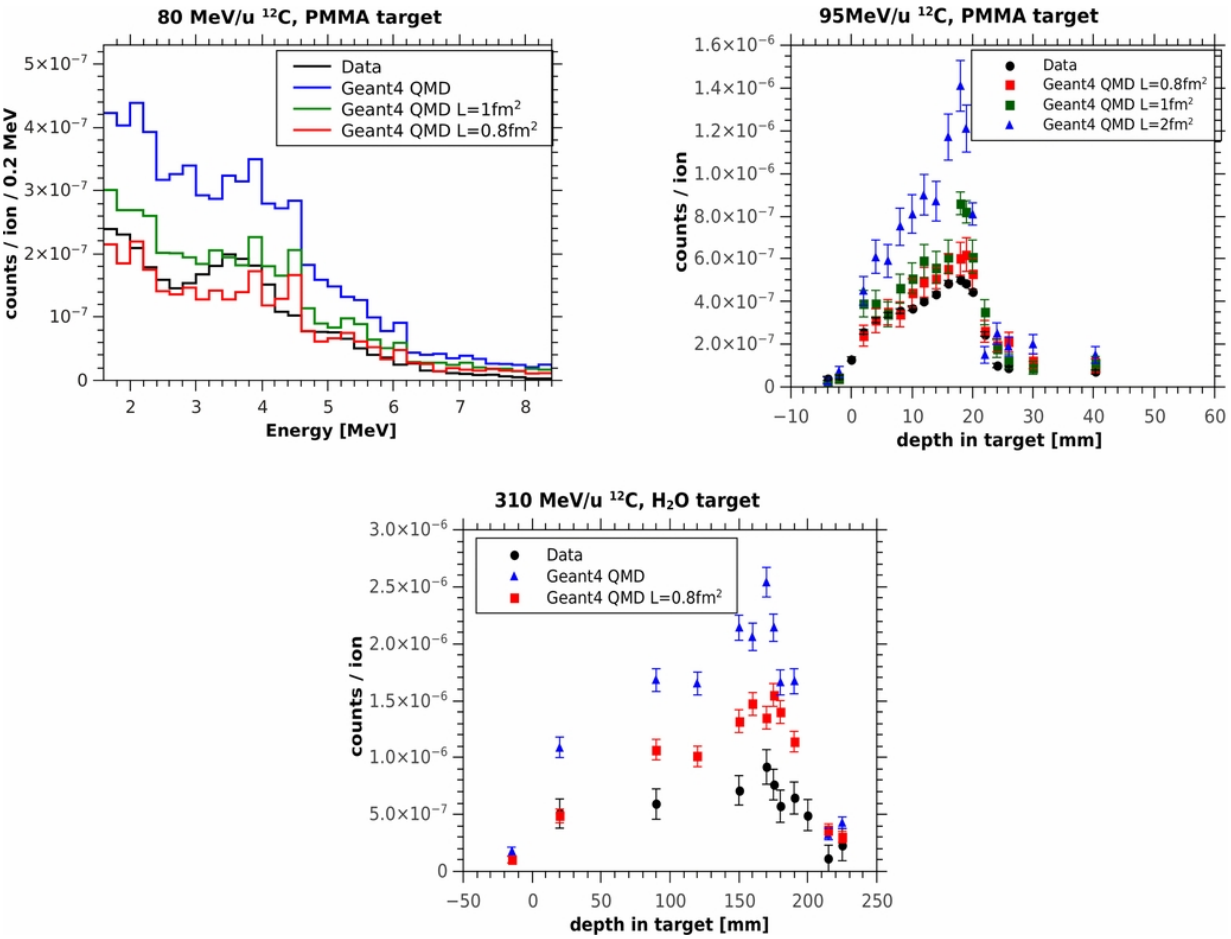

Figure 11.12: GEANT4 comparison with the experiments described in section 11.2.1. Upper left, the energy spectrum obtained in Catania (Agodi et al. 2012) by irradiating a PMMA target with a 80 MeV/u ^{12}C beam. Upper right, the prompt-gamma depth profile obtained at GANIL by the irradiation of a PMMA target by a 95 MeV/u ^{12}C beam. Bottom, the depth profile obtained from a water target with a 310 MeV/u ^{12}C beam. The error bars denote statistical error of one standard deviation.

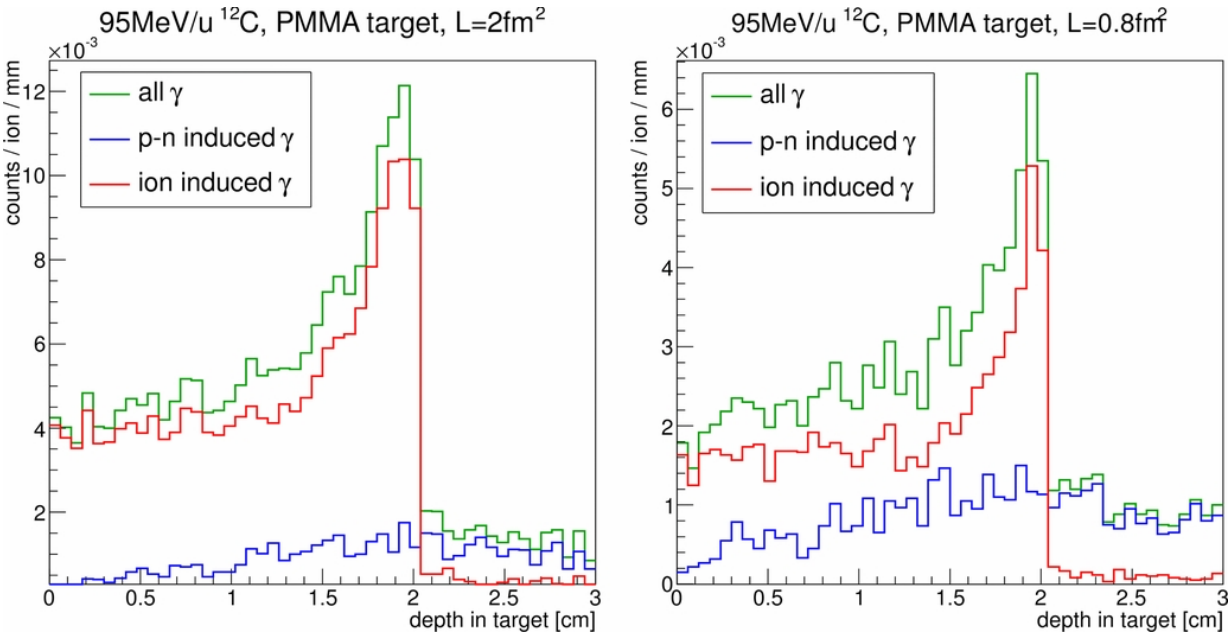

Figure 11.13: 95 MeV/u ^{12}C simulated prompt–gamma emission in a PMMA target, for $L = 2\,\mathrm{fm}^2$ (left) and $L = 0.8\,\mathrm{fm}^2$ (right). The total emission profile as well as the contributions from ion induced (QMD) and p/n induced (BIC) reactions are shown.

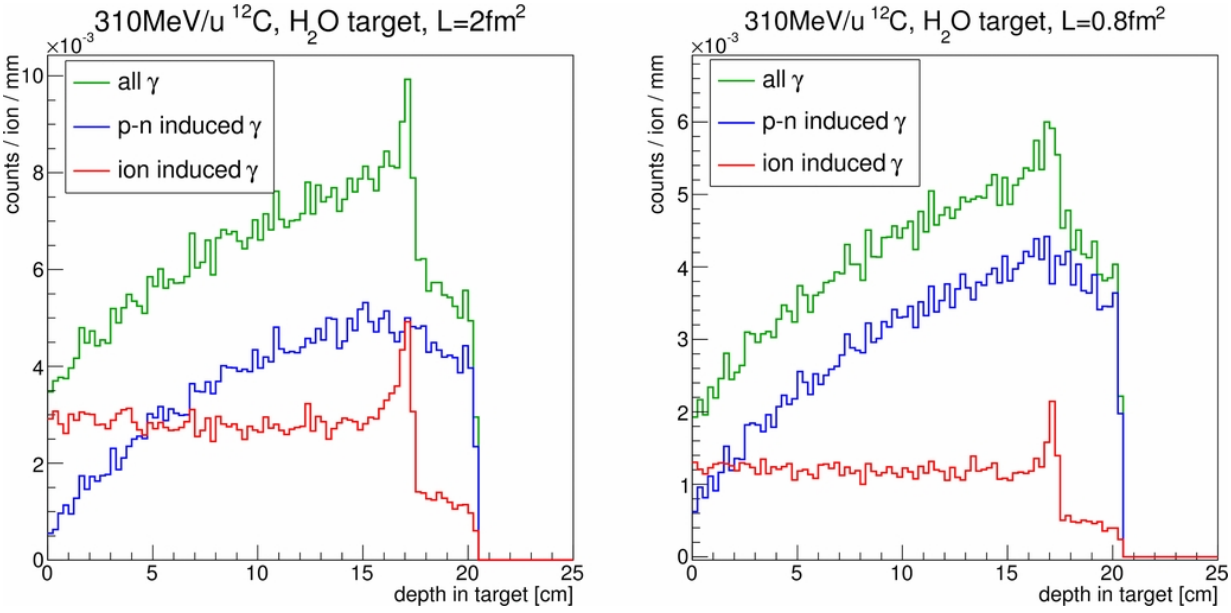

Figure 11.14: 310 MeV/u ^{12}C simulated prompt–gamma emission in a PMMA target, for $L = 2\,\mathrm{fm}^2$ (left) and $L = 0.8\,\mathrm{fm}^2$ (right). The total emission profile as well as the contributions from ion induced (QMD) and p/n induced (BIC) reactions are shown.

11.4 Conclusions

In the presented study, we have investigated the performance of some of the most widely used internal nuclear models of Geant4 in the field of hadrontherapy. Geant4 simulations are shown to overestimate the emission of prompt–gammas by a factor between 2 and 3 in the case of ^{12}C and to a lesser extent for proton beams, when using the quantum molecular dynamics (QMD) or binary light ion (BLI) reaction model for ion and the binary cascade model for proton and neutron inelastic reactions.

Furthermore, we have focused on the QMD model in order to improve the description of ion–ion nuclear reaction and the subsequent emission of prompt–gammas. The QMD calculation of ion–ion reaction dynamics involves a number of parameters, either technical or physical. As the model was never optimized for Medical Physics applications, that means for mainly light nuclei interactions, we used the nucleon Gaussian wave packet width, which is a free physical parameter in QMD, in order to better describe light ion systems and their collisions.

To define a more appropriate value for this parameter, we have benchmarked QMD with the nuclear properties of binding energy per nucleon and nuclear density. In addition, we have checked that an improved description of those properties, which yields better agreement with experimental prompt–gamma data, also maintains the same quality or improves the description of charged fragment emission.

Finally, there are remaining discrepancies between prompt–gamma depth profiles and the corresponding simulated ones, which become more important for higher beam energies and closer to the end of the primary ions' path. Those discrepancies are attributed to the contribution of secondary protons or neutrons induced nuclear reactions and are not handled by the QMD model. Therefore, a similar study/optimization for proton/neutron nuclear models has to be performed. This is also indicated by the comparison between experimental data from clinical proton beams and Geant4 simulations shown in the current study (see figure 11.2 – prompt–gamma spectrum from 160 MeV protons).

Acknowledgments

This research project has been supported by the Regional Program for Research in Hadrontherapy (PRRH, under CPER 2007–13 funding) and the ENVISION European project (grant agreement no241851). It was performed in the frame of the Labex PRIMES (ANR–11–LABX–0063) of Université de Lyon.

Chapter 12

Prompt–gamma imaging sensitivity to proton range variations

This chapter contains the work published in Physics for Medicine and Biology, in volume 60, issue 24, pages 9329–47, in December 2015, with the title *Monte Carlo study on the sensitivity of prompt–gamma imaging to proton range variations due to interfractional changes in prostate cancer patients*, by Schmid et al. (2015) – (senior author: Dedes)

12.1 Introduction

Protons are increasingly used in external beam radiotherapy due to their advantageous physical dose distribution. Recent advancements in active scanning of proton pencil beams have allowed the clinical implementation of intensity modulated proton therapy (IMPT). IMPT allows higher target conformity than broad proton beams based on passive scattering. The higher dose conformity of protons comes at the cost of increased sensitivity to anatomical variations and patient positioning errors. Therefore, to fully exploit the advantages of IMPT, an accurate and precise real–time monitoring of the spot–by–spot dose delivery, or the proton range as surrogate, is desired. To do so it may be possible to exploit secondary radiation generated at the time of proton irradiation without incurring additional imaging dose to the patient.

Different methods have been proposed for in–vivo ion range monitoring exploiting secondary radiation. Currently, the only approach being evaluated in different clinical centers (Massachusetts General Hospital–Boston USA, National Cancer Center Hospital East–Kashiwa Japan, Heidelberg Ion Beam Therapy Center–Heidelberg Germany, Gunma University Heavy Ion Medical Center–Gunma Japan) is based on positron emission tomography (PET) (Parodi et al. 2001, Enghardt et al. 2004, Knopf et al. 2008, Hsi et al. 2009, Nishio et al. 2010, Zhu et al. 2011, Bauer et al. 2013, Knopf & Lomax 2013). The method is based on the detection of coincident photons stemming from the decay of β^+–emitters. The latter are produced mainly by nuclear reactions induced by the proton

beam impinging on target nuclei in tissue. The method has drawbacks: PET scanners are costly and are challenging to install in the treatment room. Transportation from the treatment room to a dedicated PET scanner degrades signal to noise ratio due to loss of activity from physical and biological decay. Additionally, due to the nature of interaction cross sections, the β^+–emitter production fall–off is located several millimeters upstream of the Bragg peak.

Recently, alternative methods based on the detection of prompt photons emitted by nuclei undergoing de-excitation following nuclear interactions have been proposed (Stichelbaut F 2003, Min et al. 2006). These photons are generally emitted within nanoseconds or less from the time of interaction, potentially allowing for online monitoring. The interaction cross sections for the generation of excited nuclei emitting prompt–gamma are peaked at lower energies than those for the generation of β^+–emitters, suggesting a prompt–gamma emission fall–off located closer to the Bragg peak (Moteabbed et al. 2011). Different techniques have been investigated in order to exploit the physical correlation of the prompt–gammas with the range of the protons. Studies of the spatial distribution of the prompt–gamma counts, which displays a fall–off of similar shape and with a depth correlated to that of the Bragg peak, have been performed by several investigators (Kim et al. 2019, Frandes et al. 2010, Kormoll et al. 2011, Min et al. 2011, Richard et al. 2011, Smeets et al. 2012, Janssen et al. 2014, Pinto et al. 2014, Priegnitz et al. 2015). The exploitation of the additional prompt–gamma energy information has also been considered (Polf et al. 2009, Polf et al. 2013, Verburg & Seco 2014). More recently, the use of temporal distributions of prompt–gammas was also proposed as a tool to deduce proton range (Golnik et al. 2014, Hueso-González et al. 2015).

Several phantom studies based on the methods mentioned above have investigated the correlation of prompt–gammas to the proton range either with experiments or simulations, generally employing homogeneous or slab geometries. A first patient study reported simulations for passively scattered broad proton beams and few pencil beams for head and neck, prostate and thoracic spine tumours, but did not systematically investigate the method for IMPT (Moteabbed et al. 2011). In that work, differences between the proton range defined as the 50% fall–off of the dose amplitude and the prompt–gamma fall–off (also defined as the location of the 50% of the prompt–gamma signal amplitude) were analysed. Gueth et al. (2013) investigated the use of machine learning methods to detect range shifts for a single prostate cancer patient using IMPT.

In this work, we systematically investigate the sensitivity of prompt–gamma signals to detect spot–by–spot variations in IMPT dose deposition caused by daily anatomical variations. We aim at determining the accuracy and precision of prompt–gamma based range monitoring in real anatomical scenarios. Furthermore we want to assess whether tissue heterogeneities parallel to the beam direction can significantly alter the shape of prompt–gamma distributions and thus weaken their correlation to the proton range deduced by the deposited dose depth profile. To this end, five clinical datasets containing three computed tomography (CT) scans of prostate cancer patients acquired at different time points were used to generate and evaluate IMPT treatment plans in terms of proton range and prompt–gamma emission profiles in a Monte Carlo simulation study. We in-

vestigated two methods to identify the prompt–gamma fall–off based on shifting against a reference profile (Knopf et al. 2009, Helmbrecht et al. 2012, Frey et al. 2014) and by fitting a fall–off function (Janssen et al. 2014).

12.2 Materials and methods

12.2.1 Patient data

For this study planning CT scans of 5 prostate cancer patients undergoing intensity modulated radiotherapy (IMRT) were used. For each patient a set of 3 planning CT scans, labelled CT_1, CT_2 and CT_3, were acquired on consecutive days. Changes in the rectal filling and femoral heads position were the main anatomical changes between the different CT scans for each patient. The planning CT scanner was a Toshiba Aquilion LB (Toshiba Medical Systems, the Netherlands) and images were reconstructed on a $1.074\,\text{mm} \times 1.074\,\text{mm} \times 3\,\text{mm}$ grid. The three images are used clinically to estimate variations in patient anatomy over the course of treatment at the contouring and planning stage. Contouring of the clinical target volume CTV, bladder and rectum was performed individually on each scan, following rigid registration to CT_1. Treatment planning was performed on CT_1 using, in a conservative approach, the union of the three contour sets for the CTV, bladder and rectum. The photon therapy PTV was obtained by a 1 cm expansion of the union CTV and was used in this work as well.

12.2.2 Treatment planning

Intensity modulated proton therapy treatment plans were generated using a MATLAB (Mathworks, Natick, USA)–based research treatment planning system (TPS) based on CERR (Deasy et al. 2003), featuring a proton pencil beam algorithm (Schell & Wilkens 2010). The plans consisted of 2 parallel opposed scanned proton beams each delivering a uniform dose to achieve a prescription of 3 Gy effective dose (2.73 Gy physical dose) to the PTV per fraction. Dose was calculated on the same voxel grid as the CT. The CT number to relative stopping power to water (RSP) lookup table required by the pencil beam algorithm was derived using the approach of Schneider et al. (1996) adapted to the CT scanner used in this study.

The spots from one beam were exported from the TPS to a Monte Carlo dose calculation engine based on Geant4 (Agostinelli et al. 2003). Only the highest intensity spots were taken into account since they are most distal to the patient surface and deposit the largest fraction of the dose. This study assumes spot-by-spot verification of the proton range and the distal layer is the most interesting for in vivo monitoring since it contains the highest intensity spots, which should yield the best signal to noise ratio. Additionally these spots fully cross the treated anatomy. A typical spot delivered about 10^8 protons to the patient.

12.2.3 Monte Carlo simulation

Monte Carlo simulations were performed with GEANT4 version 10.0, patch–01. The developed application was based on an extension of the DICOM example. All primary and secondary particles (protons, electrons, photons, neutrons, positrons, ions and nuclear fragments) were transported and their interactions were simulated according to theoretical models or tabulated cross sections. A range cut of 1 mm was used as secondary production threshold, comparable to the CT grid size. We used the G4RegularNavigation algorithm to navigate in regular voxelized geometries, which allows for increased simulation speed with low memory consumption.

The physics used to simulate interactions of all primaries and secondaries was the predefined `QGSP_BIC_HP` GEANT4 physics list. This includes multiple scattering and ionization for charged particles, as described in the standard GEANT4 electromagnetic package, which offers a good compromise between performance and speed. For the considered energies, the binary cascade model (BIC) was used for ion–ion and hadron–ion inelastic interactions.

A lookup table for CT number to mass density and tissue composition was generated for the Monte Carlo simulation based on the work of Schneider et al. (2000). The lookup table was adapted to the CT scanner used in this study.

Dose and prompt–gamma scoring

Separate Monte Carlo simulations were performed for each distal spot. The dose of the selected spot was scored on the CT grid. Photons with energy higher than 2.5 MeV were counted in two different ways: (1) they were scored within the voxel corresponding to their emission point (these distributions are noted with PG_i where i corresponds to the CT scan number) and (2) they were detected outside the patient at a scoring plane of dimensions 540 mm × 120 mm. The lateral position of the scoring plane was centred on the beam central axis and positioned above the patient surface (170 mm above the isocenter). The location of the scoring plane is illustrated in figure 12.1. The prompt–gamma photons reaching the scoring plane were stored in a phase space distribution PS_i (i corresponds to the CT scan number) by saving their detection position, energy and momentum at the scoring plane, as well as their original emission position within the patient. For analysis, prompt–gamma photons within different angular acceptance windows were investigated ($\pm 2^o$,$\pm 3^o$ and $\pm 4^o$ normal to the scoring plane). To generate profiles, for both PG_i and PS_i, lateral integration normal to the incident beam direction was performed. For PG_i, the lateral integration was over non–zero contributions in the yz plane of figure 12.1. For PS_i the lateral integration was along z in the scoring plane of figure 12.1. Along the proton beam direction (x in figure 12.1) prompt–gamma counts were binned to the voxel grid (1.074 mm spacing).

Simulation of the full planned proton fluences (approximately 10^8 protons per spot), which is necessary to obtain realistic statistics for PS_i, is time consuming (approximately 860 h per spot on a single CPU). For this reason we decided to perform analysis on PG_i using a fraction of the full planned proton fluences corresponding to the geometrical detection

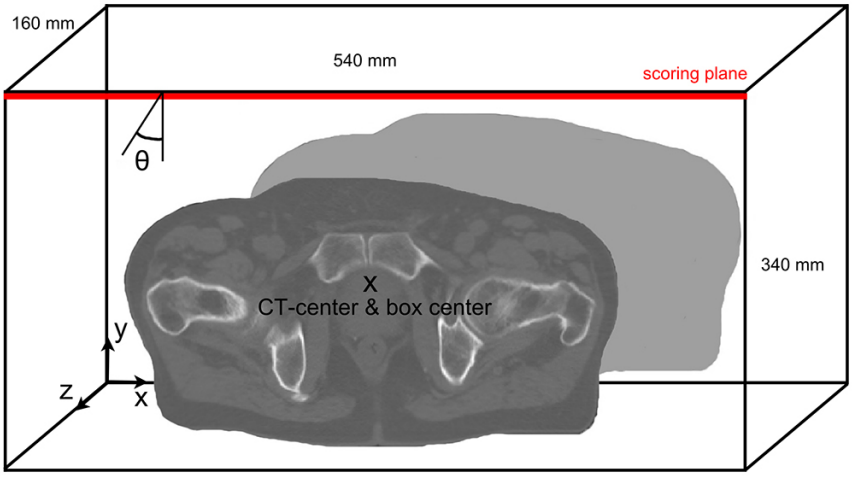

Figure 12.1: The anterior side of the box (parallel to zx) was used as scoring plane. The front edge of the plane is indicated in the figure. The angle indicated in the figure is calculated in the xy plane. The isocenter is labelled CT–center.

efficiency. The purpose is to have in the PG_i obtained with the scaled proton fluence, the same number of prompt–gammas as in the PS_i obtained from the full proton fluence. To validate this approach, 36 out of the selected 180 distal spots of one patient were simulated on CT_1 and CT_2 with full proton fluences to serve as baseline. The number of spots was limited by calculation time and the spots were chosen to sample the distribution of range shifts observed in patients. For validation, PS_1 and PS_2 of the 36 spots were compared, as well as the corresponding PG_1 and PG_2.

The highest intensity layer of the optimized plan of patient 1 used for PG calculation contained 7.8×10^7 protons corresponding to a 0.35% rescaling of the full plan fluence. This scaling factor yields PG profiles with the same number of prompt–gammas as in the PS using $\pm 3^{o}$ angular acceptance for a full plan.

12.2.4 Data analysis

In this work, the spot–by–spot proton range was estimated using the prompt–gamma emission profiles as surrogate and labelled R_{PG} . The data analysis was performed for PS_i in the same manner as for PG_i, hence all following considerations explicitly reported for PG_i are also representative for PS_i. R_{PG} was correlated with the proton range estimated directly from the dose distribution R_{dose}. The range differences ΔR_{dose} and ΔR_{PG} (ΔR_{PS}) between two fractions, corresponding to two CT scans, were analysed. To find ΔR_{PG} two methods were investigated: (1) fitting a sigmoid function to the prompt–gamma profiles PG_i in the region where the emission decreases (fall–off) and rises (patient's surface) and (2) shifting the prompt–gamma profiles $PG_{2/3}$ against a reference profile PG_1 until the root mean square (RMS) error in a region of interest is minimized.

Region of interest identification

For both methods, it is necessary to define two regions of interest covering the prompt–gamma profile fall and rise. To find the fall–off region, it is not sufficient to identify the profile maximum as this point is not necessarily followed by the fall–off as seen in figure 12.2, where the prompt–gamma production is higher in bone at 200 mm depth. To define the fall–off region an integration of the prompt–gamma profile in anti-beam direction is performed: $I_{\mathrm{PG}}(x) = \sum_{i=1}^{x} \mathrm{PG}\,(x_{\mathrm{PGend}} - i)$ where x_{PGend} is the last bin of the profile, with x indicating the bins along the profile. The intersection point X_{intersec} of linear fits to the background (shallow rise: $I_{\mathrm{PG}}(x) < 4\%$) and main profile (steep rise: $10\% < I_{\mathrm{PG}}(x) < 40\%$) defines a window [$X_{\mathrm{intersec}}$ – 40 bins, X_{intersec} + 80 bins]. Here bins refer to the width of the CT voxel in the beam direction. In this window the most distal point p out of the 5 points with the highest values is identified. Subsequently the fall–off region is defined as [p – 10 bins, p + 50 bins] when used for fitting profiles and as [p, p + 60 bins] when used for shifting profiles. The latter is illustrated in figure 12.2(B).

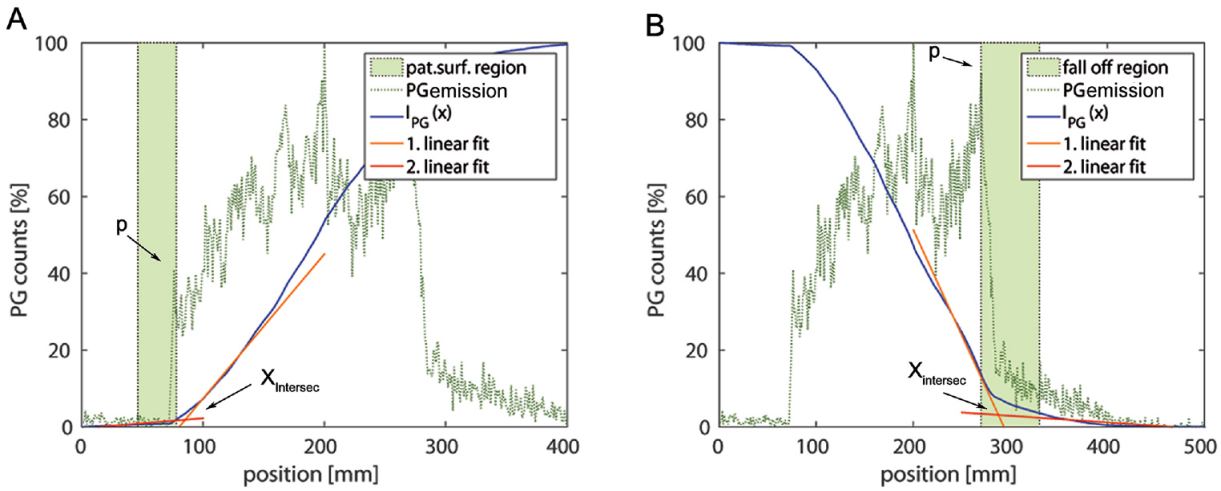

Figure 12.2: The x axis describes the position in the underlying CT scan in the beam direction. The blue curve IPG represents the integrated prompt–gamma emission profile PG in–beam (A) and anti–beam direction (B). The position where the red and orange lines intersect is called X_{intersec}. In a window around X_{intersec} the most proximal/distal out of the highest 5 points is called point p and is used to define (A) the patient's surface and (B) the fall–off region. The prompt–gamma profiles as well as the integrated profiles are normalized to their maximum value for visualization purposes. Here regions used for the shifting approach are displayed.

To identify the patient surface region from the prompt–gamma profile, the integration is performed in the in–beam direction, thus $I_{\mathrm{PG}}(x) = \sum_{i=1}^{x} \mathrm{PG}(i)$. The thresholds for fitting the shallow and steep rise were $I_{\mathrm{PG}}(x) < 2\%$ and $5\% < I_{\mathrm{PG}}(x) < 20\%$ respectively. Within the window [X_{intersec} – 20 bins, X_{intersec} + 10 bins] the most proximal out of the 5 points with highest prompt–gamma profile is used to define the point p. Subsequently

the patient surface region is defined as [p – 50 bins, p + 10 bins] when used for fitting profiles and as [p – 30 bins, p] when used for shifting profiles. The latter is illustrated in figure 12.2(A).

When using the fitting method to find changes in prompt–gamma deduced ranges ΔR_{PG}, the prompt–gamma range R_{PG} is identified for each profile separately. This is done by applying a sigmoidal fit at both the patient surface [p – 50 bins, p + 10 bins] and fall–off regions [p – 10 bins, p + 50 bins]. The fitted function $a + (1 - a)\mathrm{erf}\,[b(x - c)]$ contains three parameters, where parameter c refers to the inflection point of the error function. This point c, as estimated from the prompt–gamma distribution, is either considered the location of the patient surface, labelled $R_{\mathrm{PG\ pat.surf.}}$, or dose fall–off, labelled $R_{\mathrm{PG\ falloff}}$. The range is defined as $R_{\mathrm{PG}} = R_{\mathrm{PG\ pat.surf.}} - R_{\mathrm{PG\ falloff}}$. The slope of the prompt–gamma profile fall–off corresponds to the parameter b, and changes in b from R_{PG1} to $R_{\mathrm{PG2/PG3}}$ indicate differently shaped PG profiles which can be used for evaluation, as shown in section 12.2.4 (PG profile quality assessment).

In the shifting method, two PG of the same spot obtained from 2 CT scans of one patient, using CT_1 as spatial reference, are compared using the RMS error. The difference in depth $\Delta R_{\mathrm{PG}} = \Delta R_{\mathrm{PG\ pat.surf.}} - \Delta R_{\mathrm{PG\ falloff}}$ is found by shifting pixel by pixel $\mathrm{PG}_{2/3}$ on PG_1 until the RMS error in the fall–off region [p, p + 60 bins] and the patient surface region [p – 30 bins, p] (separately) is minimized. The resulting shifts are $\Delta R_{\mathrm{PG\ falloff}}$ and $\Delta R_{\mathrm{PG\ pat.surf.}}$, respectively. Interpolation is used to obtain sub–bin shifts. A large RMS error after alignment suggests shape differences between the two profiles (see section 12.2.4 – PG profile quality assessment).

To evaluate the suitability of the prompt–gamma surrogate for dose for in vivo range verification, the spot–by–spot proton range differences between CT_1 and $\mathrm{CT}_{2/3}$ estimated from the laterally integrated depth dose profiles are used as reference for the comparison of range differences obtained for the prompt–gamma emission. The lateral integration was over the area with non–zero dose contributions and over a single pencil beam. The proton range R_{dose} is in the literature generally defined by the 80% distal dose fall–off. However, this definition does not account for Bragg peaks distorted by range mixing effects due to tissue heterogeneities. To use a range definition which is more robust than the 80% fall–off, the shifting method was also applied to the dose. The spot–by–spot dose profiles of $\mathrm{CT}_{2/3}$ were shifted to the dose profiles of CT_1, until the RMS calculated within the fall–off window defined by the 80% level of maximum dose [d_{80} – 10 bins, d_{80} + 55 bins] was minimized, where d_{80} is the depth of the 80% dose level. The obtained shift was labelled $\Delta R_{\mathrm{dose\ falloff}}$. The difference $\Delta R_{\mathrm{dose\ pat.surf.}}$, corresponding to changes in the position of the patient surface between CT_1 and $\mathrm{CT}_{2/3}$, was estimated from the CT number profiles along the central axis of the pencil beam. A CT number threshold of −200 HU was used to define a shifting window [$d_{\mathrm{–200HU}}$ – 55 bins, $d_{\mathrm{–200HU}}$ + 10 bins] where $d_{\mathrm{–200HU}}$ is the depth where CT number becomes −200 HU. The shifting method was used to identify patient surface differences between $\mathrm{CT}_{2/3}$ and CT_1, with $\Delta R_{\mathrm{pat.surf.}}$ taking the value of the shift.

PG profile quality assessment

Figure 12.3 shows the potential impact of variations in patient anatomy on dose deposition during the course of treatment. Due to displacement of lateral heterogeneities between fractions, mainly caused by rectal filling variations and femoral head rotations, the pencil beam dose distribution in later fractions may differ significantly from the initial plan. In figure 12.4, differently shaped Bragg peaks with corresponding prompt–gamma profiles are shown for 4 different spots calculated on CT_2. Variations in the slope of the PG fall–off from CT_1 to $CT_{2/3}$ are generally related to range mixing caused by changes in the patient's anatomy. These changes were detected using the parameter b of the fitting function, and spots with $\Delta\frac{1}{b} = |\frac{1}{b_2} - \frac{1}{b_1}| > 5\,\text{mm}$ were identified as distorted spots. A corresponding definition for distorted Bragg peaks with the shifting method was established using the RMS after the prompt–gamma profile alignment, with RMS > 0.0065 corresponding to distorted spots.

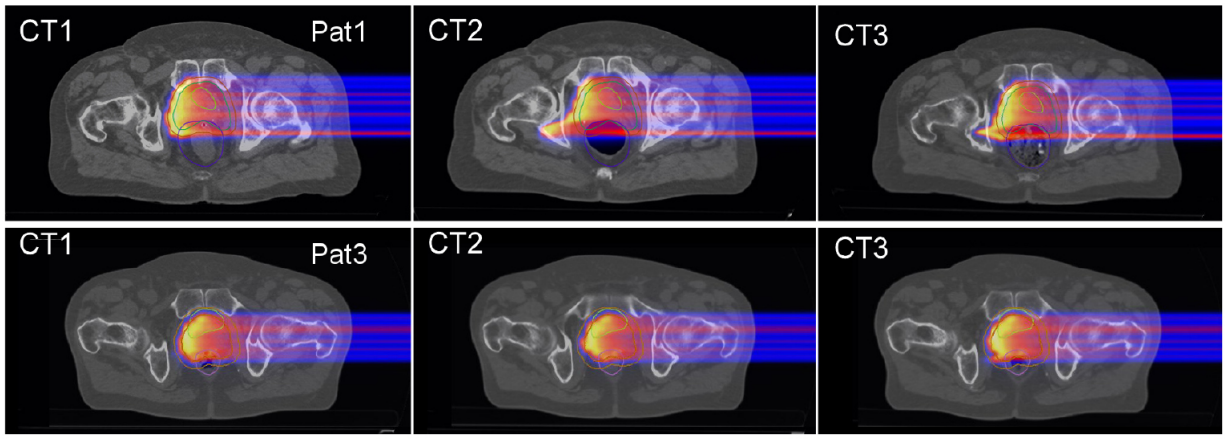

Figure 12.3: (Top row) Patient 1 (Pat1) CT scans taken on consecutive days. The plan optimized on CT_1 is applied on $CT_{2/3}$. The dose distributions of the highest intensity spots are displayed. Contours for CTV, PTV, bladder and rectum are drawn. Overshooting due to anatomical changes is visible. (Bottom row) The same for patient 3 (Pat3).

12.3 Results

12.3.1 Energy threshold and acceptance angle

In figure 12.5, the prompt–gamma spectrum from PS_1 for one pencil beam is presented. In human tissue, the main contributing elements to prompt–gamma emission are ^{16}O, ^{12}C and ^{14}N (Verburg et al. 2012). The capture of a thermal neutron by hydrogen [$^{1}H(n,\gamma)^{2}H$] can emit a gamma at an energy of about 2.2 MeV (Kozlovsky et al. 2002). Those prompt–gammas are not strongly correlated to the proton range. In addition there are annihilation processes emitting 511 keV gammas. These latter gamma lines, although correlated to the

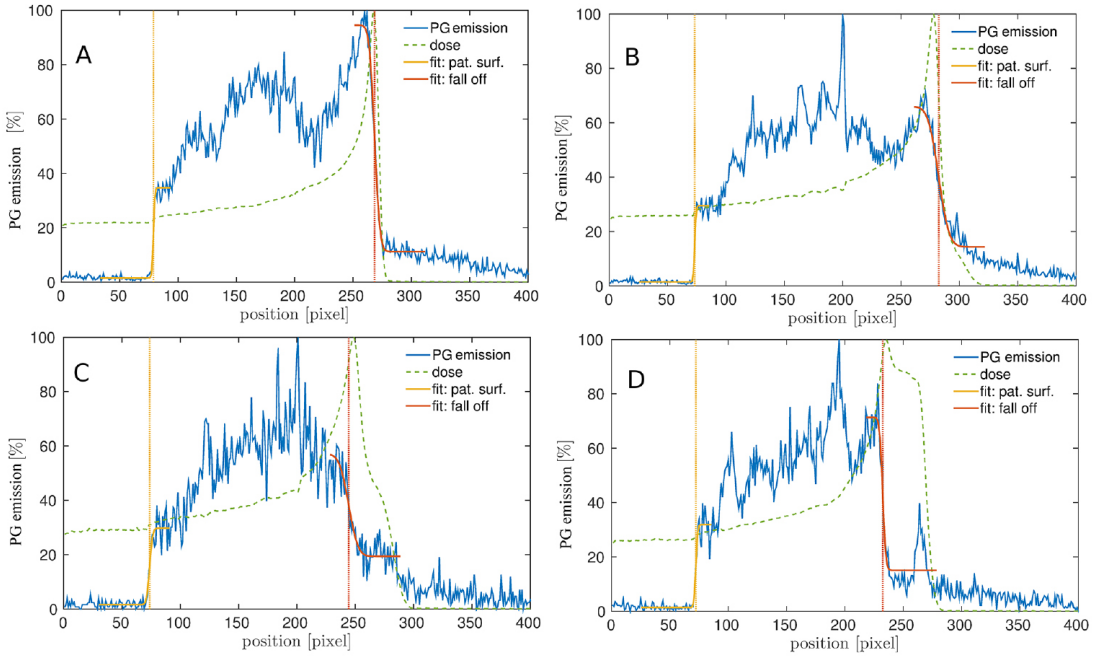

Figure 12.4: Tissue heterogeneities lead to differently shaped Bragg peaks (laterally integrated deposited dose). Four Bragg peaks with the corresponding prompt–gamma emission profiles are displayed, together with the fitted sigmoidal function. (A)–(D) represent four different spots delivered to CT_2. In (B), the dose fall–off is not as sharp as in (A), which is reflected by the prompt–gamma profile. (C) and (D) show profiles with large variation compared to a typical steep pristine Bragg peak in the patient (A).

proton range, cannot provide a real-time range verification, due to their timing dependence on the β^+–emitting nuclei half–life. Thus in this study a detection threshold of 2.5 MeV was used for all cases. In the considered example and energy window, the main contribution is from ^{12}C at 4.44 MeV, from proton induced reactions on ^{12}C or ^{16}O.

The acceptance angle of prompt–gammas reaching the phase space plane was set to $\pm 3^o$. This was based on the study of the impact of varying the acceptance angle from $\pm 2^o$ to $\pm 4^o$ on the PS_1 profile at the fall–off and at the patient surface, shown in figure 12.6. For comparison we also present PG_1. The trade–off between statistics and fall–off sharpness is quantified in table 12.1. The smaller the angle the better the profile is correlated with that from the emission point. In this study the acceptance angle of prompt–gammas reaching the phase space plane was chosen as $\pm 3^o$ as a reasonable compromise. Angles between $\pm 0.5^o$ to $\pm 3^o$ have been used previously (Lopes et al. 2012, Smeets et al. 2012, Janssen et al. 2014). The proton fluence for PG was adapted such that the prompt–gamma counts correspond to the $\pm 3^o$ acceptance angle of PS.

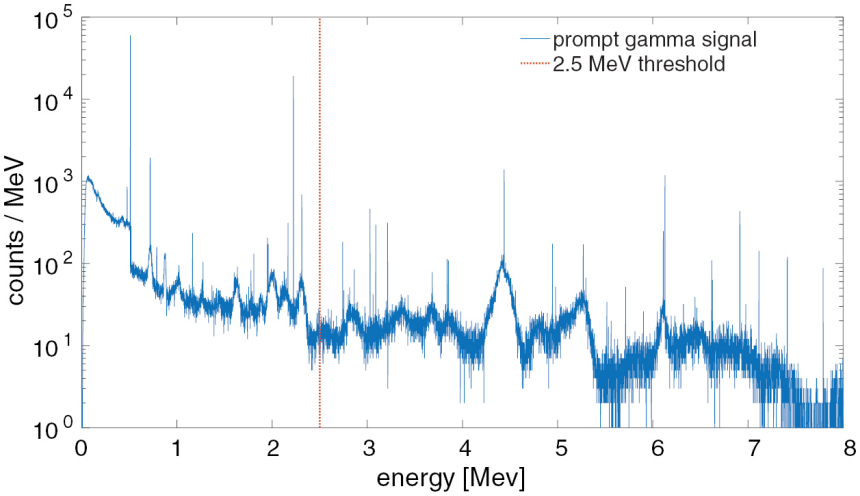

Figure 12.5: Energy spectrum of prompt–gamma photons detected at the phase space plane for a 190 MeV proton beam (2.4×10^8 primary protons) reaching the distal edge of the prostate. The vertical dashed line corresponds to the chosen threshold. No angular acceptance was applied here.

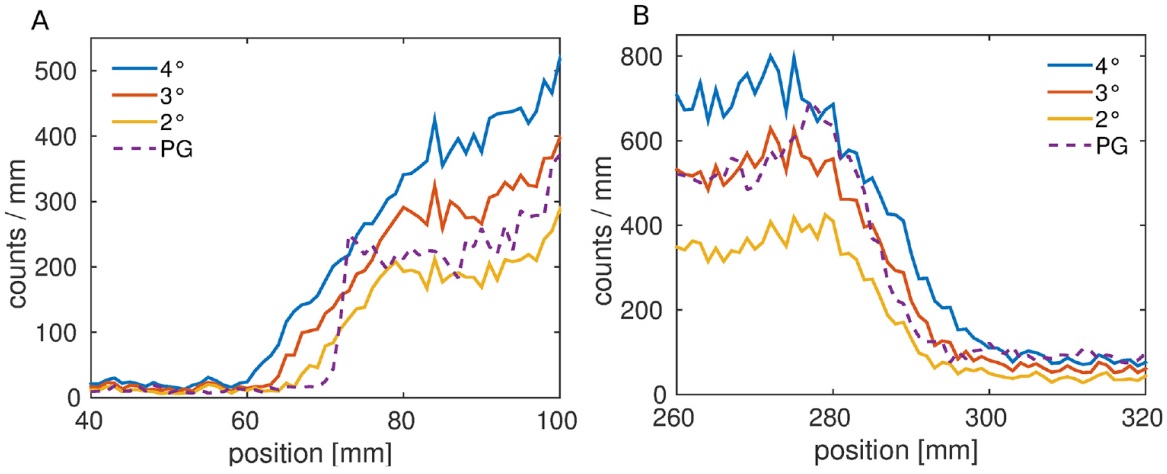

Figure 12.6: Total prompt–gamma photon counts obtained on the phase space plane (PS) for 2.4×10^8 incident protons; an energy threshold of 2.5 MeV is applied. PS for different acceptance angles ($\pm 4^o$, $\pm 3^o$ and $\pm 2^o$) are displayed for (A) the rise and (B) the fall–off region. Prompt–gamma photons scored at the emission point PG for the rescaled proton fluences are shown by the dashed line.

Table 12.1: Overview of the of the prompt–gamma profile rise (patient's surface) and fall–off deduced from the parameter b and c of the sigmoidal fit, where b is related to the slope of the curve and c to the inflection point. The number of all scored prompt–gammas for different angles are given.

Parameter of the sigmoidal function	PG	PS±2°	PS±3°	PS±4°
Pat. surf. c (mm)	72	72	71	70
Pat. surf. $1/b$ (mm)	2.5	5.1	6.1	7.5
Fall–off c (mm)	286	287	287	288
Fall–off $1/b$ (mm)	–4.6	–6.0	–7.2	–9.5
Prompt–gamma count	103 000	68 000	103 000	137 000

12.3.2 PS versus PG range assessment

In figure 12.6, PS corresponding to an acceptance angle of ±3° with full incident proton fluence is compared with PG obtained using the reduced incident proton fluence. Figure 12.7(A) shows the comparison of range difference ΔR_{PG} and ΔR_{PS} with respect to ΔR_{dose} using the shifting method for 36 spots of patient 1 for CT_1 and CT_2, where PS was calculated with the full plan proton fluence. The shift behaviour of ΔR_{PS} is compared to the shift behaviour of ΔR_{PG} , showing comparable results. In figure 12.7(B), the residual shift differences of the distributions $\Delta R_{\mathrm{PG}} - \Delta R_{\mathrm{dose}}$ and $\Delta R_{\mathrm{PS}} - \Delta R_{\mathrm{dose}}$ are displayed as boxplots, where the red line indicates the median, the extent of the box reflects the 25^{th} and 75^{th} percentiles (P_{25} and P_{75}) and the whiskers depict the range of the distribution excluding outliers. Box plot outliers were defined as values above $\mathrm{P}_{75} + 1.5(\mathrm{P}_{75} - \mathrm{P}_{25})$ and below $\mathrm{P}_{25} - 1.5(\mathrm{P}_{75} - \mathrm{P}_{25})$. The notches represent the confidence levels around the median of the particular data set. The distributions of range shift differences between prompt–gamma and dose were evaluated in terms of precision (defined as half the 95% inter–percentile range IPR) and accuracy (median). An IPR of 1.7 mm was calculated from the PS distribution ($\Delta R_{\mathrm{PS}} - \Delta R_{\mathrm{dose}}$) while for PG ($\Delta R_{\mathrm{PG}} - \Delta R_{\mathrm{dose}}$) it was found to be 1.3 mm. The same median was observed in both distributions and overlap of the notches suggests negligible differences. Hence, for further analyses, only PG was investigated.

12.3.3 Comparison of prompt–gamma ranges of consecutive CT scans

In figures 12.8 and 12.9, the proton range differences ΔR_{PG} in patient 3 between CT_1 and $\mathrm{CT}_2(\mathrm{CT}_3)$ are displayed in beam's eye view (BEV). The shifting method and the fitting method have a comparable performance. Both methods show good correlation between ΔR_{PG} and ΔR_{dose} also for large deviations between CT_1 and CT_3 which are due to variations in anatomy. Figure 12.8 shows ΔR_{PG2} (corresponding to CT_2 versus CT_1) shifts of less than 1 cm. In figure 12.9, ΔR_{PG3} suggests large anatomical changes for patient 3 between CT_1 and CT_3 with shifts of more than 3 cm (see figure 12.3). Assuming a linear correlation between ΔR_{dose} and ΔR_{PG} , the IPR was calculated excluding distorted

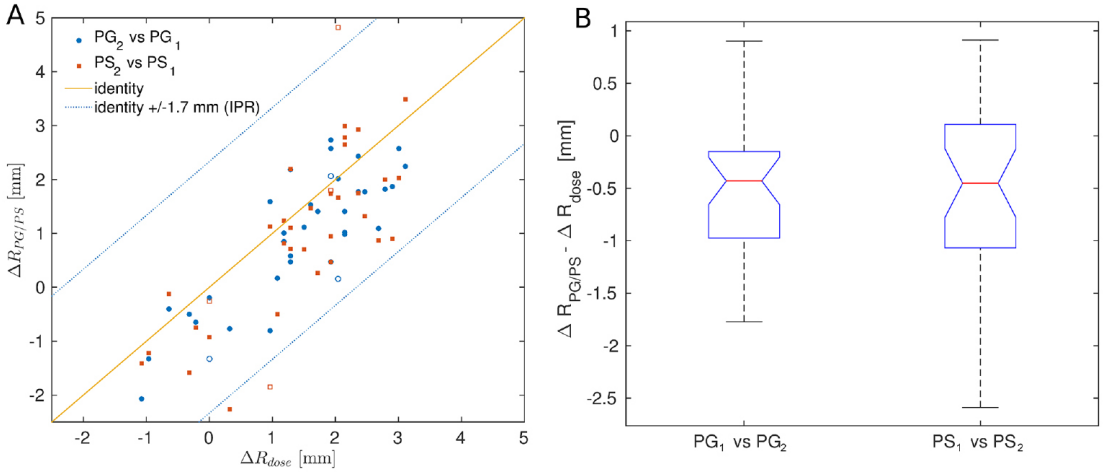

Figure 12.7: (A) Shift differences from PG_1 to PG_2 (ΔR_{PG}) marked with blue circles. The red squares refer to shift differences of PS_1 versus PS_2 (ΔR_{PS}). Spots with RMS > 0.00065 are shown with open symbols (see section 12.2.4 – PG profile quality assessment). The IPR of 1.7 mm was calculated from the ΔR_{PS} – ΔR_{dose} distribution, and is indicated by the dashed blue lines. For ΔR_{PG} – ΔR_{dose} the IPR was found to be 1.3 mm. (B) Notched boxplots of the shift difference between ΔR_{PG} – ΔR_{dose} (PG_1 versus PG_2) and ΔR_{PS} – ΔR_{dose} (PS_1 versus PS_2) for the distribution from 36 spots. The notches indicate the estimated 95% confidence level of the median.

spots as defined above. The IPR ranged from 1.7 mm to 2.7 mm for the data presented in figures 12.8 and 12.9. The medians were generally smaller than 0.5 mm.

Results of the shifting and fitting methods for all patients are summarized by box plots in figures 12.10 and 12.11, respectively. The box plots are generated after removal of distorted spots from the distributions, as defined in section 12.2.4 (PG profile quality assessment). Residual shifts between the range difference obtained by dose ΔR_{dose} and the range difference obtained by prompt–gamma emission ΔR_{PG} are displayed for the patient's surface as well as for the fall–off region separately. The combination of the range differences at the patient surface and the fall–off regions yields the total range differences. Additionally, the IPR is reported for each distribution. For the total range differences, the IPR varied between 1.4 mm and 2.9 mm for the shifting method and 1.5 mm and 3.6 mm for the fitting method.

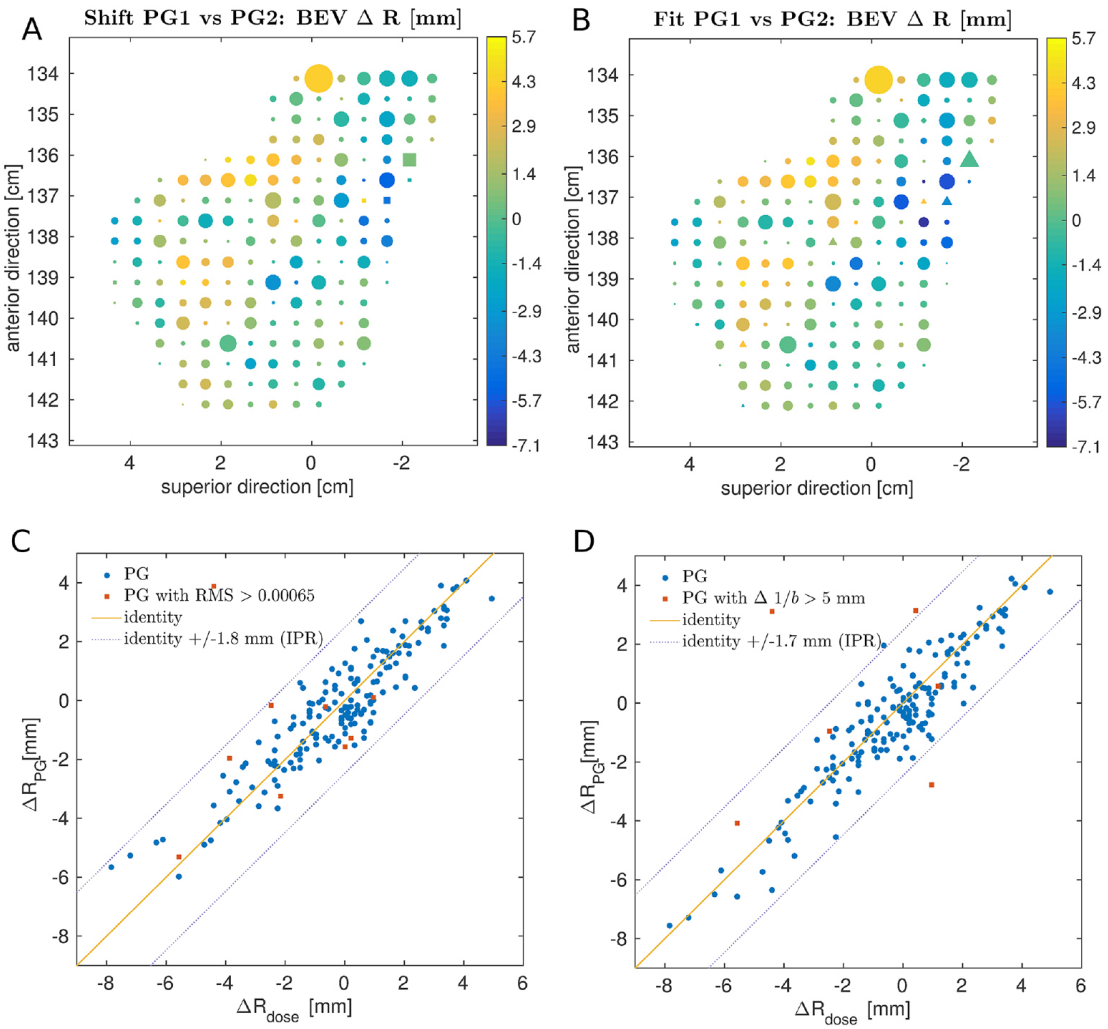

Figure 12.8: (A), (B) Beam's eye view range differences of prompt–gamma profiles (ΔR_{PG}) for patient 3 from the distal spot layer obtained (A) by shifting PG_2 on PG_1 and (B) by analysing the inflection point of the fitted function. The symbol size in (A), (B) is proportional to the number of emitted prompt–gamma photons and the magnitude of ΔR_{PG} is color coded. Square and triangular spots indicate distorted spots for the shifting and fitting methods respectively. (C), (D) Correlation of the range differences obtained from prompt–gamma profiles and dose are shown for the shifting and fitting methods respectively. The dashed lines indicate the IPR.

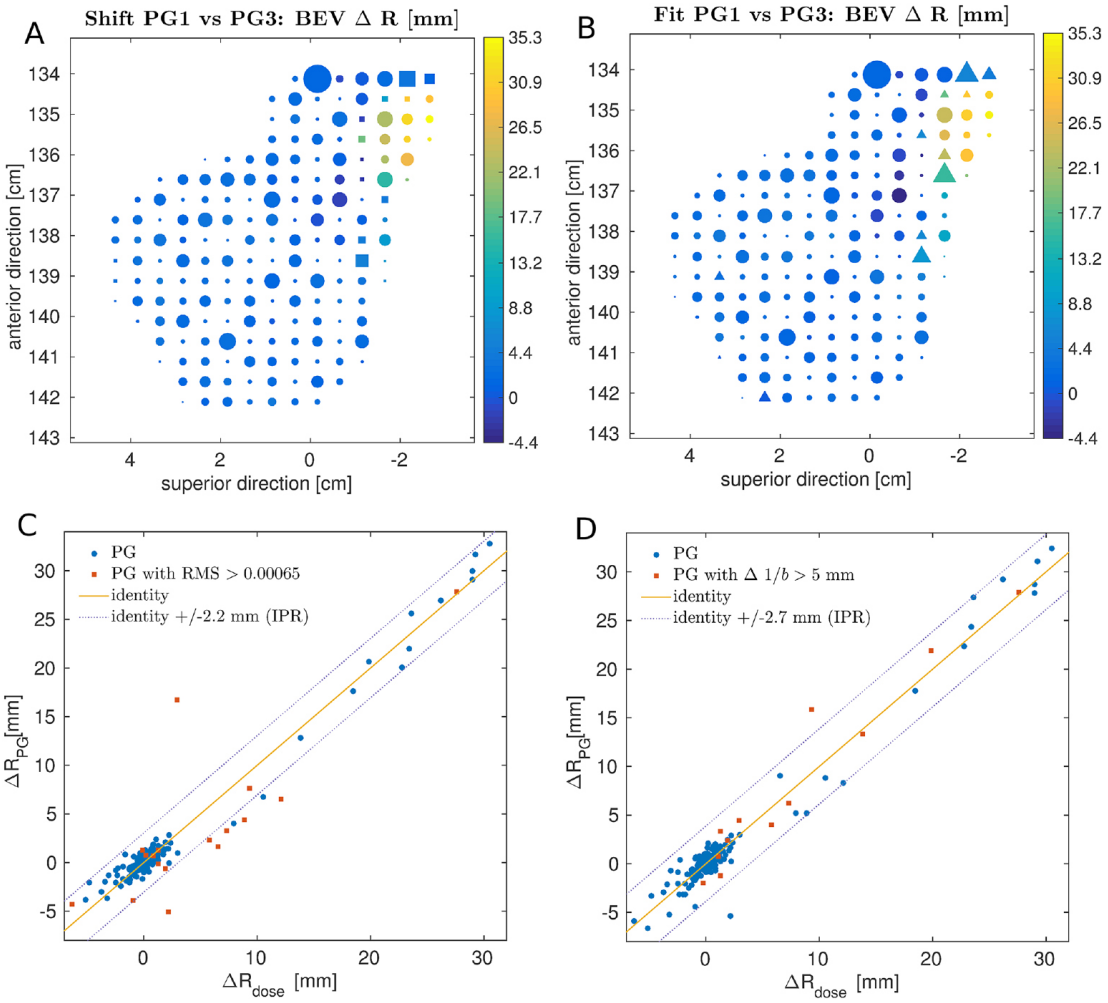

Figure 12.9: (A), (B) Beam's eye view of range differences of prompt–gamma profiles (ΔR_{PG}) for patient 3 from the distal spot layer, obtained (A) by shifting PG_3 on PG_1 and (B) by analysing the inflection point of the fitted function. The symbol size in (A), (B) is proportional to the number of emitted prompt–gamma photons and the magnitude of ΔR_{PG} is color coded. Square and triangular spots indicate distorted spots for the shifting and fitting methods respectively. (C), (D) Correlation of the range differences obtained from prompt–gamma profiles and dose are shown for the shifting and fitting methods, respectively. The dashed lines indicate the IPR.

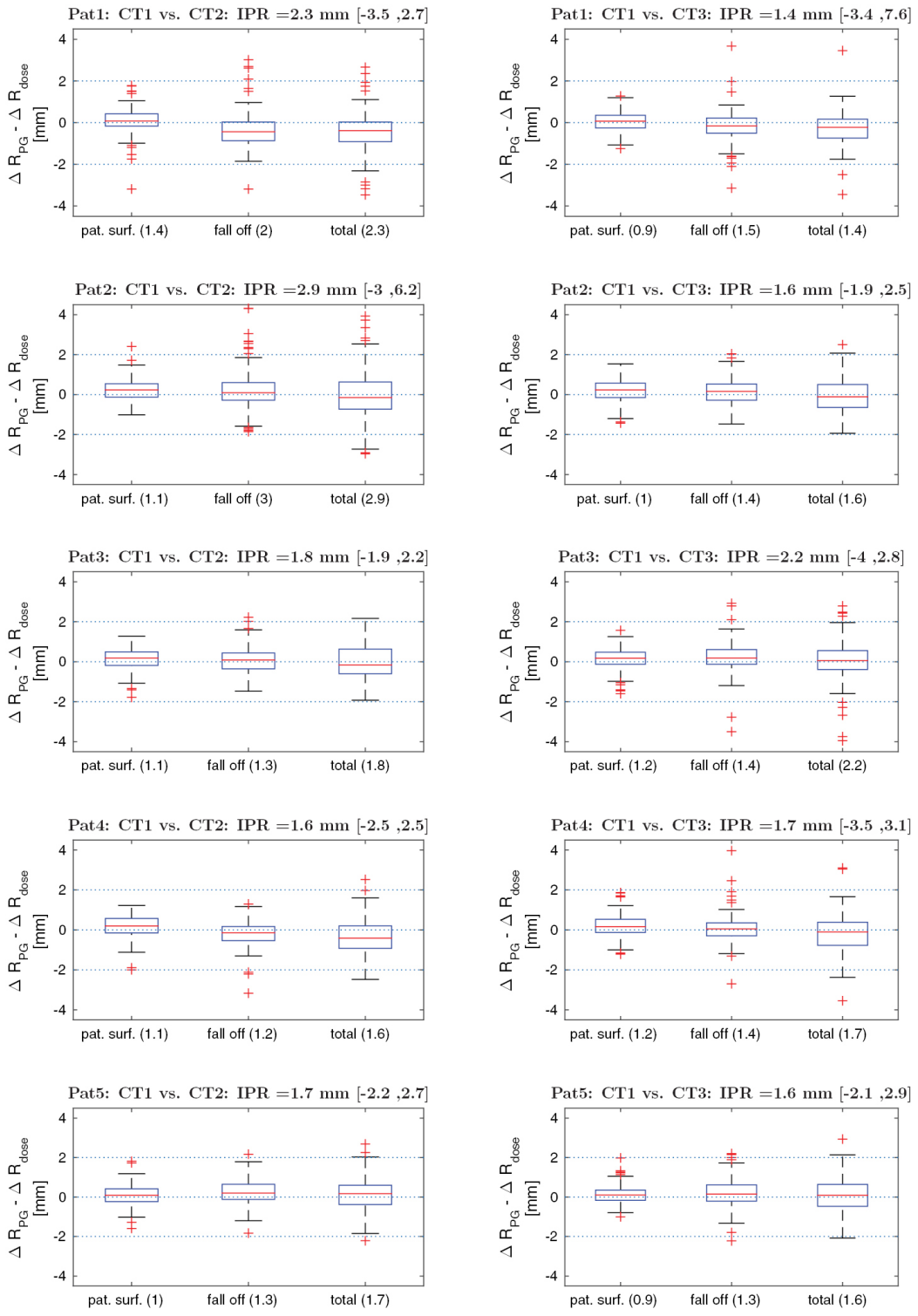

Figure 12.10: Correlation of position differences between ΔR_{PG} and ΔR_{dose} for CT_1 and $\mathrm{CT}_{2,3}$ using the shifting method. Boxplots representing the distribution of differences between ΔR_{PG} and ΔR_{dose} are displayed for all CT scan pairs. Results are shown for the patient surface and fall–off region separately, as well as their combination, resulting in total range differences. The IPR of the total shift and the range of the distribution are indicated above the plots. The IPR for the separate components are indicated in round brackets under the plots. All values are in mm.

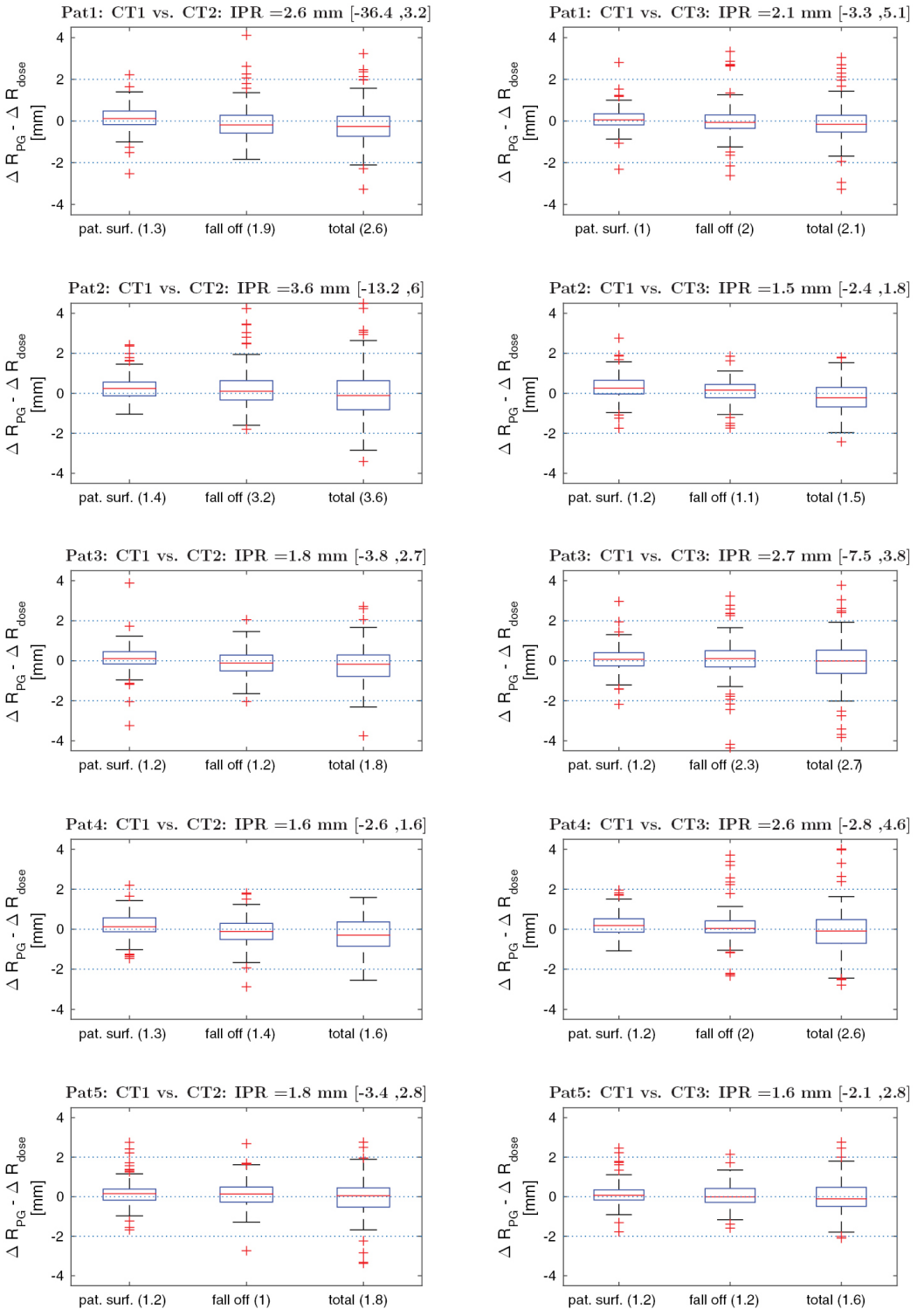

Figure 12.11: Correlation of position differences between ΔR_{PG} and ΔR_{dose} for CT_1 and $\mathrm{CT}_{2,3}$ using the fitting method. Boxplots representing the distribution of differences between ΔR_{PG} and ΔR_{dose} are displayed for the all CT scan pairs. Results are shown for the patient surface and fall–off region separately, as well as their combination, resulting in total range differences. The IPR of the total shift and the range of the distribution are indicated above the pots. The IPR for the separate components are indicated in round brackets under the plots. All values are in mm.

Figure 12.12 reports the correlation between ΔR_{dose} and ΔR_{PG} for all distributions using the shifting method. A total of 148 spots identified as distorted are not included, which corresponds to about 9% of the 1738 spots. The IPR of the combined distributions (1590 spots) was estimated to be 1.9 mm for the shifting as well as for the fitting method.

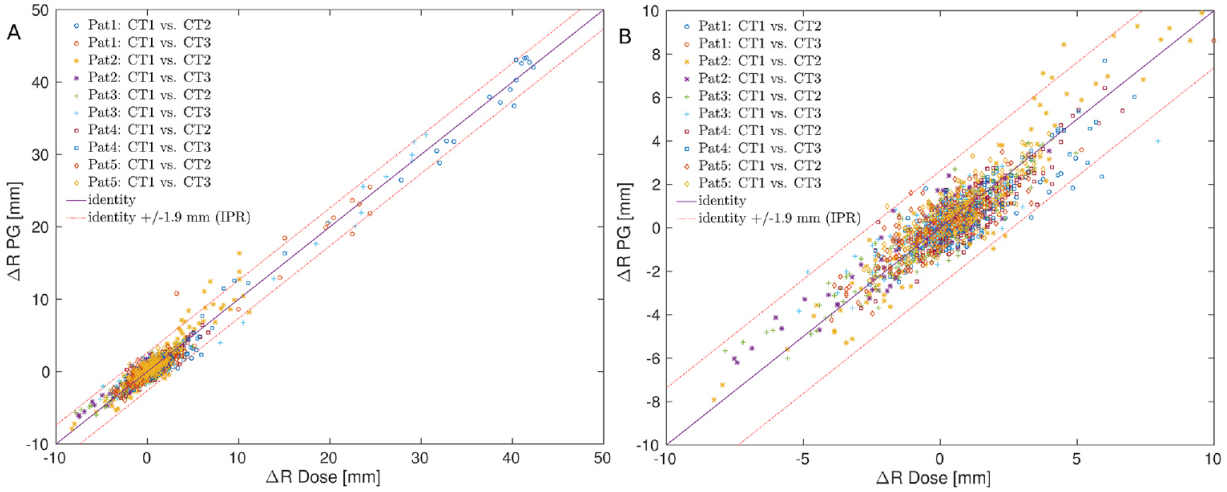

Figure 12.12: (A) ΔR_{PG} versus ΔR_{dose} for all spots considered in this study using the shifting method. The solid line is the identity and the dashed lines marks IPR for all data points excluding distorted spots. Out of 1738 spots, 143 spots are identified as distorted. (B) ΔR_{PG} versus ΔR_{dose} for spots with ΔR_{dose} between $--10$ and 10 mm.

12.4 Discussion

In this Monte Carlo study, spot–by–spot proton range verification through prompt–gamma profiles was investigated for IMPT prostate cancer patients. Two methods, based on shifting and fitting the prompt–gamma depth profiles, were developed to detect the prompt–gamma fall–off and the beam entrance at the patient's surface to derive the relative total range differences.

The differences of the location of the patient surface, of the prompt–gamma fall–off, as well as of the total range of the protons determined with both methods from the prompt–gamma profiles for the different CTs, were compared to the equivalent information obtained from the dose distributions and CT scans. For the patient's surface a precision (IPR) varying between 0.9 mm and 1.4 mm was observed for both methods. For the fall–off a precision ranging between 1.2 mm to 3.0 mm and 1.0 mm to 3.2 mm was observed for the shifting and fitting approaches, respectively, depending on the CT scan pair investigated. Medians for all distributions were below 0.5 mm. The determination of the patient entrance was more precise than that of the fall–off for both methods. This can be explained by the fact that the patient entrance is a sharper anatomical characteristic, where effects such as range mixing and the difference in the prompt–gamma emission cross sections—due to different proton energies—are negligible when compared to the fall–off region deeper in the patient. These characteristics are accentuated by making use of PG for the analysis, as discussed below. The obtained precision (IPR) of the total relative range differences ($\Delta R_{\text{dose}} - \Delta R_{\text{PG}}$) for the shifting method varies between 1.4 mm and 2.9 mm and for the fitting method between 1.5 mm and 3.6 mm. Spots which have a $\Delta R_{\text{dose}} - \Delta R_{\text{PG}}$ outside of the 2.5/97.5 percentiles used for the precision determination yield minimum and maximum deviation values of -4 mm and 6.2 mm for the shifting method, and -36.4 mm and 5.1 mm for the fitting method. For both methods, the accuracy of $\Delta R_{\text{dose}} - \Delta R_{\text{PG}}$, expressed as the median of the distributions, is below 1 mm.

To make such a study feasible in terms of computing time, we have used the profiles of prompt–gammas scored at their emission point (PG). In order to have a realistic estimation of the performance of our methods, the PG profiles analyzed had the same number of prompt–gammas as the profiles scored outside the patient with an angular acceptance of $\pm 3^{\text{o}}$ (PS). The analysis based on the PG profiles yields underestimated IPR results (1.3 mm versus 1.7 mm, section 12.3.2), due to the fact that the rise and the fall–off of the PG profiles are steeper than the ones of the PS profiles. To quantify this effect, we have compared the precision (IPR) obtained with the shifting method from the PG profiles and the PS profiles from one patient and two CT scans. Using the PG profiles, the precision (IPR) of $\Delta R_{\text{dose}} - \Delta R_{\text{PG}}$ is IPR= 1.3 mm, increasing to IPR= 1.7 mm for the PS profiles. The sub–millimeter difference in both IPR and median supports the use of PG profiles instead of the PS profiles.

The PS does not represent or model a specific camera designed for proton range verification by means of prompt–gamma detection. It is an idealized scoring plane, describing prompt–gammas outside the patient (in a location just in front of a potential camera) and with an angular acceptance compatible to that of some collimated camera designs. The

exact shape of the profile recorded by a real camera could differ significantly from that from our PS. For example, the exact collimator septa geometry and the field of view of a camera can have a very large impact on the number of recorded number of prompt–gammas (geometrical efficiency). As the detection efficiency of MeV photons from scintillating crystals is rather high, the scintillator itself is not expected to have a large effect on the detected profile in terms of statistics. Pinto et al. (2014) reported that GEANT4 simulations using the same physics parameters as in this work may overestimate prompt–gamma counts by about 15% when compared to experimental measurements (see figure 6 in that work). The same over–estimation is expected from this work. Finally, the PS profiles analyzed in this study would most likely be measured by a collimated slit or multi–slit camera based on a segmented scintillator. Our findings may not apply to different detection systems such as a Compton camera.

Large tissue heterogeneities such as air pockets close to the end of range can lead to distorted Bragg peaks (figure 12.13). In the case of figure 12.13 protons stop immediately beyond an air cavity in CT_2, causing a low intensity spike of prompt–gamma emission beyond the cavity (PG_2), and a somewhat broadened Bragg peak ($dose_2$–the cavity is not readily seen in the dose distribution since dose in air was scored). The prompt–gamma production in air is low, which causes a steep fall–off at the upstream edge of the cavity. Our method is identifying in PG_2 the steep main fall–off as the one related to the proton range (see section 12.2.4 – Region of interest identification), instead of the distal fall–off of the prompt–gamma spike. Therefore, it tries to align the steep main fall–off from PG_2 to the unique fall–off of PG_1. This results in a very small shift between PG_1 and PG_2 (see figure 12.13, where the shift between PG_2 and PG_2 aligned is almost negligible). Conversely, a large shift is seen in the dose. In such cases, the difference of range shifts obtained from the dose distribution and the prompt–gamma is large, resulting in a poor correlation in ΔR_{dose} – ΔR_{PG} . Such problematic spots can be identified from the corresponding prompt–gamma profiles by using the RMS values and the slope of the fitted function for shifting and fitting methods, respectively.

The above–mentioned precision values express the statistical effect of a large number of spots. Another source of uncertainty is the statistical uncertainty on each profile (random noise due to the stochastic nature of a Monte Carlo simulation algorithm). This has not been taken into account explicitly in our results. Nevertheless, the large number of spots (more than 1000) partially emulates this statistical uncertainty, due the fact that many spots cross similar anatomical characteristics, therefore having differences of mainly statistical nature.

For the data analysis, definitions of the patient entrance and fall–off regions are required. The extension of those regions was not the same for PG and dose, or patient surface and fall–off (see section 12.2.4 – Region of interest identification). The main criterion for this optimized selection of the extension was the reliable identification of the region for a large number of spots. Small changes to the extension of the regions have been tested and had no noticeable impact on the obtained results.

Finally, the findings of this study may be different for other anatomical sites where larger amounts of tissue heterogeneities are observed, such as head and neck or lungs.

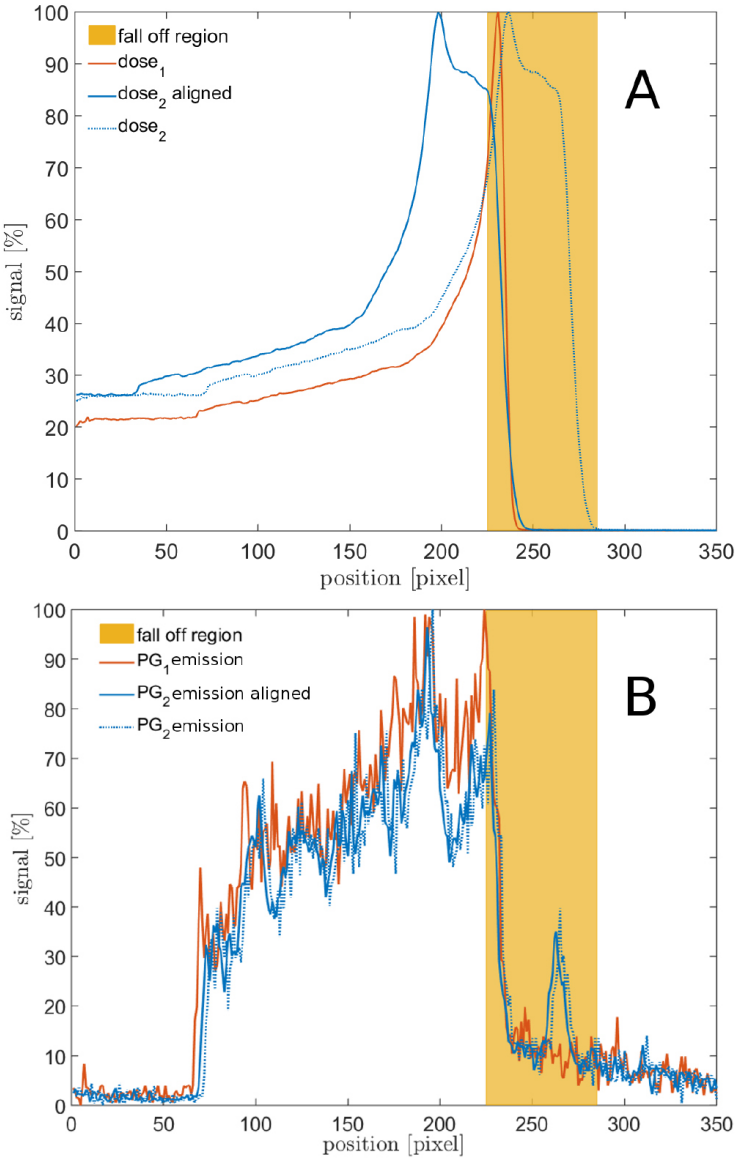

Figure 12.13: (A) Original pencil beam dose distributions on CT_1 and CT_2, labelled $dose_1$ and $dose_2$. The shifted $dose_2$ is labelled $dose_2$ aligned. The shift is obtained by aligning $dose_2$ to $dose_1$ in the shaded fall–off region. A large shift is required to align $dose_2$, due to the very different shape of the distributions caused by an air pocket (located between position 225 pixels and 250 pixels). (B) The corresponding PG_1 and PG_2 emission profiles are shown, in addition to the shifted PG_2 emission profile (labelled PG_2 emission aligned). The different shifts between (A) and (B) result in a bad correlation between ΔR_{PG} versus ΔR_{dose} for this pencil beam.

12.5 Conclusion

In this work the accuracy and precision of the detection of relative inter–fractional range shifts in prostate IMPT using the detection of prompt–gamma emission profiles was investigated using a comprehensive set of clinical data. Our results suggest that for 95% of distal spots in prostate IMPT treatments, relative range shifts from one fraction to another can be recovered by measurement of the prompt–gamma profile within $\pm 2\,\mathrm{mm}$ using either profile fitting or profile shifting methods. Methods to identify unreliable spots were presented and were shown to increase the reliability of prompt–gamma based range verification. Although translation of these results to the clinical setting will strongly depend on the performance of available collimated prompt–gamma detector systems, the theoretical detectability shown in this work confirms prompt–gamma to be a promising modality for in vivo range verification on a spot–by–spot basis.

Acknowledgments

This work was supported by the Federal Ministry of Education and Research of Germany (BMBF), grant number 01IB13001 (SPARTA), and by the German Research Foundation (DFG) Cluster of Excellence Munich–Centre for Advanced Photonics (MAP). The authors would like to thank Professor Dr Jan Wilkens and Dr Florian Kamp for sharing their CERR implementation of proton treatment planning; Dr Jonathan Orban de Xivry for sharing registration tools; Dr Reinoud Nijhuis for help with data management as well as Fiere Janssen and Dr Marco Pinto for stimulating discussion.

Chapter 13

Radiation protection modeling of the CALA facility

The investigation presented in this chapter was published in the Journal of Radiological Protection, in volume 40, issue 4, page 1048, in September 2020, with the title *Radiation protection modeling for 2.5 petawatt-laser production of ultrashort x–ray, proton and ion bunches: Monte Carlo model of the Munich CALA facility*, by Englbrecht et al. (2020) – (senior author: Dedes).

13.1 Introduction

Since the theoretical prediction of electron acceleration driven by high peak power femtosecond (fs) lasers in 1979, the development of techniques to increase the peak power of lasers, led to the development of fs short petawatt (PW) class lasers for particle acceleration (Tajima & Dawson 1979, Strickland & Mourou 1985).

The "Centre for Advanced Laser Applications" (CALA) in Garching, near Munich (Germany) is intended for laser–based acceleration of electron beams for brilliant x–ray generation ("Laser–driven Undulator x–ray Source" beamline (LUX), "Electron and Thomson Test Facility" beamline (ETTF), "Source for Powerful Energetic Compact Thomson Radiation Experiments" beamline (SPECTRE)) and laser-driven nanosecond bunches of protons and heavy ions ("Laser–driven Ion Acceleration" beamline (LION), "High Field" beamline (HF)) for the investigation of the laser–acceleration and application of energetic protons and ions.

Laser–accelerated ion bunches have been proposed for use in radiation therapy of cancer (Linz & Alonso 2016). Albeit the source characteristics of laser–accelerated ion bunches differ significantly from those of conventional sources, the former can be utilized in a wide range of novel biomedical applications and offer distinct advantages: reasonable compactness and cost effectiveness, synchronisation to laser pulses at a picosecond level, simple target and hence particle species changeability, broad energy spectrum, small source size,

down to nanosecond bunch duration, allowing to use new bunch detection techniques and to study ultrashort biological effects (Schreiber et al. 2016, Esarey et al. 2009, Richter et al. 2016, van de Water et al. 2019).

Laser–accelerated electron beams can be used for various imaging applications (e.g. through the generation of x–ray and gamma ray secondary sources) and, at the same time, enable and validate innovative radiation generation mechanisms (Wenz et al. 2015, Wenz et al. 2019).

The source characteristics can be changed through the choice of laser parameters and targets. Based on electron bunches accelerated in a plasma, these sources can be tailored to have the necessary properties of being compact and of delivering collimated, incoherent, and fs pulses of x–ray radiation (Esarey et al. 2009, Gonsalves et al. 2015, Lu et al. 2007).

The ionizing radiation emerging from the experiments using the up to 2.5 PW laser pulses with 25 fs duration will be mixed particle species of high intensity, high maximum energy, broad energy spectrum and short duration; thus posing new challenges on shielding and monitoring compared to conventional radiation protection. The mixed particle species' nature of the bunches impedes the usage of conventional particle specific shielding concepts. Conventional radiation protection mazes are designed such that particles need to undergo at least two scatters to escape the room. Such designs, which avoid straight penetrations as straight paths, can not be applied in CALA, since the optical laser pulses need to enter the experimental caves though straight penetration holes. Their diameter should be also large enough to keep the power density on the laser mirrors below their damage threshold.

As part of the legal regulations, areas in the CALA building have to be subdivided in the radiation protection categories unclassified, supervised and exclusion areas, based on the predominant prompt dose–rate to be expected (BfUNnS 2018). Personal electronic dosimeters, as required in controlled radiation protection areas, have been reported to be unusable due to under-response already for bunches of millisecond duration (Borne et al. 2002, Ankerhold et al. 2009, Ambrosi et al. 2010).

Detailed dose rate estimations from Monte Carlo simulations can provide the insight necessary to categorize the local radiation exposure, especially since the envisioned particle source terms are still subject of active research and have not yet been reached experimentally.

We report here the results of Monte Carlo simulations of different source types relevant to various CALA experiments, in terms of calculated dose rates, with respect to established dose limits.

13.2 Materials and methods

13.2.1 CALA laser systems

The Centre for Advanced Laser Applications hosts two unique laser systems to drive the ionizing radiation produced in laser–plasma experiments: ATLAS–3000 and PFS–pro.

Table 13.1: Laser systems installed in CALA and some of their main operation parameters.

	ATLAS–3000	PFS–pro
Power [10^{12}W]	2500	10
Pulse length [fs]	25	≤ 30
Pulse Energy [J]	60	0.1
Wavelength [nm]	750–850	700–1400
Shot frequency [Hz]	1	1000

ATLAS–3000 laser system

The workhorse of CALA is the "Advanced Titanium-Sapphire Laser" system ATLAS–3000, which uses the chirped pulsed amplification (CPA) technique and has been operated for 20 years now and subsequently upgraded to reach a peak power of up to 2.5 PW. Its main specifications are summarized in table 13.1. This peak power makes ATLAS the highest peak power laser system operated by a university in Europe in 2020. ATLAS is used to drive the experiments in LION, HF, ETTF, LUX and SPECTRE.

Its infrared pulses (central wavelength $\lambda = 800\,\text{nm}$) contain an energy (after pulse compression) of up to 60 Joule in a duration of 25 fs and are guided as a 28 cm diameter beam via the evacuated laser beam delivery from the laser hall to five dedicated experimental rooms called caves. There, they are focused by off–axis parabolic mirrors down to micrometer small spot sizes.

By placing solid targets or gas jets in the focus, the intensity of up to 10^{23} W/cm^2 drives electrons or ions to relativistic velocities (see table 13.2). Those targets absorb the laser and represent the source of energetic particles, i.e. the source of the accelerator. The pulse repetition frequency on target can reach up to 1 Hz.

PFS–pro laser system

The "Petawatt Field Synthesizer laser system" (PFS–pro) is the second laser system under development in CALA with unique broadband laser characteristics (see table 13.1). Once completed, PFS–pro will seed the SPECTRE experiment.

Using a repetition frequency of up to 1 kHz, the SPECTRE experiment will use the highly energetic photons (50 keV – 200 keV) to generate x–rays by acceleration of electrons up to 70 MeV (see table 13.2). The x–rays will be used for medical imaging, among other applications.

13.2.2 CALA building and cave infrastructure

The five caves corresponding to the five CALA experiments are arranged around a central beam dump of 5.85 m thickness (figure 13.1). Due to the design of the CALA facility, the walls of the experimental caves, the roof and the floor as well as the central beam dump are the critical components that shield the exterior and the other experimental

Table 13.2: Simulated particle source parameters of the five experiments in CALA: LION and HF (Schreiber et al. 2016, Daido et al. 2012, Macchi et al. 2013, Lindner et al. 2018) as well as LUX and ETTF (Esarey et al. 2009, Gonsalves et al. 2015, Lu et al. 2007) are seeded by ATLAS–3000. SPECTRE can be seeded by ATLAS and PFS–pro.

	Beam component	Energy spectrum	Spectral shape	Full divergence angle	Charge per pulse	Shot frequency
LION	^{12}C	10 – 400 MeV/u	Box	180 mrad	0.016 nC	1 Hz
	p^+	10 – 200 MeV	Box	180 mrad	0.16 nC	1 Hz
	e^-	10 – 1000 MeV	e^{-x}	10^3 / E[MeV] mrad	1.6 nC	1 Hz
HF	^{97}Au	1 – 10 MeV/u	Box	200 mrad	0.016 nC	1 Hz
	^{12}C	10 – 200 MeV/u	Box	200 mrad	0.016 nC	1 Hz
	p^+	10 – 200 MeV	Box	200 mrad	0.16 nC	1 Hz
	e^-	10 – 1000 MeV	e^{-x}	10^3 / E[MeV] mrad	1.6 nC	1 Hz
LUX	e^-	5000 MeV	Mono	2 mrad	1 nC	1 Hz
	e^-	10 – 5000 MeV	Box	2 mrad	1 nC	1 Hz
ETTF	e^-	5000 MeV	Mono	2 mrad	1 nC	1 Hz
	e^-	10 – 5000 MeV	Box	2 mrad	1 nC	1 Hz
SPECTRE	e^-	70 MeV	Mono	10 mrad	0.025 nC	1000 Hz
	e^-	500 MeV	Mono	10 mrad	0.25 nC	1 Hz
	e^-	10 – 500 MeV	Box	10 mrad	0.5 nC	1 Hz

areas. Surrounding the five experimental caves are: a hall southeast of the central beam dump (labeled as "Experimental hall" in figure 13.1) that will host x–ray cabins for x–ray experiments, on the south a laboratory area (labeled as "Laboratory" in figure 13.1) and several corridors. The laser beam line (indicated by the dashed lines in figure 13.1) enters the five experimental caves through several penetration holes through the interior walls of CALA.

The central beam dump, walls, roof and cave doors consist of magnetite concrete (orange in figure 13.2) and magnetite aggregate of various water content (purple in figure 13.2) in a sandwich–like structure between either concrete or steel container walls. Such technically easy to realise sandwich–like construction was used for radiation protection buildings previously (Forster 2020). For CALA, the magnetite was compacted to a density of $\rho = 4.0\,\mathrm{t/m^3}$ and allowed for thinner shielding walls at lower cost. The used magnetite filling without cementitious interconnection is sustainable, since it can be reused after the operation of CALA will have ended, in contrast to regular heavy concrete (Forster 2020). The floor is made of 75 cm standard concrete using Portland cement to shield the groundwater from the radiation. The roof on top of the caves is 90 cm of compacted magnetite aggregate filled into rectangular steel pipes.

All caves (except of HF) are 18 m long, 3 m wide and 4.25 m high. The shorter HF cave is 13.81 m long, 4.3 m wide and 4.25 m high. A false floor is placed at 1 m above the concrete floor. The 20 cm radius laser penetration holes are located 50 cm below the false floor (see figure 13.3 for details of the modelled geometry at the beam line level). These holes are locations where dose can potentially leak outside of the experimental caves and are of particular interest in our study. Movable steel doors of 1 m thickness filled with magnetite aggregate seal the entrance to the caves during experiments.

Due to the high power density of the laser pulses, the interaction between pulse and

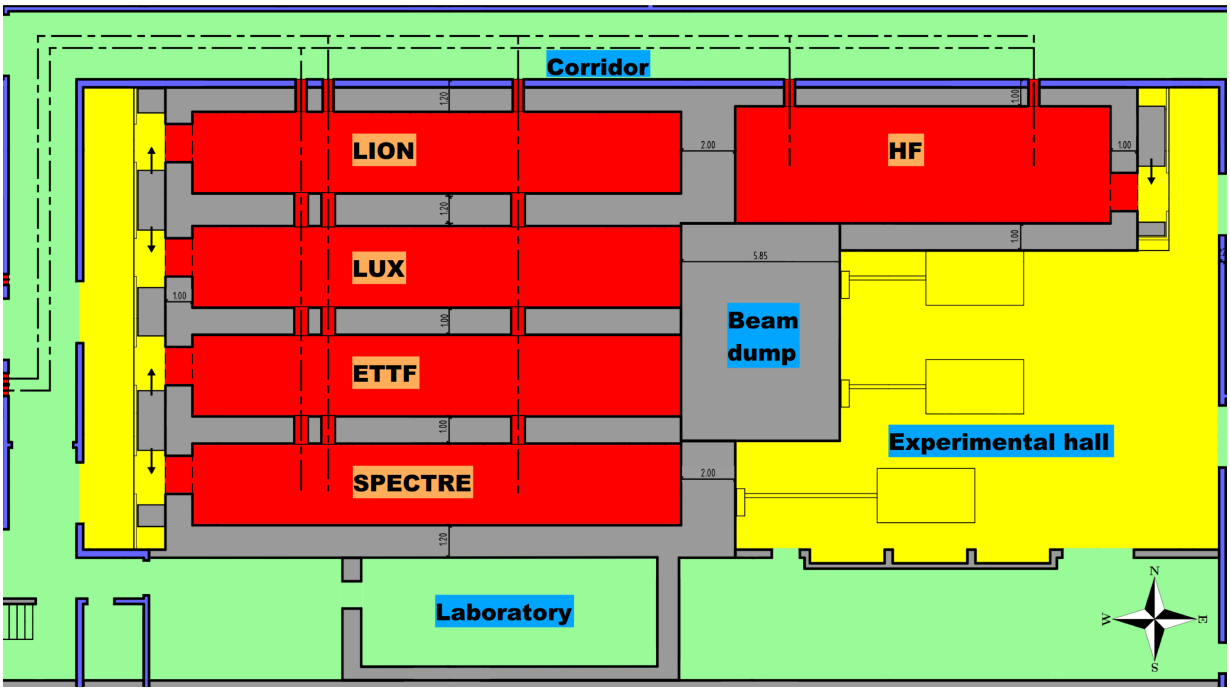

Figure 13.1: Structural and radiation zoning layout of the CALA facility. The five experimental caves are arranged around the central beam dump. East of the beam dump is an experimental hall hosting experiments inside x–ray cabins. Each active cave is an exclusion area (red), supervised areas are access limited (yellow). Areas like the transport corridor and a laboratory room are unclassified areas (green) based on the expected dose rate estimate. Sliding doors (moving in the direction of the arrows) can lock the active caves. Laser pulses can enter the caves through the transport beam line (dashed lines).

solid or gas target is performed in high vacuum (10^{-6} mbar). Above the false floor, the caves contain large vacuum chambers (depicted as white boxes in figure 13.2). The cubic modules ($1.21 \times 0.98 \times 1\text{m}^3$) are made of 2.5 cm thick aluminium plates mounted on steel frames and are connected to the beamline vacuum pipe system. Each chamber weighs 1.2 t.

The details of the structures in each experimental cave relevant to this radiation protection investigation are described in the following.

13.2.3 CALA experiments

Laser–driven Ion Acceleration (LION)

The LION experiment (see figure 13.1) places 10 nm – 1000 nm thin plastic, metal or diamond–like carbon foils targets into the focused ATLAS pulses. The targets are located 232 cm above the concrete floor. The high electric field of the laser pulse drives the electrons out of the foil. They form a sheath at the target rear boundary (Snavely et al. 2000). The arising field gradient between displaced electrons and foil nuclei can reach up to TeV/m

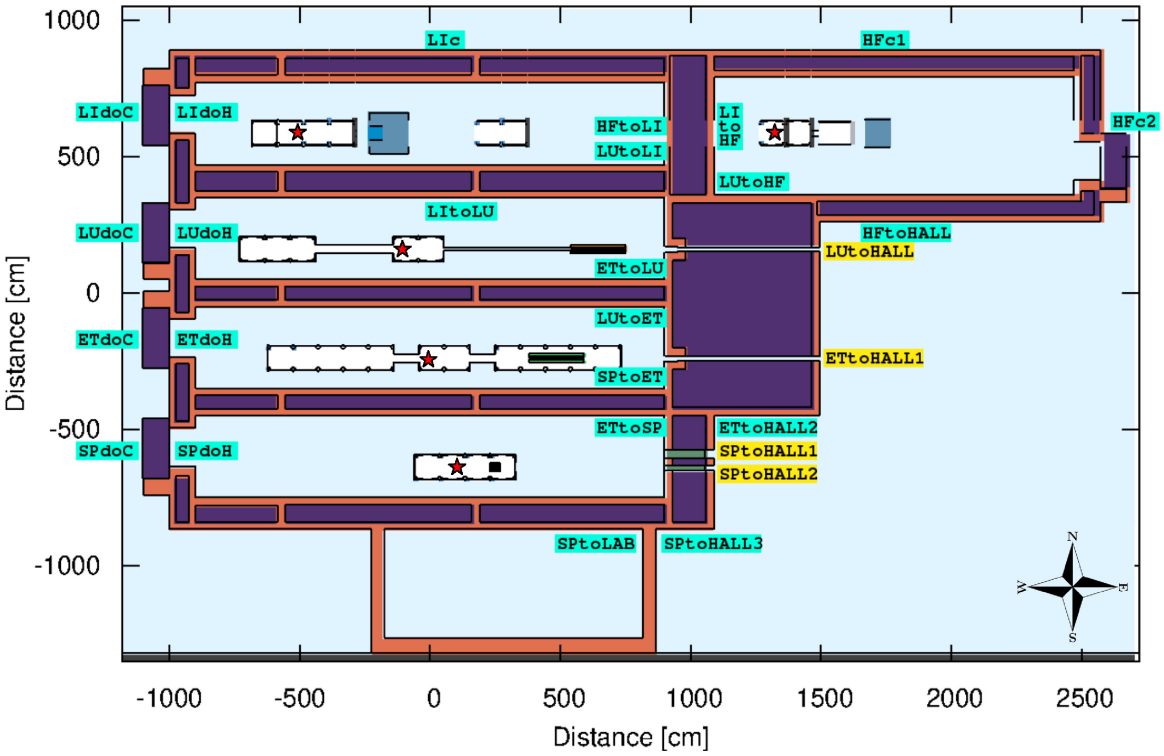

Figure 13.2: Layout of the CALA geometry model implemented in FLUKA. The 2D cross section was taken at the height in which the laser impinges onto the targets. These positions are indicated with a star in each cave and the beams are propagated in general eastwards. The vacuum chambers (in white) as well as the LION (gray–blue) and HF beam dumps (gray) are also shown. The labels indicate critical locations where the dose was quantified within cylindrical volumes. Turquoise and yellow colours correspond to two different radii of these cylindrical volumes.

and accelerates the hydrogen or carbon contamination on the foil surface up to hundreds of MeV (see table 13.2), which then propagate to the east towards the LION beam dump.

On its south side the LION cave borders with the LUX cave and on the east side with the HF beam line. On the north and west side it adjoins the north corridor and the entrance hall to the experimental areas, respectively.

Walls of 1.0 m – 1.2 m thickness are employed in order to keep the dose rates outside the cave below the designated levels. The wall in forward direction towards the east is 2 m thick and shields the HF cave during beam operation in LION. Three laser penetration holes (see figure 13.3) are running through the north (corridor – LION) and south walls (LION – LUX).

The cave geometry model (see figure 13.2) contains a set of four vacuum chambers for the acceleration experiment and two for offline tests without ATLAS.

A dedicated, hybrid beam dump (see figure 13.2) was designed for LION with the

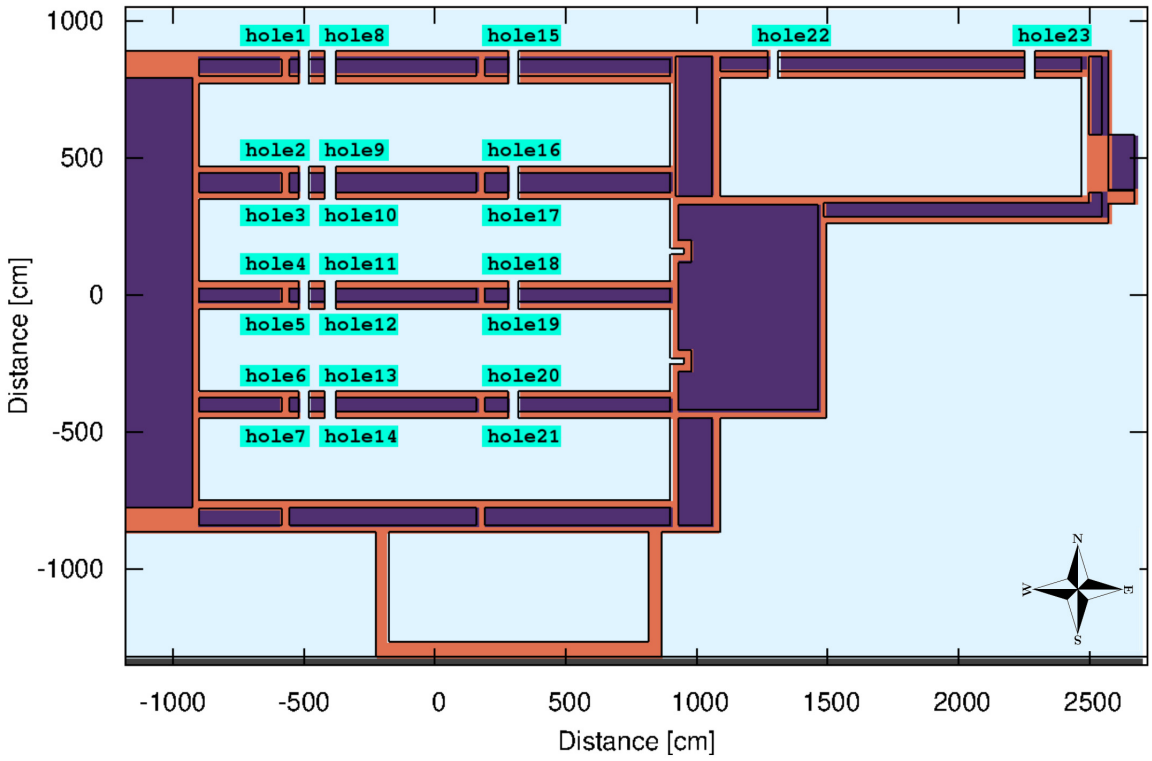

Figure 13.3: Layout of the CALA geometry model implemented in FLUKA. The 2D cross section was taken at the height in which the laser enters the caves. The labels indicate critical locations where the dose was quantified within cylindrical volumes.

purpose of stopping the beam and containing as much as possible of the primary and secondary particles' fluence and dose. The size of the beam dump is $1.5\,\mathrm{m} \times 1.5\,\mathrm{m} \times 1.5\,\mathrm{m}$. It is made of heavy concrete, except for a cubic volume of $0.5\,\mathrm{m} \times 0.5\,\mathrm{m} \times 0.5\,\mathrm{m}$ filled with water. The water tank is placed in the middle of the beam dump on the vertical and horizontal dimension and starts at the front face of the beam dump.

High–Field (HF)

The HF experiment aims to explore the acceleration and interaction characteristics of ultrashort and intense ion bunches (Lindner et al. 2017). Electrons, protons and heavier ions will be accelerated using foil targets (see table 13.2). The goal is preparatory research for the generation of extremely neutron–rich isotopes to answer questions in astrophysical nucleosynthesis of heavy elements like gold and uranium by experimentally realizing the fission–fusion reaction mechanism, which is yet inaccessible using conventional particle accelerators (Habs et al. 2010).

On its south side, the HF cave borders with the experimental hall for x–ray experiments outside the ETTF and LUX cave and the central beam dump. On the west side the LION

cave is behind the 2 m thick wall. On the north and east side it adjoins the north corridor and the entrance hall to the experimental areas, respectively.

Two vacuum chambers (see figure 13.2) are connected by a 25 cm long pipe (r = 15.9 cm) to a special chamber for the HF experiment ($0.85 \times 0.8 \times 1.25$ m^3).

Since the primary radiation from the laser–driven source in HF is similar to LION–protons and heavier ions, the HF beam dump is similar. The heavy concrete beam dump is 1 m × 1 m × 1 m and stops the ion bunches and secondary particles, which are emitted 227 cm above the concrete floor. The beam dump geometry model is shown in figure 13.2.

The outer walls of HF are 1 m thick. Two laser penetration holes (see figure 13.3) are running through the north (corridor – HF) wall.

Laser–driven Undulator x–ray Source (LUX)

The LUX experiment is set for generating brilliant x–ray undulator radiation using a magnetic undulator seeded by up to 5 GeV laser accelerated electrons. In full operation, it may serve as a prototype laser–plasma–accelerator based free–electron laser (FEL) for biomedical x–ray imaging experiments.

On its north side, LUX is separated from LION by a 1.2 m wall. On its south side it borders the ETTF cave via a 1 m wall and the east side of the cave terminates with the central 5.85 m thick central beam dump. Three laser penetration holes are running through the north and three through the south walls (see figure 13.3).

Two vacuum chambers for laser preparation are connected by a 3 m long pipe (r = 35 cm) to three chambers for the interaction of pulse and gas target (220 cm above the concrete floor and propagating to the east). A 2.1 m long permanent magnet ($B = 0.85$ T) follows 4.87 m after the chamber. The magnet serves as diagnostic element for the electron spectrum and as a radiation protection device by bending the electrons downwards into the 5.85 m central beam dump.

Only the generated x–rays pass the magnet travelling straight and may traverse the central beam dump through a cylindrical 7.5 cm radius channel which is on axis with the laser propagation. All the above elements have been implemented in the LUX cave geometry model (see figure 13.2).

Electron and Thomson Test Facility (ETTF)

The ETTF experiment is a fundamental research experiment to generate brilliant hard x–rays via Thomson–backscattering and betatron radiation from ATLAS–3000 pulses. Up to 5 GeV laser–accelerated electrons (see table 13.2) and their characteristics like charge, space charge, timing and acceleration mechanism are probed experimentally, in order to provide x–rays best suited for biomedical experiments like phase contrast imaging or ultrafast phenomena in solids and matter in the plasma state (Wenz et al. 2015, Wenz et al. 2019).

On its north side, ETTF borders with LION and south with SPECTRE, both separated from ETTF by a 1 m thick wall. Like LUX, the east end of the cave is the 5.85 m central beam dump with a 7.5 cm radius cylindrical channel for x–rays. Three laser penetration

holes have the same dimensions as for LUX and are running through the north and through the south walls (see figure 13.3).

Three sets of chambers (five, two and five chambers in each set) connected by pipes were implemented into the geometry model (see figure 13.2). The interaction point of the laser pulse and the target is located 220 cm above the concrete floor and the beam propagates eastwards. The last set of chambers hosts the same 2.1 m permanent magnet ($B = 0.85$ T) as in LUX. The magnet again is used for electron energy diagnostic and as a radiation protection device by bending the electrons downwards.

For one special experiment only, the magnet will be removed, causing the 5 GeV beam to be directly aiming at the x–ray penetration hole in the 5.85 m central beam dump. In order to keep the experimental hall dose within the legal dose rate limits, the channel will be shielded using lead bricks before and after the channel.

Source for Powerful Energetic Compact Thomson Radiation Experiments (SPECTRE)

The SPECTRE experiment uses both ATLAS (up to 300 TW) and PFS–pro as driving laser. Using ATLAS, SPECTRE aims to accelerate electron bunches of 0.25 nC – 0.5 nC up to 500 MeV for x–ray generation. At the PFS–pro repetition rate of up to 1 kHz, a tunable source of x–rays for biomedical experiments using up to 70 MeV electron beams will be established. Further source characteristics are listed in table 13.2.

On its north side, SPECTRE borders with ETTF and on the south with a laboratory, both separated from SPECTRE by 1 m thick walls. The north wall has three laser penetration holes (see figure 13.3). The east wall of 1.9 m thickness shields the experimental hall from radiation created in SPECTRE and has two cylindrical holes to allow for the extraction of the x–rays. For most experiments, the two holes will be plugged with S235JR steel bars.

A set of four chambers is installed and uses a 40 cm permanent magnet ($B = 0.85$ T) for energy diagnostic and radiation protection. The chambers and the magnet are included in the SPECTRE geometry model as shown in figure 13.2.

The particle bunches will be emitted 219.5 cm above the concrete floor and propagate eastwards.

13.2.4 Monte Carlo simulations

All radiation protection calculations presented in this study were performed with the FLUKA Monte–Carlo code version FLUKA2011.2x.2. FLUKA is a general purpose tool for calculations of particle transport and interactions with matter, covering an extended range of applications (Ferrari et al. 2005, Böhlen et al. 2014). It can simulate with high accuracy the interaction and propagation in matter of about 60 different particles, with energies ranging from the keV to the TeV scale.

The PRECISIOn physics settings were chosen for all simulations. Neutrons were transported down to thermal energies, electrons down to 1.5 MeV, photons to 1 keV and all

other particles down to 100 keV. Option EMF was used to request a detailed transport of electrons, positrons and photons.

A detailed geometry model of the five experimental caves and surrounding structures was implemented. The geometry model included the walls, doors, roof, vacuum chambers, spectrometer magnets and beam dumps, all characterised by their dimensions, elemental composition, mass density and for the magnet, field strength. The vacuum pipes for the laser and other chamber content have not been included in the model.

Two horizontal cross sections of the detailed geometry model are shown: one at the height where the laser pulses impinged on the target (figure 13.2) and one at the height where the laser enters the experimental caves (figure 13.3).

As sources, electrons ($< 5\,\mathrm{GeV}$), protons ($< 200\,\mathrm{MeV}$), ^{12}C–ions ($< 400\,\mathrm{MeV/u}$) and ^{197}Au–ions ($< 10\,\mathrm{MeV/u}$) were simulated using spectra, divergences and bunch–charges based on expectations from recent experimental results (Gaillard et al. 2011, Gonsalves et al. 2015). Source characteristics are summarised in table 13.2. The assumed particle energies and bunch charges were optimistic upper boundaries that account for the most challenging scenarios in terms of radiation protection.

The bunch divergences expected experimentally for proton or laser ion acceleration experiments (LION, HF) are, almost entirely, covered by the in–cave beam dumps and have hence been modelled using the scaling law in table 13.2 or worst case values. There are ions also emitted at larger angles, but due to their low energy fluence they would not be relevant for radiation protection. For laser–driven electron acceleration the generated particle beam is very directional, with a divergence of typically one mrad (Döpp et al. 2018). The angular pointing jitter is of a similar order and thus approximated using a fixed value (table 13.2).

For LION and HF, the simulated sources were mixed radiation fields (ions plus electrons). To simplify the simulations while allowing for detailed interpretation of the results, each initial component of a mixed radiation field was simulated separately and summed up in post processing to yield the total equivalent dose.

For the dosimetric evaluation of the results, the equivalent dose in units of pSv/primary particle was scored in a three dimensional Cartesian mesh (USRBIN scorer) with a voxel size of $10\,\mathrm{cm} \times 10\,\mathrm{cm} \times 10\,\mathrm{cm}$. To allow for a detailed investigation of different contributions, the total equivalent dose was decomposed to dose from neutrons, photons and, when deemed necessary, from electrons.

At critical locations, where elevated doses were expected, the dose equivalent was also scored in cylindrical air volumes of various radii and 5 cm thickness. The locations of these cylindrical scorers is indicated in figure 13.2, and 13.3, bearing the scorer's name. In turquoise are indicated scorers with 50 cm radius and in yellow with radii of 15 cm to 32.4 cm, depending on the size of the hole preceding them.

The radiation dose limits (see next section 13.2.5 for more details) are expressed in units of micro–Sievert per hour (μSv/h). The scored dose in pSv/primary needed therefore to be converted to μSv/h via a conversion factor. The conversion factor c was generally different for each simulation and source component and was calculated as follows:

$$c = N \cdot R \cdot U. \tag{13.1}$$

where N is the number of particles produced per laser shot, R the number of laser shots per hour (repetition rate) and $U = 10^{-6}$ is a dimensionless conversion factor from pSv to μSv.

13.2.5 Radiation limits

The CALA facility, in terms of radiation protection, can be regarded in three categories: exclusion, supervised and unclassified areas. The upper radiation dose rate limits were defined in agreement with the Bavarian Environment Agency (Bayerisches Landesamt für Umwelt).

In the CALA plan showed in figure 13.1, the north and west corridors and the southern laboratory are the only unclassified areas. The radiation dose rate limit in an unclassified area is $0.5\,\mu$Sv/h in order to ensure an accumulated dose of less than 1 mSv per year, assuming a 2000 h annual occupancy. A local maximum of about $2\,\mu$Sv/h is tolerable if monitored and/or not accessible.

All experimental caves, the air space above them and the experimental hall were categorised as supervised areas. The CALA design goal for supervised areas was $2.5\,\mu$Sv/h with a local maximum allowed to reach $7.5\,\mu$Sv/h. These locations will then be made inaccessible or marked with a warning sign.

When the laser is in full operation in a cave, then that particular cave is closed and classified as an exclusion area, while the other caves remain supervised areas. There is no upper dose rate limit specified in exclusion areas, as long as the dose rate limits in surrounding supervised or unclassified areas are not violated.

During the present trial operation of the CALA facility, continuous dose measurements are performed with passive thermoluminescent dosimeters (TLDs) on up to ten locations and evaluated monthly. Their values have to be reported quarterly to the Bavarian Environment Agency.

Thus, controlled areas with the duty to wear personal electronic dosimeters can be avoided. This might change in future operation because of activations and the resulting dose rate they might cause. However, this has not been observed yet in CALA but would cause no basic problems since the radiation would not be pulsed and dosimeters are available.

13.3 Results

Our simulations showed that, due to the high particle fluence ($\approx 10^{10}$particles/cm^2) in a cave during operation, the dose rate can exceed 1.5 kSv/h. The peak dose rate occurred, for each cave, in the respective beam dump, but dose rates higher than $10\,\mu$Sv/h were present in large sections of each cave during operation of the laser (e.g. figure 13.4).

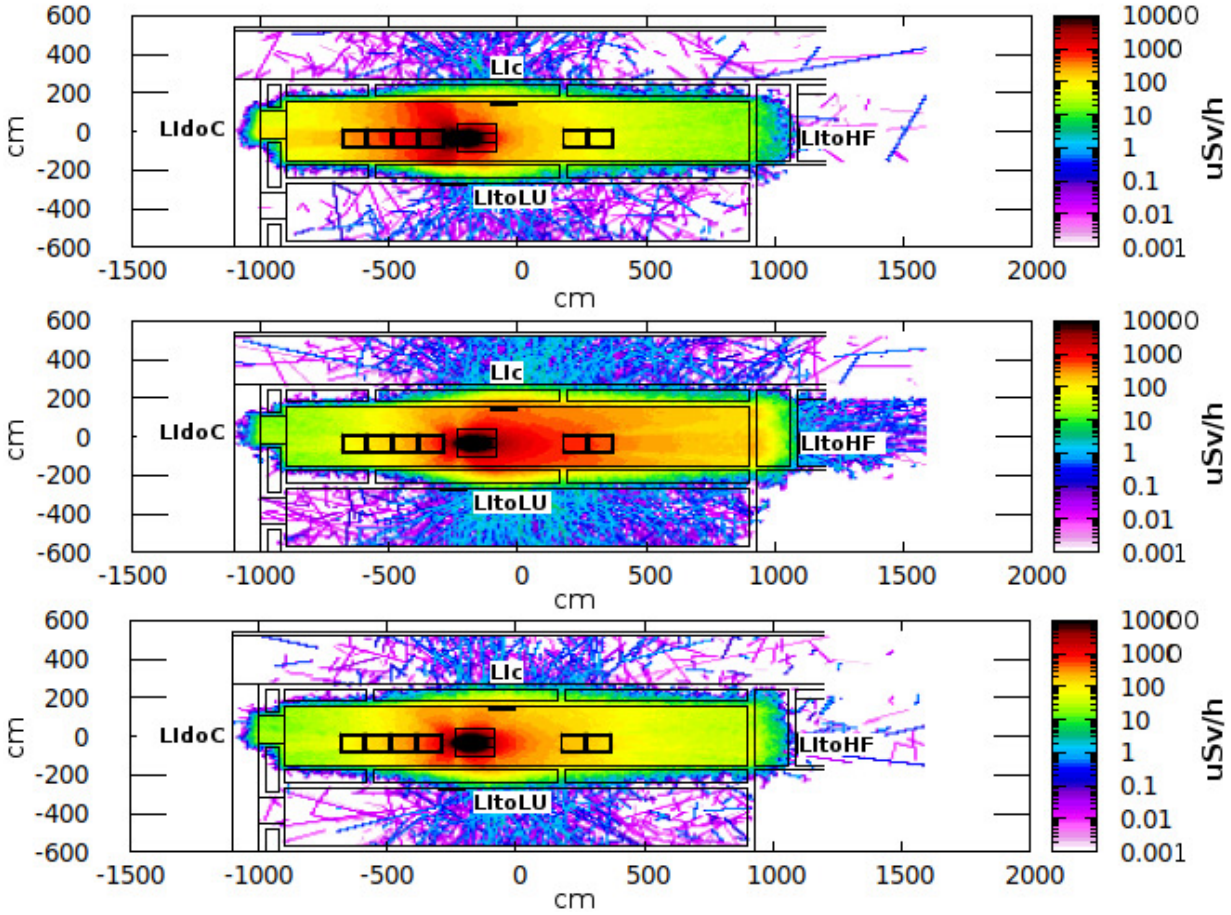

Figure 13.4: Dose rate distributions in and around the LION cave, plotted at the source height for different versions of the LION beam dump and for the ^{12}C ion beam component. (*top*) a version of the beam dump made entirely of concrete, (*middle*) the same beam dump volume filled with water and (*bottom*) the optimised beam dump, combining both water and concrete volumes. Some of the scoring locations relevant to LION are also indicated.

Depending on the cave and the beam dump geometry and location, different critical points were identified as indicated in figure 13.2 and 13.3. These critical points were in general located either behind open holes (vacuum pipes) or behind walls, but in positions where large fluences of particles were directed. The dose escaping a cave, during operation in that particular cave, and quantified at these critical locations, were summarised in tables 13.3 – 13.7. Dose rates above the dose rate limit are highlighted as bold numbers in the tables. For the experiments producing protons and heavier ions (LION and HF), the dose rate is listed as an electron beam component (same for all sources) and the respective proton or heavy ion beam component.

All estimated dose rates listed in tables 13.3 – 13.7 entail statistical uncertainties. These uncertainties were strongly dependent on the absolute dose rate and on the particle

fluence in the location under consideration. For calculated dose rates larger than 1 μSv/h, the statistical uncertainty ranged from 1% to 9%, with the mean error being 6%. For calculated dose rates between 0.1 μSv/h and 1 μSv/h the error ranged from 1% to 40%, with the mean error being 15% and errors above 20% associated uniquely to dose rates below 0.5 μSv/h.

13.3.1 LION

As summarised in table 13.3, two source types were simulated for the LION cave: ^{12}C ions and protons. The same electron beam component was assumed for each of the two source types. For each beam component, the major contributions to dose rates (by neutrons and photons) alongside with the total dose rate are listed. Finally, the overall total dose rate for a given source type is shown in the right most column.

In the vast majority of critical locations the dose from neutrons was at least one order of magnitude higher than that from photons, which was the next highest contribution. Dose imparted to these critical points by other particles was negligible.

For the proton source, the total dose rate at all critical locations was below the 0.5 μSv/h and 2.5 μSv/h limit, for unclassified and supervised areas, respectively. For the ^{12}C ion source, the dose rate exceeded the limits in two locations: opposite to the beam dump on the corridor (LIc) the dose rate was estimated to be 1.06 μSv/h (0.5 μSv/h limit) and on one location on the roof of the cave it was calculated to be 3.31 μSv/h (2.5 μSv/h limit).

The results in table 13.3 were obtained with the hybrid (water – concrete) beam dump geometry, which helped to reduce the dose in the north corridor and through the laser penetration holes. Exemplary results are shown in figure 13.4, comparing a full concrete beam dump (figure 13.4 – *top*), a full water beam dump of the same volume (figure 13.4 – *middle*) and the hybrid beam dump of the same volume (figure 13.4 – *bottom*).

13.3.2 HF

Table 13.4 lists the dose rate calculations for HF, at the critical points shown in figures 13.2 and 13.3. Three source types were simulated for the HF cave: protons, ^{12}C ions and ^{197}Au ions. The same electron beam component was assumed for each of the three source types. For each beam component, the major contributions to dose rates from neutrons and photons and the total dose rate are listed.

The proton source calculations showed that the dose rate limit of 0.5 μSv/h was mildly exceeded in two locations, at the corridor outside the HF cave (0.80 and 0.52 μSv/h, respectively). Behind one of the laser penetration holes leading to the north corridor (hole22), the dose rate was estimated to be 1.26 μSv/h, substantially higher than the 0.5 μSv/h, but well below 2.5 μSv/h. In the same locations the dose rate for the ^{12}C source was found to be 0.77, 0.74 and 0.52 μSv/h*h*. The dose rate outside of the HF cave for the ^{197}Au source has been calculated to be always below 0.5 μSv/h. In all cases the dose rate due to neutrons was at least an order of magnitude higher than due to photons.

Table 13.3: Dose rate calculations for LION, at critical locations indicated in figures 13.2 and 13.3. The upper half of the table lists results for the ^{12}C source, while the lower half presents the same for the proton source.

	Dose rate		Dose rate [μSv/h]						
	limit	Detector	Electron beam component			Carbon beam component			
	[μSv/h]	position	n	γ	Total	n	γ	Total	Total
	0.5	LIc	6.86×10^{-2}	1.07×10^{-3}	7.11×10^{-2}	$\mathbf{9.62\times10^{-1}}$	1.55×10^{-2}	$\mathbf{9.90\times10^{-1}}$	**1.06**
	0.5	hole1	1.76×10^{-2}	1.29×10^{-2}	3.25×10^{-2}	1.95×10^{-1}	9.10×10^{-3}	2.07×10^{-1}	2.40×10^{-1}
	0.5	hole8	8.01×10^{-2}	1.45×10^{-3}	8.21×10^{-2}	2.73×10^{-1}	4.10×10^{-3}	2.81×10^{-1}	3.63×10^{-1}
	0.5	hole15	1.63×10^{-3}	2.89×10^{-3}	4.53×10^{-3}	4.75×10^{-1}	8.22×10^{-3}	4.92×10^{-1}	4.97×10^{-1}
LION	2.5	LItoHF	6.39×10^{-4}	1.14×10^{-5}	6.55×10^{-4}	2.04×10^{-1}	2.60×10^{-3}	2.13×10^{-1}	2.13×10^{-1}
Carbon	2.5	LItoLU	8.54×10^{-2}	1.28×10^{-3}	8.82×10^{-2}	1.61	2.70×10^{-2}	1.65	1.73
	2.5	LIroof	2.10×10^{-1}	1.72×10^{-3}	2.13×10^{-1}	**3.03**	3.79×10^{-2}	**3.10**	**3.31**
	2.5	LIdoC	1.22×10^{-2}	1.38×10^{-4}	1.24×10^{-2}	6.37×10^{-3}	2.90×10^{-4}	6.66×10^{-3}	1.94×10^{-2}
	2.5	hole3	3.54×10^{-2}	2.74×10^{-2}	6.29×10^{-2}	2.32×10^{-1}	1.07×10^{-2}	2.44×10^{-1}	3.07×10^{-1}
	2.5	hole10	9.44×10^{-2}	5.85×10^{-3}	1.26×10^{-1}	2.87×10^{-1}	4.20×10^{-3}	4.00×10^{-1}	5.26×10^{-1}
	2.5	hole17	6.47×10^{-3}	1.87×10^{-4}	6.68×10^{-3}	4.96×10^{-1}	7.35×10^{-3}	5.08×10^{-1}	5.15×10^{-1}
			Electron beam component			Proton beam component			
	0.5	LIc	6.86×10^{-2}	1.07×10^{-3}	7.11×10^{-2}	1.75×10^{-2}	3.79×10^{-4}	1.80×10^{-2}	8.91×10^{-2}
	0.5	hole1	1.76×10^{-2}	1.29×10^{-2}	3.25×10^{-2}	7.43×10^{-2}	3.22×10^{-3}	7.94×10^{-2}	1.12×10^{-1}
	0.5	hole8	8.01×10^{-2}	1.45×10^{-3}	8.21×10^{-2}	2.72×10^{-2}	7.29×10^{-4}	2.85×10^{-2}	1.12×10^{-1}
	0.5	hole15	1.63×10^{-3}	2.89×10^{-3}	4.53×10^{-3}	1.13×10^{-2}	3.07×10^{-4}	1.17×10^{-2}	1.62×10^{-2}
LION	2.5	LItoHF	6.39×10^{-4}	1.14×10^{-5}	6.55×10^{-4}	6.25×10^{-4}	7.71×10^{-6}	6.41×10^{-4}	1.30×10^{-3}
Proton	2.5	LItoLU	8.54×10^{-2}	1.28×10^{-3}	8.82×10^{-2}	4.46×10^{-2}	7.32×10^{-4}	4.55×10^{-2}	1.34×10^{-1}
	2.5	LIroof	2.10×10^{-1}	1.72×10^{-3}	2.13×10^{-1}	8.52×10^{-2}	1.34×10^{-3}	8.72×10^{-2}	3.00×10^{-1}
	2.5	LIdoC	1.22×10^{-2}	1.38×10^{-4}	1.24×10^{-2}	6.34×10^{-4}	2.85×10^{-5}	6.62×10^{-4}	1.30×10^{-2}
	2.5	hole3	3.54×10^{-2}	2.74×10^{-2}	6.29×10^{-2}	6.82×10^{-2}	2.44×10^{-3}	7.39×10^{-2}	1.37×10^{-1}
	2.5	hole10	9.44×10^{-2}	5.85×10^{-3}	1.26×10^{-1}	1.09×10^{-1}	3.73×10^{-3}	1.14×10^{-1}	2.40×10^{-1}
	2.5	hole17	6.47×10^{-3}	1.87×10^{-4}	6.68×10^{-3}	5.28×10^{-3}	5.71×10^{-4}	6.13×10^{-3}	1.28×10^{-2}

13.3.3 LUX

The simulation results in table 13.5 summarise the dose rates at various positions outside the LUX cave for the two different electron spectra listed in table 13.2.

Since the primary source consists of highly penetrating electrons, the total dose rate is reported alongside with the individual contributions from neutrons, gammas and electrons.

For most scoring positions and for both the monoenergetic and the broad energy spectrum, the neutron dose rate was found to be the highest contribution to the total dose rate. The calculated dose rates from the broad spectrum electron source were mostly lower than the corresponding ones from the monoenergetic 5 GeV electron beam and always lower than the 2.5 μSv/h design limit.

For the 5 GeV monoenergetic electron source, at all positions, except for the critical location behind the channel of the central beam dump (LUtoHALL), the dose rate was kept below the 2.5 μSv/h design limit. At that particular location, the dose rate calculated was 6.95 μSv/h, which is still below the 7.5 μSv/h maximum local dose rate limit.

Table 13.4: Dose rate calculations for HF, at critical locations indicated in figures 13.2 and 13.3. The *upper third* of the table lists results for the proton source, the *middle third* for the ^{12}C source, while the *lower third* holds for the ^{197}Au source.

	Dose rate		Dose rate [μSv/h]						
	limit	Detector	Electron beam component			Proton beam component			
	[μSv/h]	position	n	γ	Total	n	γ	Total	Total
	0.5	hole22	2.83×10^{-1}	1.65×10^{-2}	3.07×10^{-1}	$\mathbf{9.01\times10^{-1}}$	1.32×10^{-2}	$\mathbf{9.56\times10^{-1}}$	**1.26**
	0.5	hole23	2.70×10^{-2}	5.83×10^{-3}	3.34×10^{-2}	1.16×10^{-1}	4.86×10^{-3}	1.22×10^{-1}	1.55×10^{-1}
HF	0.5	HFc1	1.70×10^{-1}	3.21×10^{-3}	1.76×10^{-1}	$\mathbf{6.12\times10^{-1}}$	1.10×10^{-2}	$\mathbf{6.27\times10^{-1}}$	$\mathbf{8.03\times10^{-1}}$
Proton	0.5	HFc2	4.11×10^{-1}	7.37×10^{-3}	4.26×10^{-1}	9.57×10^{-2}	1.06×10^{-3}	9.69×10^{-2}	$\mathbf{5.23\times10^{-1}}$
	0.5	HFtoHALL	8.22×10^{-2}	1.52×10^{-3}	8.58×10^{-2}	2.65×10^{-1}	5.76×10^{-3}	2.73×10^{-1}	3.59×10^{-1}
	0.5	HFtoLI	$\mathcal{O}(10^{-5})$	$\mathcal{O}(10^{-6})$	$\mathcal{O}(10^{-5})$	0.00	1.14×10^{-8}	1.30×10^{-8}	1.30×10^{-8}
	2.5	HFroof	2.95×10^{-1}	4.87×10^{-3}	3.54×10^{-1}	4.49×10^{-1}	4.84×10^{-3}	4.56×10^{-1}	8.10×10^{-1}
			Electron beam component			Carbon beam component			
	0.5	hole22	2.83×10^{-1}	1.65×10^{-2}	3.07×10^{-1}	2.02×10^{-1}	6.05×10^{-3}	2.10×10^{-1}	$\mathbf{5.17\times10^{-1}}$
	0.5	hole23	2.70×10^{-2}	5.83×10^{-3}	3.34×10^{-2}	1.49×10^{-1}	3.67×10^{-3}	1.56×10^{-1}	1.89×10^{-1}
HF	0.5	HFc1	1.70×10^{-1}	3.21×10^{-3}	1.76×10^{-1}	$\mathbf{5.70\times10^{-1}}$	1.09×10^{-2}	$\mathbf{5.91\times10^{-1}}$	$\mathbf{7.66\times10^{-1}}$
Carbon	0.5	HFc2	4.11×10^{-1}	7.37×10^{-3}	4.26×10^{-1}	2.98×10^{-1}	6.57×10^{-3}	3.15×10^{-1}	$\mathbf{7.41\times10^{-1}}$
	0.5	HFtoHALL	8.22×10^{-2}	1.52×10^{-3}	8.58×10^{-2}	3.65×10^{-1}	7.14×10^{-3}	3.83×10^{-1}	4.69×10^{-1}
	0.5	HFtoLI	$\mathcal{O}(10^{-5})$	$\mathcal{O}(10^{-6})$	$\mathcal{O}(10^{-5})$	1.34×10^{-5}	2.09×10^{-7}	1.36×10^{-5}	1.36×10^{-5}
	2.5	HFroof	2.95×10^{-1}	4.87×10^{-3}	3.54×10^{-1}	4.97×10^{-1}	5.05×10^{-3}	5.05×10^{-1}	8.59×10^{-1}
			Electron beam component			Gold beam component			
	0.5	hole22	2.83×10^{-1}	1.65×10^{-2}	3.07×10^{-1}		1.25×10^{-4}		3.07×10^{-1}
	0.5	hole23	2.70×10^{-2}	5.83×10^{-3}	3.34×10^{-2}		1.97×10^{-5}		3.34×10^{-2}
HF	0.5	HFc1	1.70×10^{-1}	3.21×10^{-3}	1.76×10^{-1}		1.29×10^{-6}		1.76×10^{-1}
Gold	0.5	HFc2	4.11×10^{-1}	7.37×10^{-3}	4.26×10^{-1}		8.19×10^{-6}		4.26×10^{-1}
	0.5	HFtoHALL	8.22×10^{-2}	1.52×10^{-3}	8.58×10^{-2}		$\mathcal{O}(10^{-6})$		8.58×10^{-2}
	0.5	HFtoLI	$\mathcal{O}(10^{-5})$	$\mathcal{O}(10^{-6})$	$\mathcal{O}(10^{-5})$		$\mathcal{O}(10^{-6})$		$\mathcal{O}(10^{-5})$
	2.5	HFroof	2.95×10^{-1}	4.87×10^{-3}	3.54×10^{-1}		$\mathcal{O}(10^{-4})$		3.54×10^{-1}

Table 13.5: Dose rate calculations for LUX, at critical locations indicated in figures 13.2 and 13.3. The upper half of the table lists results for the 5 GeV monoenergetic electron source, while the lower half holds for the broad spectrum electron source.

	Dose rate limit		Dose rate [μSv/h]			
	[μSv/h]	Detector position	n	γ	e^-	Total
	2.5	LUroof	7.64×10^{-1}	1.16×10^{-2}	4.27×10^{-3}	7.95×10^{-1}
	2.5	LUtoLI	1.45×10^{-1}	3.43×10^{-3}	0.00	1.49×10^{-1}
	2.5	LUtoET	8.51×10^{-1}	1.15×10^{-2}	8.43×10^{-3}	9.13×10^{-1}
	2.5	LUdoC	1.36×10^{-1}	9.72×10^{-4}	0.00	1.37×10^{-1}
	2.5	LUtoHF	1.69	2.60×10^{-2}	8.02×10^{-3}	1.82
LUX	2.5	LUtoHALL	7.49×10^{-1}	8.11×10^{-2}	**6.08**	**6.94**
5 GeV	2.5	hole2	7.35×10^{-2}	2.41×10^{-3}	0.00	7.59×10^{-2}
	2.5	hole5	7.45×10^{-2}	1.47×10^{-2}	2.79×10^{-3}	9.19×10^{-2}
	2.5	hole9	6.40×10^{-2}	4.94×10^{-3}	0.00	6.89×10^{-2}
	2.5	hole12	1.14×10^{-1}	1.15×10^{-2}	5.79×10^{-3}	1.31×10^{-1}
	2.5	hole16	3.71×10^{-1}	4.55×10^{-2}	7.73×10^{-3}	4.25×10^{-1}
	2.5	hole19	7.86×10^{-1}	4.64×10^{-2}	5.71×10^{-3}	8.42×10^{-1}
	2.5	LUroof	4.66×10^{-1}	5.60×10^{-3}	9.76×10^{-4}	4.72×10^{-1}
	2.5	LUtoLI	4.20×10^{-2}	1.25×10^{-3}	1.83×10^{-2}	6.51×10^{-2}
	2.5	LUtoET	3.88×10^{-1}	5.95×10^{-3}	0.00	3.94×10^{-1}
	2.5	LUdoC	7.04×10^{-2}	1.05×10^{-3}	3.36×10^{-6}	7.14×10^{-2}
	2.5	LUtoHF	6.29×10^{-1}	1.18×10^{-2}	0.00	6.87×10^{-1}
LUX	2.5	LUtoHALL	5.52×10^{-2}	2.76×10^{-2}	2.31	2.40
Broad	2.5	hole2	8.84×10^{-2}	3.97×10^{-3}	5.46×10^{-4}	9.45×10^{-2}
	2.5	hole5	1.49×10^{-1}	2.92×10^{-3}	2.59×10^{-3}	1.54×10^{-1}
	2.5	hole9	2.36×10^{-2}	4.30×10^{-3}	6.82×10^{-4}	2.86×10^{-2}
	2.5	hole12	8.17×10^{-2}	6.32×10^{-3}	4.51×10^{-3}	9.25×10^{-2}
	2.5	hole16	2.66×10^{-1}	1.72×10^{-2}	6.78×10^{-3}	2.90×10^{-1}
	2.5	hole19	3.05×10^{-1}	2.77×10^{-2}	8.63×10^{-3}	3.44×10^{-1}

13.3.4 ETTF

The simulation results in table 13.6 summarise the dose rates at various positions outside the ETTF cave for three experiments: the broad spectrum electron source with the spectrometer magnet in place and a 5 GeV monoenergetic electron source with and without the spectrometer magnet.

Since the primary source is the same as in LUX and consisting of highly penetrating electrons, the total dose rate is again reported alongside with the individual contributions from neutrons, gammas and electrons.

For the broad spectrum source, the dose rates in all scoring locations remained below the 2.5 μSv/h design limit. Out–of–cave doses from neutrons were in general found to be one order of magnitude higher than photon doses and two orders lower compared to electrons.

The lower energy electrons from the broad spectrum source bent downwards, at large angles after travelling in the spectrometer magnet, caused the highest total dose rate of 1.11 μSv/h to occur behind the nearest laser penetration hole (hole21), inside LUX. Similar to LUX, electrons are scattered through the 5.85 mm central beam dump channel (ETtoHALL1) and gave rise to a calculated dose rate of 1.04 μSv/h.

For the 5 GeV monoenergetic source scenario with the spectrometer magnet, the highest dose rate of 3.76 μSv/h was calculated to occur behind the central beam dump hole (ETtoHALL1), as depicted in figure 13.5. That was higher than the design limit of 2.5 μSv/h, but less than the maximum allowed local dose rate value of 7.5 μSv/h.

For that particular source and experimental cave, the highest simulated dose rate for CALA of 1.6 kSv/h was recorded inside the central beam dump. In this inaccessible location, the particle fluence locally was calculated to be up to 1.6×10^{10}particles/cm^2.

The monoenergetic 5 GeV beam without the spectrometer magnet presented a challenging scenario for radiation protection. This is highlighted in figure 13.6, where the dose rate at beam line level is shown for the case with (left) and without (right) spectrometer magnet.

The employed 50 cm lead bricks at both ends of the central beam dump channel were able to hinder direct electron penetration through the channel. The highest dose rates in the scenario were 2.61 μSv/h on the roof of the building (ETroof), 3.84 μSv/h in the SPECTRE cave (ETtoSP), 5.33 μSv/h and 6.83 μSv/h in the two of the eastern laser penetration holes (hole18 and hole21) to LUX and SPECTRE respectively.

Although the aforementioned dose rates exceeded the design limit of 2.5 μSv/h, they were below maximum local dose limit of 7.5 μSv/h.

Table 13.6: Dose rate calculations for ETTF, at critical locations indicated in figures 13.2 and 13.3. The upper third of the table lists results for the 5 GeV monoenergetic electron source, the middle third for the same source without magnet but with lead shielding, while the lower third lists the results for the broad spectrum electron source.

	Dose rate limit		Dose rate [μSv/h]			
	[μSv/h]	Detector position	n	γ	e^-	Total
	2.5	ETroof	7.85×10^{-1}	1.28×10^{-2}	3.87×10^{-2}	8.37×10^{-1}
	2.5	ETtoLU	2.46×10^{-1}	4.91×10^{-3}	3.45×10^{-3}	2.63×10^{-1}
	2.5	ETtoSP	7.58×10^{-1}	1.49×10^{-2}	3.21×10^{-3}	7.83×10^{-1}
	2.5	ETdoC	5.35×10^{-2}	5.03×10^{-4}	0.00	5.40×10^{-2}
	2.5	ETtoHALL1	1.10×10^{-1}	7.51×10^{-2}	**3.46**	**3.76**
ETTF	2.5	ETtoHALL2	6.34×10^{-1}	1.09×10^{-2}	1.29×10^{-2}	6.67×10^{-1}
5 GeV	2.5	hole4	3.83×10^{-2}	5.30×10^{-3}	5.66×10^{-3}	5.34×10^{-2}
	2.5	hole7	4.83×10^{-2}	4.48×10^{-3}	0.00	5.28×10^{-2}
	2.5	hole11	4.88×10^{-2}	5.78×10^{-3}	0.00	5.46×10^{-2}
	2.5	hole14	1.04×10^{-1}	4.87×10^{-3}	0.00	1.09×10^{-1}
	2.5	hole18	8.02×10^{-1}	5.09×10^{-2}	6.50×10^{-3}	8.66×10^{-1}
	2.5	hole21	5.34×10^{-1}	5.42×10^{-2}	8.47×10^{-3}	6.29×10^{-1}
	2.5	ETroof	**2.50**	8.47×10^{-2}	6.32×10^{-3}	**2.61**
	2.5	ETtoLU	2.07	4.80×10^{-2}	6.10×10^{-2}	2.20
	2.5	ETtoSP	**3.50**	8.69×10^{-2}	3.17×10^{-2}	**3.84**
	2.5	ETdoC	3.96×10^{-3}	8.92×10^{-5}	0.00	4.05×10^{-3}
	2.5	ETtoHALL1	3.95×10^{-1}	5.50×10^{-4}	0.00	4.11×10^{-1}
ETTF	2.5	ETtoHALL2	1.69×10^{-2}	1.29×10^{-4}	0.00	1.87×10^{-2}
5 GeV	2.5	hole4	2.63×10^{-1}	3.84×10^{-2}	0.00	3.02×10^{-1}
no Magnet	2.5	hole7	4.39×10^{-1}	3.49×10^{-2}	8.98×10^{-2}	5.64×10^{-1}
	2.5	hole11	2.57×10^{-1}	3.34×10^{-2}	1.19×10^{-3}	2.91×10^{-1}
	2.5	hole14	3.62×10^{-1}	4.21×10^{-2}	1.71×10^{-2}	4.22×10^{-1}
	2.5	hole18	**4.77**	3.42×10^{-1}	2.12×10^{-1}	**5.33**
	2.5	hole21	**5.86**	3.93×10^{-1}	5.43×10^{-1}	**6.83**
	2.5	ETroof	6.86×10^{-1}	9.94×10^{-3}	8.19×10^{-4}	7.00×10^{-1}
	2.5	ETtoLU	2.20×10^{-1}	3.80×10^{-3}	2.86×10^{-3}	2.28×10^{-1}
	2.5	ETtoSP	4.63×10^{-1}	1.03×10^{-2}	2.02×10^{-3}	4.86×10^{-1}
	2.5	ETdoC	2.92×10^{-2}	3.16×10^{-4}	0.00	2.95×10^{-2}
	2.5	ETtoHALL1	3.52×10^{-2}	5.01×10^{-2}	8.84×10^{-1}	1.04
ETTF	2.5	ETtoHALL2	2.27×10^{-1}	3.81×10^{-3}	2.83×10^{-3}	2.35×10^{-1}
Broad	2.5	hole4	4.51×10^{-2}	4.40×10^{-3}	4.68×10^{-3}	6.38×10^{-2}
	2.5	hole7	3.92×10^{-2}	8.15×10^{-3}	7.16×10^{-4}	4.80×10^{-2}
	2.5	hole11	6.41×10^{-2}	8.23×10^{-3}	1.04×10^{-3}	7.34×10^{-2}
	2.5	hole14	4.93×10^{-2}	5.30×10^{-3}	7.62×10^{-4}	5.53×10^{-2}
	2.5	hole18	7.01×10^{-1}	8.60×10^{-2}	2.35×10^{-2}	8.19×10^{-1}
	2.5	hole21	9.86×10^{-1}	7.78×10^{-2}	3.98×10^{-2}	1.11

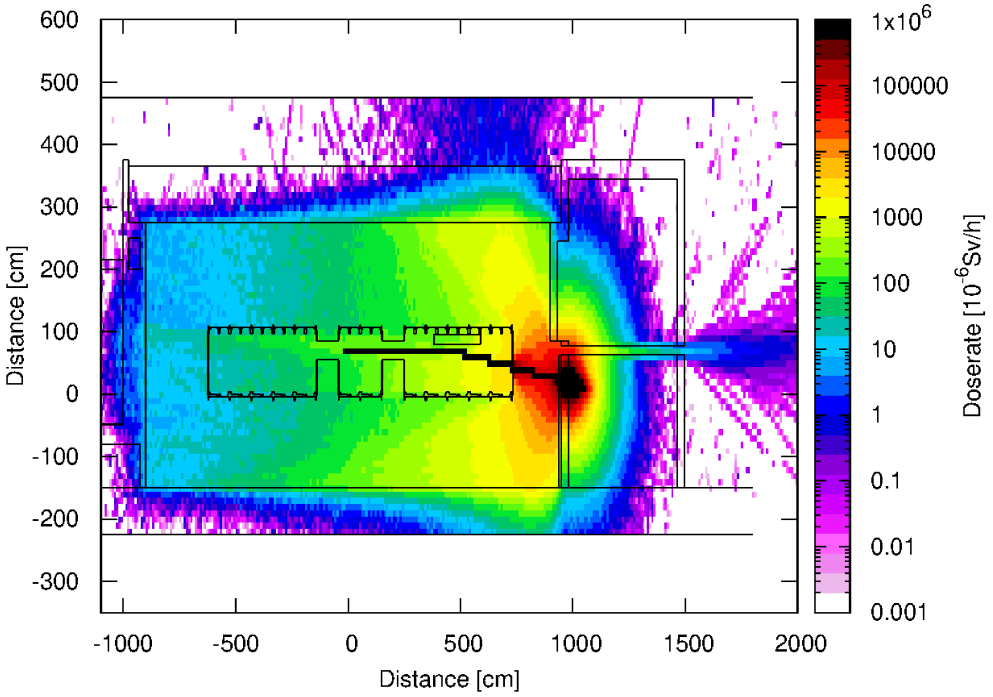

Figure 13.5: Dose rate in ETTF shown on a vertical plane, for a 5 GeV monoenergetic electron source intersected by a spectrometer magnet.

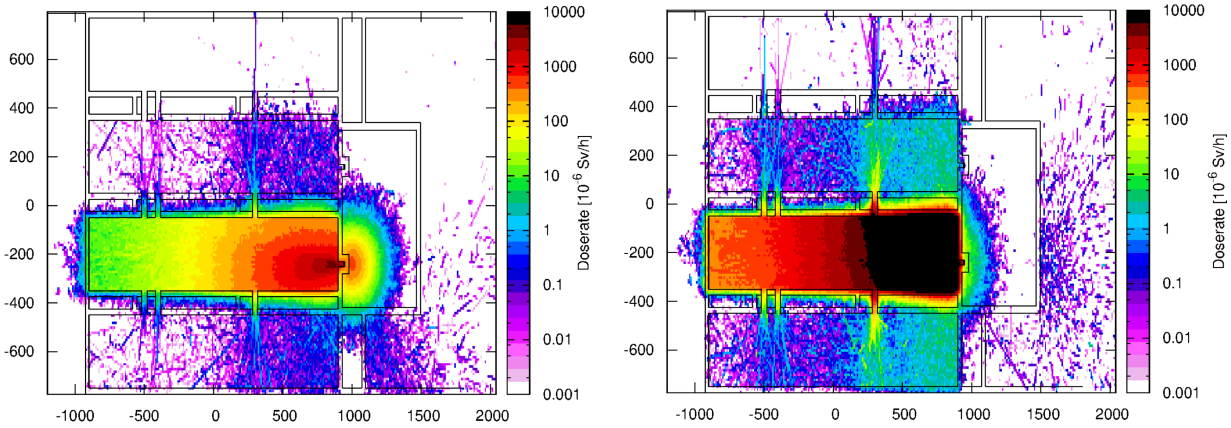

Figure 13.6: Dose rate distributions for the ETTF cave at the beam line level for 5 GeV monoenergetic electron source. (*Left*) normal setup with spectrometer magnet, (*right*) without spectrometer magnet. In the latter 50 cm lead were employed at both ends of the beam dump channel as shielding.

13.3.5 SPECTRE

For the SPECTRE cave, three source types were simulated (for details see table 13.2): a 500 MeV electron source at a repetition rate of 1 Hz. A broad spectrum electron source with energies up to 500 MeV, also generated with 1 Hz repetition frequency. Finally, 70 MeV monoenergetic electrons with 1 kHz repetition frequency.

Table 13.7 summarises the dose rates outside the cave in all scoring locations for the three source configurations. Except for one case, the calculated dose rates were at least one order of magnitude below the respective limits, either that of 2.5 μSv/h for supervised areas or that of 0.5 μSv/h for unclassified areas. The sole exception to this was in the critical location behind one of the laser penetration holes leading to ETTF (hole20), for the 70 MeV monoenergetic electrons with 1 kHz repetition frequency. There the dose rate was quantified to exceed the 2.5 μSv/h design limit (3.62 μSv/h), but still below the maximum allowed local dose rate of 7.5 μSv/h. Neutrons in general dominated again the out–of–cave doses.

Table 13.7: Dose rate calculations for SPECTRE, at critical locations indicated in figures 13.2 and 13.3. The upper third of the table lists results for the 70 MeV monoenergetic electron source at 1 kHz, the middle third for the 500 MeV monoenergetic electron source at 1 Hz, while the lower third represents the broad spectrum electron source at 1 Hz.

	Dose rate limit		Dose rate [μSv/h]			
	[μSv/h]	Detector position	n	γ	e^-	Total
	2.5	SProof	5.07×10^{-3}	2.33×10^{-3}	0.00	7.40×10^{-3}
	2.5	SPtoET	6.92×10^{-4}	8.54×10^{-4}	0.00	1.55×10^{-3}
	2.5	SPdoC	0.00	1.47×10^{-4}	0.00	1.47×10^{-4}
	2.5	SPtoHALL1	0.00	0.00	0.00	0.00
SPECTRE	2.5	SPtoHALL2	0.00	0.00	0.00	0.00
70 MeV	2.5	SPtoHALL3	0.00	0.00	0.00	0.00
	0.5	SPtoLAB	0.00	0.00	0.00	0.00
	2.5	hole6	2.03×10^{-2}	9.05×10^{-3}	0.00	2.94×10^{-2}
	2.5	hole13	1.54×10^{-2}	1.05×10^{-2}	0.00	2.59×10^{-2}
	2.5	hole20	1.30	1.39	9.25×10^{-1}	**3.62**
	2.5	SProof	4.82×10^{-2}	8.19×10^{-4}	1.15×10^{-4}	4.91×10^{-2}
	2.5	SPtoET	6.29×10^{-2}	1.34×10^{-3}	6.55×10^{-4}	6.67×10^{-2}
	2.5	SPdoC	5.89×10^{-3}	9.22×10^{-5}	0.00	5.98×10^{-3}
	2.5	SPtoHALL1	5.15×10^{-2}	1.42×10^{-3}	1.90×10^{-3}	5.48×10^{-2}
SPECTRE	2.5	SPtoHALL2	5.09×10^{-2}	6.96×10^{-4}	0.00	5.16×10^{-2}
500 MeV	2.5	SPtoHALL3	1.63×10^{-2}	2.02×10^{-4}	0.00	1.65×10^{-2}
	0.5	SPtoLAB	3.20×10^{-2}	6.03×10^{-4}	1.79×10^{-4}	3.28×10^{-2}
	2.5	hole6	1.53×10^{-2}	1.54×10^{-3}	4.12×10^{-4}	1.73×10^{-2}
	2.5	hole13	1.62×10^{-2}	1.91×10^{-3}	5.66×10^{-4}	1.87×10^{-2}
	2.5	hole20	1.43×10^{-1}	2.05×10^{-2}	1.06×10^{-2}	1.79×10^{-1}
	2.5	SProof	3.21×10^{-2}	5.85×10^{-4}	0.00	3.27×10^{-2}
	2.5	SPtoET	3.05×10^{-2}	6.61×10^{-4}	0.00	3.12×10^{-2}
	2.5	SPdoC	1.69×10^{-3}	9.00×10^{-5}	0.00	1.78×10^{-3}
	2.5	SPtoHALL1	1.47×10^{-2}	9.63×10^{-4}	0.00	1.56×10^{-2}
SPECTRE	2.5	SPtoHALL2	1.43×10^{-2}	1.07×10^{-3}	0.00	1.54×10^{-2}
Broad	2.5	SPtoHALL3	6.87×10^{-3}	8.69×10^{-5}	2.74×10^{-4}	7.23×10^{-3}
	0.5	SPtoLAB	1.29×10^{-2}	2.52×10^{-4}	0.00	1.34×10^{-2}
	2.5	hole6	9.22×10^{-3}	1.71×10^{-3}	7.53×10^{-4}	1.17×10^{-2}
	2.5	hole13	1.62×10^{-2}	2.16×10^{-3}	3.05×10^{-3}	2.14×10^{-2}
	2.5	hole20	2.01×10^{-1}	1.51×10^{-1}	3.77×10^{-1}	8.33×10^{-1}

13.4 Discussion

As presented in section 13.3 and detailed in tables 13.3 – 13.7, by employing adequate beam dumps matched to beam–divergence, magnets, passive shielding and laser pulse repetition limits, the simulated dose rates remained mostly below the design limits of the CALA facility ($0.5\,\mu$Sv/h for unclassified and $2.5\,\mu$Sv/h for supervised areas).

The highest dose areas in each cave were directly correlated to the location of the respective beam dump. For the LUX, ETTF and SPECTRE caves, the beam was terminated in the central beam dump, at the eastern end of each cave. This meant that the dose was higher in that part of these caves and in the experimental hall behind it. LION and HF contain individual beam dumps, approximately located at the center of these caves. Therefore, the high dose areas in these caves were estimated to be in their middle, which in its turn resulted in the highest dose rates outside these caves to occur behind the walls in the direct vicinity of the beam dumps.

Common to all caves, as expected, was the trend for high dose rates to be present behind the laser penetration holes that are near beam dumps. The locations behind the holes, however, are below the false floor, hence not readily accessible.

Secondary neutrons were in the majority of the cases the dominant contribution to dose rates calculated outside a cave in operation. The initial beam and charged secondaries were usually fully contained in the beam dumps, but secondary neutrons created in these volumes with highly directional fluences, gave rise to considerable doses outside the caves. A different situation was observed when beams were hitting the central beam dump and the x–ray extraction channels were open (LUX, ETTF, SPECTRE caves). In these cases, high energy electrons not adequately bent by the spectrometer magnets were the main component of the calculated out–of–cave high dose rate. The leakage of electrons through the x–ray extraction channels can be further increased due to scattering in the chamber walls intersecting the beam. Thicker walls significantly increase the angular distribution of the electron beam, allowing a larger part of it to go through the beam dump channels.

The high energy electrons produced in ETTF and SPECTRE yield a non–negligible fluence of muons (up to about $1000\,\text{muons/cm}^2$), approximately two orders of magnitude lower than the maximum neutron fluence. These muons are predominantly produced at shallow locations of the central beam dump. They display a broad energy spectrum with energies reaching up to a few GeV. Muons with such energies have a maximum range in water of $3 - 4$m, which means that they cannot escape the central 5.85 m thick concrete beam dump. Therefore, their contribution to the dose rate outside the caves is negligible.

The maximum quantified dose rate outside a cave in operation was $6.94\,\mu$Sv/h and was estimated for a monoenergetic 5 GeV electron source, behind the central beam dump of LUX (LUtoHALL). That location is one of the critical points of the CALA design. Despite the massive central beam dump, two channels through it are required in order to allow for the transport of x–ray beams from the experimental caves LUX and ETTF to the experimental hall to the east. Behind these two channels dose rates were expected to be high. The almost 3–fold excess of the $2.5\,\mu$Sv/h design limit for supervised areas was mainly attributed to high energy electrons escaping through one of the central beam dump

channels. The dose was below the maximum locally allowed dose rate of 7.5 μSv/h, which means that this particular location can be made temporarily inaccessible.

A similar dose rate (6.83 μSv/h) was calculated to leak through one of the laser penetration holes towards SPECTRE (hole21), in the case of the most demanding configuration in ETTF, that of a monoenergetic 5 GeV electron source without a spectrometer magnet. Due to the lack of magnetic deflection, almost the entire electron beam hit a small area of the central beam dump (the lead blocking the beam dump channel) and created a high radiation environment in a small section of the ETTF cave. The creation of secondary neutrons there was the primary dose rate component leaking through the laser penetration hole most proximal to the beam dump. As in the previous case, this dose rate was also below the maximum locally allowed rate of 7.5 μSv/h.

Excess of the dose rate limit for unclassified areas (0.5 μSv/h limit) was estimated for a few occasions concerning the LION and HF caves simulated operation. For LION the highest dose rate (1.06 μSv/h) occurred at the north corridor (LIc) in the ^{12}C source case and was mainly attributed to the directional neutron fluence escaping the hybrid beam dump. For HF, the respective excess dose rate of 1.06 μSv/h was calculated behind one of the laser penetration holes (hole22) for the proton source case. Again the main component of the dose rate originated from secondary neutrons, backscattered from the concrete beam dump. The fact that for HF, contrary to LION, it was the proton rather than the ^{12}C source the most demanding scenario in terms of radiation protection was due to the fact that the maximum ^{12}C ion energy in HF was assumed to be half of that in LION (for details see table 13.2).

Due to the ^{12}C–ion energy spectrum in LION, elevated dose rates, mostly stemming from secondary neutrons were calculated. This necessitated the design of an optimised beam dump for that cave. A full concrete beam dump of the same volume would stop the beam before penetrating deep enough and most of the produced secondaries would back–scatter (see figure 13.4 – *top*). A full water beam dump would stop the beam deeper in the beam dump, but the lower density and stopping power of water compared to concrete would still allow a large fraction of the generated secondaries to escape, especially towards the forward direction (see figure 13.4 – *middle*). The hybrid beam dump stops the beam approximately in the middle of its volume, containing more efficiently the secondaries by enclosing them in a water cavity surrounded by concrete (see figure 13.4 – *bottom*).

One of the main goals of the CALA radiation protection design was to allow caves and experimental areas neighbouring a cave, in which the laser is in operation, to remain accessible. This goal was in general achieved in our simulation study, with a few exceptions at a handful of locations (dose rates $> 2.5\,\mu$Sv/h): in the experimental hall, when LUX or ETTF (only for 5 GeV monoenergetic) were in operation, due to the electron leakage through the central beam dump channels. In SPECTRE and LUX when ETTF was in operation with 5 GeV monoenergetic electrons without spectrometer magnet due to neutron dose. In ETTF when SPECTRE was in operation with 70 MeV at the high repetition rate of 1 kHz, again due to neutrons.

In this study we have calculated the dose rate from each individual cave being in operation, without assuming cumulative effects due to operation of multiple caves in parallel.

Four out of the five CALA caves (LION, HF, LUX and ETTF) are seeded by the ATLAS–3000 laser and will not be operated simultaneously. ETTF and SPECTRE (seeded by the PSF–pro laser) can in principle be operated at the same time, which means that their cumulative dose rate should be taken into account. A critical location which would be accessible during a simultaneous operation of ETTF and SPECTRE is the vicinity of ETtoHALL2 and SPtoHALL1. This is the area with the highest overlap of dose leakage originated by ETTF and SPECTRE. At this location, the cumulative dose rate for the worst case scenario sources (ETTF 5 GeV monoenergetic, SPECTRE 500 MeV monoenergetic) would be $7.22 \times 10^{-1} \mu$Sv/h ($6.67 \times 10^{-1} \mu$Sv/h + $5.48 \times 10^{-2} \mu$Sv/h), which is much lower than the required dose rate limit.

Additional, beyond the nominal, material and equipment intersecting the electron beam might significantly change the source characteristics. This in its turn can have a large impact on the dose rate outside the caves. An extreme example of such a scenario is the 5 GeV ETTF electron beam passing through the thick 2.5 cm vacuum chamber walls instead of the 3 mm extraction window. This was estimated to increase the dose rate in the critical location ETtoHALL1 by up to a factor of 10. Therefore, constant monitoring of the dose rate at multiple locations outside the caves is envisaged.

There are a few factors that justify the design choice of the facility without a maze at the entrance to each cavern, maximising cavern space. The first is the fact that the use of thick concrete doors mitigates the need for a maze (in most radiation treatment facilities, the maze makes up for the shielding required due to the use of glass and thin metal doors). In addition, for all experimental caverns except for HF, the entrance is located at 180° with respect to the laser propagation direction towards the target. This is the direction with the lowest particle fluence, thus resulting to a low dose rate. Furthermore, the source parameters used in this MC study and yielding in some cases dose rates near the allowed limits, are rather optimistic (worst case scenario for shielding). Especially for biomedical applications, higher energies/fluences are not necessary.

Finally, in the unlikely case that the laser and targetry enable us to reach higher energies/fluences in the far future of CALA, decreasing the repetition rate for some particular experiments/sources would allow us to remain within the radiation protection limits.

13.5 Conclusion

Monte Carlo simulations were used to design and validate the radiation protection setup of the multi–purpose laser–particle acceleration facility CALA. Our study showed that dose from secondary neutrons was the major contribution in dose rates calculated outside experimental caves in operation. In addition to that, any holes or extraction channels would allow parts of the beam to escape the caves, also resulting in high dose rates. The presented design achieved the goal of dose rates below 7.5 μSv/h outside any cave in operation. The present trial operation of CALA is used to assure that the provisions concerning the dose limits are met.

Acknowledgments

The work presented in this documents was supported by the German research foundation (DFG) Cluster of Excellence Munich Centre for Advanced Photonics (EXC 158). We acknowledge the work performed by TÜV SÜD Nuclear Technologies, Warrington, UK as the starting point for our studies and Florian Saran and Klaus Wirgler for help with the preparation of technical drawings.

Chapter 14

Summary, conclusions and future outlook

In this final chapter, the main findings of this manuscript are briefly presented. Emphasis is put on the limitations or open questions left in each study and, wherever possible, directions for future investigations are outlined.

Imaging for proton treatment planning

The current clinical practice for obtaining the relative stopping power (RSP) patient maps required for proton treatment planning is based on single energy x–ray CT. The linear attenuation coefficient map is converted to RSP via a semi–empirical calibration, resulting in large errors (up to 3.5%) for some tissue types. Dual energy x–ray CT (DECT) and proton CT (pCT) both outperform single energy x–ray CT in terms of RSP accuracy, and are therefore considered as candidates for imaging for proton treatment planning. While there have been indirect comparisons between the two modalities, often in the simulation level and excluding real detector effects for pCT, the first direct experimental comparison of the two was presented in this manuscript. For that purpose, an experimental campaign was organized in Munich and Chicago, utilizing the most advanced preclinical pCT scanner and a commercial DECT scanner.

Image artifacts of an amplitude exceeding 2% in RSP were found to deteriorate the overall RSP accuracy achievable with pCT. By establishing an improved scanner calibration algorithm, with the aid of MC simulations, the overall pCT RSP accuracy improved from 0.87% to 0.55%, approaching the RSP accuracy achievable with an idealized pCT scanner (0.17%). The experimental pCT RSP accuracy was equivalent to that reached by a commercially available state–of–the–art DECT scanner (0.67%).

The fact that a prototype pCT scanner already achieved comparable or better RSP accuracy with a state–of–the–art commercial DECT scanner highlights the potential of the former for proton treatment planning. Improvements of the scanner design might push the RSP accuracy limit closer to that of ideal simulations (0.1-0.2%). The work with

the phase II scanner prototype and its five–stage energy detector showed that avoiding stopping protos near stage interfaces could improve image quality. This could for example be realized with the same scanner design by modulating the energy of the initial beam when scanning the object, in such a way that protons reaching the detector would stop away from stage boundaries. That approach was recently demonstrated experimentally by Dickmann et al. (2021)(Dedes–senior author). Such a solution would require though either an approximate prior knowledge of the object or image acquisition at two energies, thus twice the imaging dose. A dedicated scanner design operating exclusively with energy modulation was constructed by the ProtonVDA company and is currently tested (Sarosiek et al. 2021). Alternatively, optimization of the calibration procedure could also result in an increase of the scanner performance. An empirical artifact correction calibration adapted from x–ray CT (Kachelrieß et al. 2006) was applied to pCT experimental data and resulted in significant artifact reduction and RSP accuracy of 0.46% (Dedes–senior author, under review).

With a sub–1% RSP accuracy, DECT is a commercially available imaging solution that can used for proton treatment planning. In addition to the RSP accuracy, it inherently achieves better spatial resolution compared to proton imaging. A drawback of DECT is the increased dose to the patient. A factor of 5–10 higher is not negligible for ionizing radiation imaging dose. Furthermore, pCT images can be easily acquired with the patient placed at the treatment position, thus minimizing positioning uncertainties and in conjunction to low dose, be utilized for daily image guidance and plan adaptation. Finally, although RSP accuracy was quantified in this study, a full dosimetric evaluation that will show how the image quality achieved with the different modalities translates to clinical proton treatment plans should be performed. In a first step, Meyer et al. (2019) have quantified this in a simulation study assuming ideal detection systems and such investigations should be extended to account for realistic scanners.

Proof–of–concept FMpCT

The trade–off for decreasing image noise is increasing imaging dose, the latter being a crucial aspect for imaging modalities employing ionizing radiation. For particle therapy image guidance and plan adaptation though, only a fraction of the patient anatomy is required to be imaged accurately. A paradigm is the need for accurate imaging of the region which the therapeutic particle beam will traverse to reach the tumor, whereas the rest of the patient is to be spared as much as possible from receiving either therapeutic or imaging dose. To circumvent this trade–off, images of spatially varying quality can be acquired and reconstructed. This is the basic concept behind fluence modulated proton CT (FMpCT), where the fluence of the proton beam is modulated within every projection.

A proof–of–concept study employed ideal simulations (no detector effects), in which proton pencil beams were used in order to modulate the fluence within a projection. A very simple fluence modulation algorithm considered either an intersection of the central beam axis with a region of interest (ROI) – thus no pencil beam fluence modulation, called full fluence (FF) – or no intersection with that ROI resulting in pencil beam fluence

modulation with a fluence modulation factor (FMF) with respect to FF. Images were reconstructed by means of a filtered backprojection (FBP) algorithm. It was confirmed that FMpCT is feasible and that in general the image quality in the ROI was not affected by the high noise outside. For simple homogeneous objects, FMpCT retained the FF image quality (better than 0.3% in RSP accuracy and 1.2% in noise) for FMF down to 5% of the FF, while in uniform low fluence images noise increased considerably. Similar results were observed for simulations on patient anatomies, with FMpCT retaining the same noise and RSP accuracy as the FF down to 5% of FF, while uniform fluence image noise increases rapidly with decreasing fluence. Calculating treatment plans on FMpCT resulted in proton range which was in very good agreement with that calculate on FF images. Supporting the concept of FMpCT, significant imaging dose savings were quantified in simulations, ranging from 30% to about 60% outside of the ROI, where healthy tissue sparing is desired. Emulating fluence modulation on previously acquired experimental data of a pediatric head phantom, by discarding protons, confirmed the hypothesis FMpCT. It was also observed that the lowest level of FMF, below which RSP accuracy deteriorates significantly rose from 1% in ideal simulation to 30% emulated FMpCT on real data.

The concept was demonstrated experimentally using the pencil beam scanning capability of a clinical proton facility. For a cylindrical uniform phantom FMF of 50% and 20% was probed. Inside the central ROI, FMpCT and uniform fluence images achieved a sub–1% RSP accuracy and as planned, FMpCT noise remained at the level of FF image noise, in contrast to uniform 50% and 20% images, whose noise in the ROI increased significantly.

The simulation study and the experimental demonstration of FMpCT confirmed the postulated potential of the technique for preserving high image quality in a specific region of the image, while decreasing imaging dose outside. The simplicity of the fluence modulation algorithm which took into account intersection of the ROI by the beams in a binary fashion was a limiting factor that would not allow for more conformal shaping of the noise and dose inside and outside of the ROI, respectively. Furthermore, in this forward approach, the noise level in the ROI could not be a–priori predicted or prescribed. This would necessitate a fluence optimization algorithm and detailed theory of pCT noise, both subsequently developed. Specific limitations of the FMpCT concept were outlined. When employing a weak modulation (FMF>50%) the gain compared to uniform fluence in terms of preserving noise in the ROI and reducing dose outside of it would be low. This limitation would be aggravated by a large sized ROI. At the opposite extreme, strong modulation (FMF<30%) resulted in deterioration of the image quality in the ROI, when comparing FMpCT and FF images.

Image noise reconstruction for pCT

The quantification of pCT image noise can be estimated in a straightforward manner from the RSP variance in homogeneous regions of an image. The drawback of this simple method is that it is constrained to small regions due to the spatial distribution of noise in pCT and it is of limited use in the case of highly heterogeneous objects, such as human patients. Therefore, a voxelwise pCT noise map is necessary in order to fully capture all

noise patterns, allowing for prescribing and achieving spatially varying image quality. A voxelwise pCT noise map could be obtained by n–realizations of an image, but this would require either time consuming simulations of the imaging process based on a prior image, or repeated image acquisitions, thus increasing the imaging dose to the patient. An elegant solution to this problem is the direct reconstruction of noise, from the same data acquired for a single pCT RSP image.

A 2D pCT noise reconstruction algorithm was formulated, yielding the voxelwise RSP variance maps. Due to the linearity of the problem in FBP reconstruction, RSP variance images are reconstructed from WEPL variance projections, in analogy to the RSP image reconstructed from WEPL projections. In the projection level, protons are binned in pixels. For RSP reconstruction, the mean WEPL of all the protons in every projection pixel is determined. For variance reconstruction, the WEPL variance in every projection pixel needs to be determined. The WEPL variance (σ^2_{WEPL}) in a projection pixel was found to be proportional to the variance of the energy distribution (σ^2_E) of the protons binned in a projection pixel and anti–proportional to the number of protons (N) in that projection pixel. Given that for a fixed initial proton beam energy σ^2_E in a projection pixel determined by the object, the number of protons in a projection pixel is the only free parameter allowing for controlling the σ^2_{WEPL}. One of the main findings of the 2D pCT noise formalism is that multiple Coulomb scattering plays an important role in noise formation. Protons binned in the same projection pixel can previously traverse completely different paths in the object due to multiple Coulomb scattering. This results into increased σ^2_E and thus increased noise near heterogeneities and at the object boundaries. This effect was found to be reduced when taking into account multiple Coulomb scattering in the image reconstruction. Subsequently, a realistic MC simulation of the scanner was used in order to further validate the 2D noise reconstruction. Furthermore, by enhancing the simulation platform to account for non–linearities in the detection process and using a realistic beam model derived from experimental data, noise reconstructed from experimental and simulated pCT images were found to be in good agreement (less than 10%), providing a reliable tool for further pCT studies. Using it allowed for disentangling contributions to noise. In addition to the highly object dependent noise from multiple Coulomb scattering, the next most important contribution to noise came from energy straggling in the object and in the detector. For the phase II pCT scanner the sum of energy straggling in the object and in the detector is almost constant and object independent, as the less material a proton traverses in the object, the longer it travels in the energy detector until it stops. The energy spread of the initial proton beam also adds a constant amount of noise.

From the conclusion that noise in pCT is highly object dependent, an obvious limitation that arises is that it is not possible to assume generic patient models for the prediction of noise. In order to acquire FMpCT images optimized for spatially varying image quality, prior knowledge of the object is necessary. This can come from a previous pCT image. A convenient alternative would be a diagnostic x–ray CT image that can be used as an input to the pCT simulation platform, which was shown to reproduce accurately noise in experimental pCT images. Furthermore, the formulation of a robust pCT noise reconstruction unveiled that previous models capable of predicting noise at the center of a homogeneous

object significantly underestimate noise in clinically relevant geometries. Therefore, the dose efficiency (imaging dose for a given noise level) advantage of pCT over x–ray CT might not be as large as it was assumed in the past.

Noise optimized pCT

A further step towards optimized FMpCT is the capability to achieve a prescribed spatially varying image quality. It necessitates determining a fluence pattern for each acquisition angle, such that the resulting pCT reconstruction achieves a given image variance target.

The validated pCT MC simulation platform was used in order to develop a variance prescription algorithm. Finding relative modulation factors for each pencil beam such that the summed fluence pattern results in a prescribed image variance map is a computationally expensive problem, generally requiring alternating between the reconstructed image domain and the projection domain. The proposed image variance prescription method was formulated as an optimization algorithm. For a given image variance target, the variance projections were are found via forward projection. The set of solutions is large and constraints such as non–negativity have to be imposed. The variance image is reconstructed and the difference to the variance target is derived via forward projection and added to the variance projections of the previous iteration. Using this algorithm it was shown that pCT images with a contrast of four in variance could be achieved for homogeneous and anthropomorphic phantoms, resulting in dose savings ranging from 25.7% to 40.5% outside the low variance ROI. On the other hand, an increase of the dose of 9.2% to 19.2% was observed in the ROI. This can be considered as a negligible increase, considering that this effect occurs in the area where subsequently a therapeutic amount of dose will be deposited. The methodology was demonstrated experimentally. Fluence patterns optimized based on MC simulations were delivered in real pCT scans. The prescribed variance contrast of four was achieved also in experimental scans, showcasing the applicability of the whole workflow.

One significant limitation of the presented algorithm was that it does not explicitly account for dose reduction. Dose savings were achieved indirectly by prescribing higher noise, but a more controlled and direct dose prescription and optimization would be desirable. This was achieved in a subsequent study, with the formulation of a joint noise and dose optimization (Dickmann et al. 2020a). Furthermore, there is a large set of solutions in the projection level, with only a subset being realizable, resulting in a few constraints: the projection noise cannot be negative, there should be no projection pixels with zero protons for the FBP reconstruction to work and finally the number of protons should be finite. All these constraints limit the achievable variance contrast. In the experimental demonstration a slight degradation of the RSP accuracy was observed for the case of FMpCT, but as the ROI contained only two inserts the RSP accuracy needs to be more carefully quantified in the future. In general RSP accuracy was always at the level of 1% or better. In addition, as fluence patterns are optimized in the (sub–)mm scale, the whole process is sensitive to misalignments and beam size changes of that order. Finally, for technical reasons the interfacing with the beam delivery system allowed only step–and–shoot image acquisition.

This led to a tenfold increase of the image acquisition time, for the same beam–on time. This problem can be circumvented by using a continuous rotation of the object and continuous image acquisition, which would require synchronising the rotation stage with the fluence patterns to be delivered.

Biological uncertainties in proton treatment planning

In proton radiation therapy, a constant relative biological effectiveness (RBE) of 1.1 with respect to the reference photon radiation is currently assumed. Biological experiments show that the RBE is not constant throughout therapeutic proton fields and several models have been proposed to calculate it. Many of these models predict the RBE based on the delivered dose D, the dose weighted linear energy transfer $\mathrm{LET_d}$, the ratio of the coefficients of the linear quadratic model of describing cell survival for photon irradiation $(\alpha/\beta)_x$ and a free parameter q used for fitting the $\mathrm{LET_d}$ proton linear quadratic model coefficients to experimental cell survival data. Both $(\alpha/\beta)_x$ and q determination entails large uncertainties.

In this manuscript a study was presented in which these uncertainties for a published biological model were quantified and their impact on proton treatment plans (TPs) were explored. For two patients cases, intensity modulated proton therapy (IMPT) TPs were optimized and simulated assuming constant RBE and the dose was recalculated using the variable RBE model. For the nasopharyngeal patient, differences in dose in the tumor between the constant and variable RBE model was within the model uncertainties, but in two out of the three considered healthy organs at risk (OAR) the dose in the constant RBE model was considerably underestimated. For the prostate patient, the dose was found to be significantly different between the constant and variable RBE models in the tumor and in one of the two OARs considered. The findings also suggested that the uncertainty on both $(\alpha/\beta)_x$ and q can be important, depending on the patient, and that they should both taken into account when estimating non–constant RBE. Subsequently, a direct dose optimization accounting for uncertainties of the variable RBE model was performed and resulted in reducing dose in the OARs while maintaining tumor coverage with respect to the constant RBE plans.

Although uncertainty on q parameter could be reduced by fitting the models on a broader range of experimental data, the uncertainty on $(\alpha/\beta)_x$ for different organs and tumors in different patients is unlikely to be reduced in the foreseeable future. Therefore, robust TP optimization taking into account these uncertainties is beneficial. One limitation of the presented work is that it was performed on only two patients. A large and more diverse cohort would be necessary in order to reach more definitive conclusions. Furthermore, $\mathrm{LET_d}$ which is one of the parameters of the RBE model can be different for very similar dose distributions. This means that results might differ when the same therapeutic dose is delivered by a different configuration of treatment fields. Nevertheless the $\mathrm{LET_d}$ enters the RBE formula in the square root and consequently small $\mathrm{LET_d}$ variations cause even smaller RBE variations. The baseline RBE weighted dose strongly depends on the exact RBE functional dependence to dose, as described by a specific model. Therefore,

any conclusions about the agreement or disagreement between the constant RBE dose and the RBE weighted dose can vary when different models are used. Finally, the basis of the presented comparison is the well established linear quadratic model. Although this is a valid basis for describing cell survival determining the effect of radiation in the tumor, it does not account for complications and second cancer risks in healthy organs.

Prompt–gamma range monitoring

The finite range of protons and heavy ions in matter results in steep dose gradients which are favorable for use of radiation therapy. The advantage of highly conformal dose distributions and sparing of surrounding healthy tissues comes with the drawback of high sensitivity to range uncertainties. To overcome this problem, monitoring of the ion range in the patient is deemed necessary. As the therapeutic beam does not exit the patient, range monitoring has to rely on secondary radiation, induced by the therapeutic beam. Prompt–gamma (PG) emission, which occurs almost instantaneously with the proton or heavy ion beam interacting with the patient is an excelled candidate for in–vivo, real–time range monitoring.

MC simulations are used extensively for estimating the applicability of PG range monitoring and will be likely used for comparing the measured PG distribution in a patient to the expected one according to the treatment plan. Their accuracy in modelling the PG emission and detection is therefore of high importance. It has been observed that GEANT4 overestimates the production of prompt–gammas by 80%–180% for ^{12}C beams of therapeutic energies. In the work presented in this manuscript the focus was put on the assessment and improvement of the QMD inelastic nuclear interaction model for a better description of prompt–gamma emission. One of the free parameters of the QMD model (the nucleon wave packet width L) which was identified to strongly affect prompt–gamma emission, was initially optimized for heavy nuclei not relevant to particle therapy. Benchmarking the QMD model against binding energy per nucleon and nuclear density data, an optimal value for this parameter was found. Using that value, QMD produces lighter fragments of lower binding energy, which resulted in less frequent de–excitation via photon emission. In addition, the optimized L parameter value resulted in an overall improvement of the prediction of charged secondary angular yields, benchmarked against experimental data. With these settings, the discrepancy in predicting measured prompt–gamma emission reduced from 86%–165% to a few % up to 30% for all cases except for around the Bragg peak in the case of high energy ^{12}C beams where it was reduced from 180% down to 70%.

The remaining discrepancies in the prompt–gamma emission originates from the nuclear reactions induced by the secondaries lighter fragments, mostly protons. This is why they are concentrated around and beyond the Bragg peak. They cannot therefore be improved from the optimization of the QMD model which only handles nucleus–nucleus reactions. Furthermore, what should also be investigated in the future is a more holistic optimization taking into account other particle therapy related observables such as β^+ emitters.

In an effort to systematically investigate the accuracy and precision of range detection from prompt–gamma emission profiles, a set of five prostate cancer patient inten-

sity modulated proton therapy (IMPT) treatments were simulated. For each patient, the treatment was planned and simulated on an initial CT and resimulated in two subsequent CTs containing realistic interfractional anatomical changes. Prompt–gamma depth profiles were also simulated in order to detect spot–by–spot variations due to the aforementioned anatomical changes. Two methods, shifting and fitting, were applied on prompt–gamma depth profiles in order to identify range shifts from one CT to another due to anatomical changes. These prompt–gamma deduced range shifts were compared to range shifts deduced from the dose.

The accuracy with which range shifts in could be detected in clinical relevant anatomies and treatment plans was well below 1 mm and the precision was within ±2 mm, estimated over 1738 highest intensity spots. These results confirmed the potential of PG proton–range monitoring in clinical scenarios. To achieve this level of accuracy and precision a method of identification of distorted spots based on PG metrics was developed. The method discarded approximately 10% of the spots in anatomical locations where severe profile distortions were caused, impeding a reliable range shift identification.

In order to reduce computation time, this study was performed on the emitted prompt–gamma profiles and not on the profiles of the prompt–gammas exiting the patient. The emission profiles are idealized as photons do not cross the patient and display much sharper rise and fall–off. Nevertheless, the two type of profiles were compared and their differences in terms of deduced range were found to be sub-mm, which was also confirmed by a subsequent study (Tian et al. 2018). Furthermore, no specific detection system was simulated. This means that the conclusions drawn from this study could potentially be affected by detection effects such as efficiency, energy and spatial resolution, geometrical coverage and others. A follow up prompt–gamma range monitoring study based on a real detection system (Tian et al. 2021) confirmed the validity of the above mentioned conclusions. The simulated proton treatment plans were constrained to the pelvic region, but no significant deterioration of the quantified performance was observed for head and neck cancer patients (Tian et al. 2018, Tian et al. 2020). Finally, a minimum number of emitted prompt–gammas per spot, meaning a lower dose limit per spot, is required in order to achieve good sensitivity. It has been shown (Tian et al. 2018, Tian et al. 2020, Tian et al. 2021), this can be achieved by creating treatment plans that take the prompt–gamma monitoring aspect into account while retaining their dosimetric quality.

Radiation protection for laser–based ion accelerators

Particle therapy relies on compact, reliable and cost effective accelerators for the production of therapeutic beams. In particular for proton therapy the accelerator work horse is the cyclotron. Conventional accelerators convert the high voltage pulses from the modulator into pulsed radio frequency (RF) energy. The RF pulses are the sent to the accelerating structure to setup an electric field which is used for charged particle acceleration. The strength of this field is in the order of MV/cm. Alternatively, tens or even hundreds GV/cm can be achieved when focusing high–power short–pulsed laser on specific solid state targets. Laser–based ion and electron acceleration for biomedical application is explored in the

Centre for Advanced Laser Applications (CALA). 2.5 PW laser pulses with 25 fs duration yield mixed particle–species of high intensity, high energy and pulsed, thus posing new challenges compared to conventional radiation protection. The radiation protection aspect of CALA was investigated in a simulation study, assuming biomedically relevant beams of electrons ($< 5\,\mathrm{GeV}$), protons ($< 200\,\mathrm{MeV}$), $^{12}\mathrm{C}$ ($< 400\,\mathrm{MeV/u}$) and $^{197}\mathrm{Au}$ ($< 10\,\mathrm{MeV/u}$) ions. The goal was to achieve the radiation protection design specifications ($< 0.5\,\mu\mathrm{Sv/h}$ for unclassified areas, $< 2.5\,\mu\mathrm{Sv/h}$ for supervised areas, $< 7.5\,\mu\mathrm{Sv/h}$ maximum local dose rate).

By employing adequate beam dumps matched to beam–divergence, magnets, passive shielding and laser pulse repetition limits, the simulated dose rates remained mostly below the design limits of the CALA facility. In the majority of the cases, secondary neutrons were the dominant contribution to dose rates calculated outside a cave in operation, while the initial beam and charged secondaries were usually fully contained in the beam dumps. High dose rates outside the caves were observed behind holes envisaged for guiding the laser beam in the caves or for extracting produced x–rays to neighbouring experimental halls. The high energy $^{12}\mathrm{C}$ beam produced in one of the experimental halls necessitated the design of an optimised beam dump, consisting of a central water compartment surrounded by concrete. This optimized geometry results into the beam stopping deep into the water compartment with minimal backscattering, while the secondaries which are then directed mostly forward are essentially surrounded by the concrete compartment.

This radiation protection study has demonstrated a design adequate for beams intended for biomedical applications. For different applications requiring higher energies and fluxes at the same repetition rate, a revised design would be required. Nevertheless, such beam characteristics are not expected to be achieved in the near future. Another aspect that emerged from that work is that the typical maze design for radiation shielding can be mitigated by the use of thick and heavy concrete doors (requiring dedicated mechanical solutions) and vice versa. Finally, it was found that any additional, beyond the nominal, material and equipment intersecting the particle beams might have a significant effect on the dose rate escaping a cave. For example, as it was simulated, a variation of the thickness of the vacuum chamber walls intersecting a high energy electron beam led to a large increase of the out–of–cave dose rate. Therefore, constant monitoring of the dose rate at multiple locations outside the caves is required and envisaged for CALA.

Appendix A

Appendix to "Noise reconstruction in proton CT"

A.0.1 Approximate variance reconstruction

In this section, we describe how equation 6.15 can be approximated through equation 6.20. Consider the two convolution kernels $g^2(j\Delta\xi)$ (equation 6.3) and $g_{\mathrm{C}}(j\Delta\xi)$ (equation 6.18) (shown in figure A.1) of the convolutions necessary for the reconstruction of the variance in V_{γ_n} (equation 6.16) and C_{γ_n} (equation 6.19) respectively. With the Ram-Lak filter, $g_{\mathrm{C}}(j\Delta\xi)$ takes an especially simple form. Due to the alternating structure of $g(j\Delta\xi)$, a shift by $1\Delta\xi$ cancels all side lobes. Only adjacent pixels mutually influence each other (different apodization windows may have more complex convolution kernels).

In figure A.1, one can see that $g^2(j\Delta\xi)$ and $g_{\mathrm{C}}(j\Delta\xi)$ have both very limited reach. As a consequence, each filtered projection value is approximately only a weighted sum of its nearest neighbors (for the Ram-Lak covariance kernel, it is exactly only one nearest neighbor). Under the assumption that the projections are locally approximately constant, we are able to approximate all projection values (noise values) with the one at the center of the convolution kernel

$$V_{\gamma_n}(j\Delta\xi) \approx \sigma^2_{\gamma_n}(j\Delta\xi) \sum_{m=-D/2}^{D/2-1} g^2(m\Delta\xi) \tag{A.1}$$

$$C_{\gamma_n}(j\Delta\xi, (j+1)\Delta\xi) \approx \sigma^2_{\gamma_n}(j\Delta\xi) \sum_{m=-D/2}^{D/2-1} g_{\mathrm{C}}(m\Delta\xi). \tag{A.2}$$

As the sum of equation A.1 converges quickly, we can extend it to infinity. Thus, with the expression of the Ram-Lak filter (equation 6.3), we are able to evaluate the sum analytically (for example with the Fourier transform, as described in (Martin 2010))

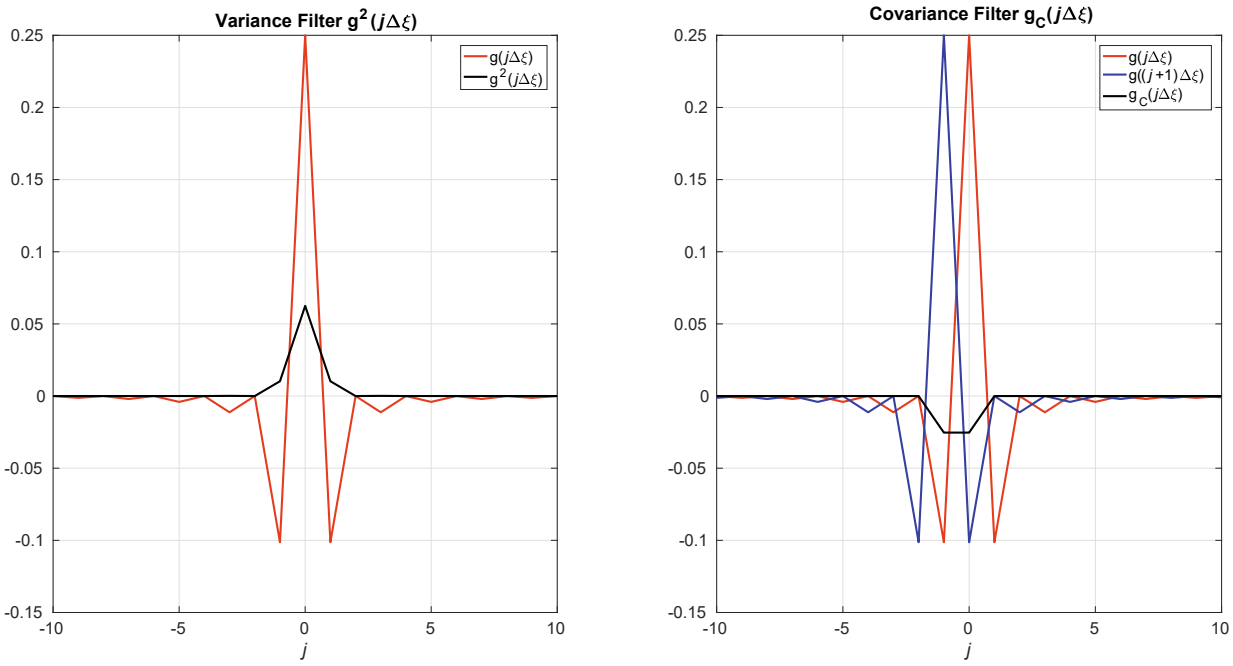

Figure A.1: Variance and covariance filters (black lines) for unit detector spacing $\Delta\xi = 1$.

$$\begin{aligned}\sum_{m=-D/2}^{D/2-1} g^2(m\Delta\xi) &\approx \sum_{m=-\infty}^{\infty} g^2(m\Delta\xi) \\ &= \frac{1}{(2\Delta\xi)^4} + \frac{2}{(\pi\Delta\xi)^4} \cdot \underbrace{\sum_{m=1}^{\infty} \frac{1}{(2m-1)^4}}_{\pi^4/96} = \frac{1}{12(\Delta\xi)^4}.\end{aligned} \tag{A.3}$$

With the Ram-Lak filter, the sum in equation A.2 has only two (equal) terms (see figure A.1)

$$\sum_{m=-D/2}^{D/2-1} g_C(m\Delta\xi) \overset{\text{equation 6.18,equation 6.3}}{=} -\frac{1}{2\pi^2(\Delta\xi)^4}. \tag{A.4}$$

Finally, we can factorize equation 6.15:

$$\mathrm{Var}\left[f(x_p, y_p)\right] = \left(\frac{\pi}{N_p}\Delta\xi\right)^2 \sum_{n=1}^{N_p} \frac{\sigma^2_{\gamma_n}(j\Delta\xi)}{12(\Delta\xi)^4} \left\{(1-u)^2 + 2(1-u)u\frac{-12}{2\pi^2} + u^2\right\} \tag{A.5}$$

The factorized term is the approximation of the convolution, given by equation A.1, while the term in curly brackets comprises the interpolation effect on the noise. The reconstruction in terms of the interpolation effect is then given through

$$\mathrm{Var}\left[f(x_p, y_p)\right] = \left(\frac{\pi}{N_p}\Delta\xi\right)^2 \sum_{n=1}^{N_p} V_{\gamma_n}(j\Delta\xi) f_{\mathrm{interp}}(u), \tag{A.6}$$

where

$$f_{\mathrm{interp}}(u) = (1-u)^2 + 2(1-u)u\frac{-12}{2\pi^2} + u^2. \tag{A.7}$$

Since there is no preferred query point for the interpolation, we assume u to be uniformly distributed in $[0, 1]$, therefore f_{interp} can be approximated by its mean

$$f_{\mathrm{interp},\mu} = \int_0^1 f_{\mathrm{interp}}(u)\mathrm{d}u = \frac{2}{3} - \frac{2}{\pi^2}. \tag{A.8}$$

Replacing $f_{\mathrm{interp}}(u)$ by $f_{\mathrm{interp},\mu}$ in equation A.6 yields the expression given in equation 6.20.

In reality, the distribution of u-values is not perfectly uniform. Figure A.2 shows the mean of equation A.7 for a finite set of projections. The resulting structures in figure A.2 are caused by the 'interference' between the 2D image pixel grid and the 1D projections grid. It is an inherent property of accurate noise reconstruction, which eventually superimposes with the noise projections.

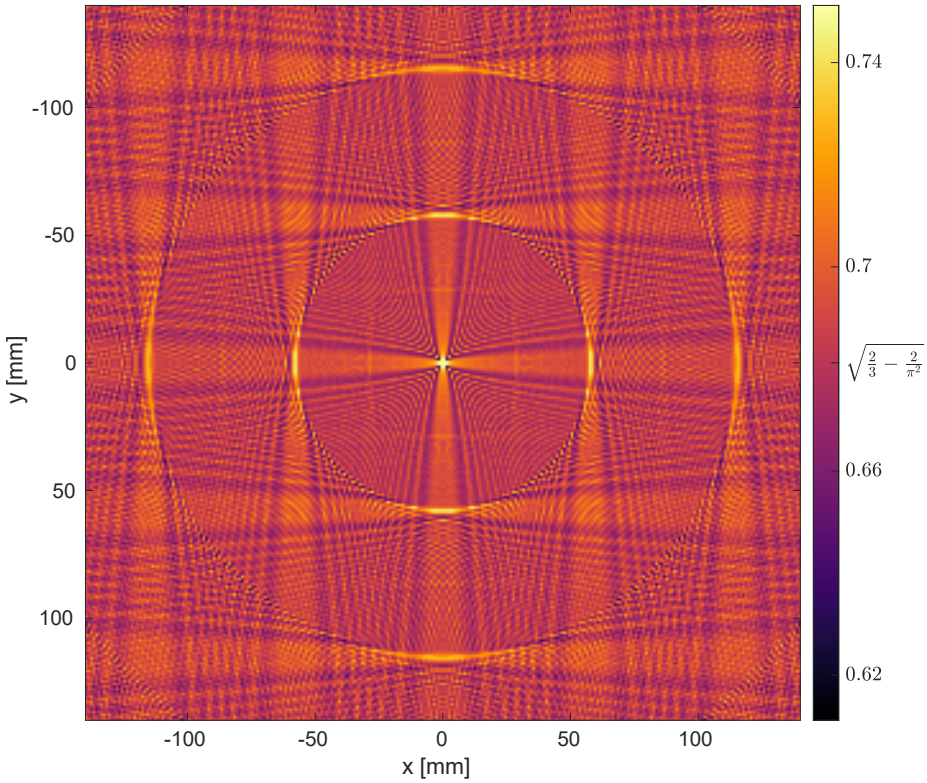

Figure A.2: Exact noise reduction per pixel for an image, reconstructed with linear interpolation and the Ram-Lak filter. The image measures 280mm $\times$ 280mm with 1mm $\times$ 1mm voxel size. The structures are a consequence of the interference with the 1mm spaced projections. Notice that the center is particularly high, as its pixel center is for most angles close to a sampled projection value and thus profits from the interpolation less. Pixels with higher values happen to fall on the discrete projection values more often than in between. Deviations from the approximation are, however, in general quite small.

A.0.2 2D noise reconstruction including DDB

Here we will present a more thorough discussion of the 2D interpolation involved in the DDB variance reconstruction, as mentioned in section 6.2.2. We will discuss the effect of this two–dimensional interpolation on the pCT noise based on bilinear interpolation (see figure A.3). With the bilinear interpolation, the reconstruction becomes

$$\begin{aligned} f(x_p, y_p) = \frac{\pi}{N_p}\Delta\xi \sum_{n=1}^{N_p} \sum_{m=-D/2}^{D/2-1} & p_{\gamma_n}(m\Delta\xi, k\Delta\eta) g((j-m)\Delta\xi)\,[1-u]\,[1-v] \\ & + p_{\gamma_n}(m\Delta\xi, k\Delta\eta) g((j+1-m)\Delta\xi) u\,[1-v] \\ & + p_{\gamma_n}(m\Delta\xi, (k+1)\Delta\eta) g((j-m)\Delta\xi)\,[1-u]\,v \\ & + p_{\gamma_n}(m\Delta\xi, (k+1)\Delta\eta) g((j+1-m)\Delta\xi) uv, \end{aligned} \tag{A.9}$$

where

$$v = v(\eta_n) = \frac{\eta_n - k\Delta\eta}{\Delta\eta}, \tag{A.10}$$

and $\eta_n = \eta_n(x_p, y_p) = -x_p \sin(\gamma_n) + y_p \cos(\gamma_n)$. Just as in equation 6.15, the dependencies $j = j(\xi_n)$ and $k = k(\eta_n)$ are implicit.

Just as we had before, projection values from different angles γ_n as well as along the ξ-coordinate $m\Delta\xi$ are independent. However, this holds only true for $m\Delta\xi$–values binned at the same depth η. Since, due to the bilinear interpolation, we sum up projection values from different depths, we have to take their covariance into account.

$$\begin{aligned} \mathrm{Cov}\left[p_{\gamma_n}(m\Delta\xi, k\Delta\eta), p_{\gamma_{n'}}(m'\Delta\xi, k'\Delta\eta)\right] &= \delta_{n,n'}\delta_{m,m'}\mathrm{Cov}\left[p_{\gamma_n}(m\Delta\xi, k\Delta\eta), p_{\gamma_n}(m\Delta\xi, k'\Delta\eta)\right] \\ &\equiv \delta_{n,n'}\delta_{m,m'} C_{\gamma_n}(m\Delta\xi, k\Delta\eta, k'\Delta\eta) \end{aligned} \tag{A.11}$$

Without further specifying this covariance term and following the procedure from above (equation 6.11, equation 6.14, equation 6.15), the variance of equation A.9 becomes

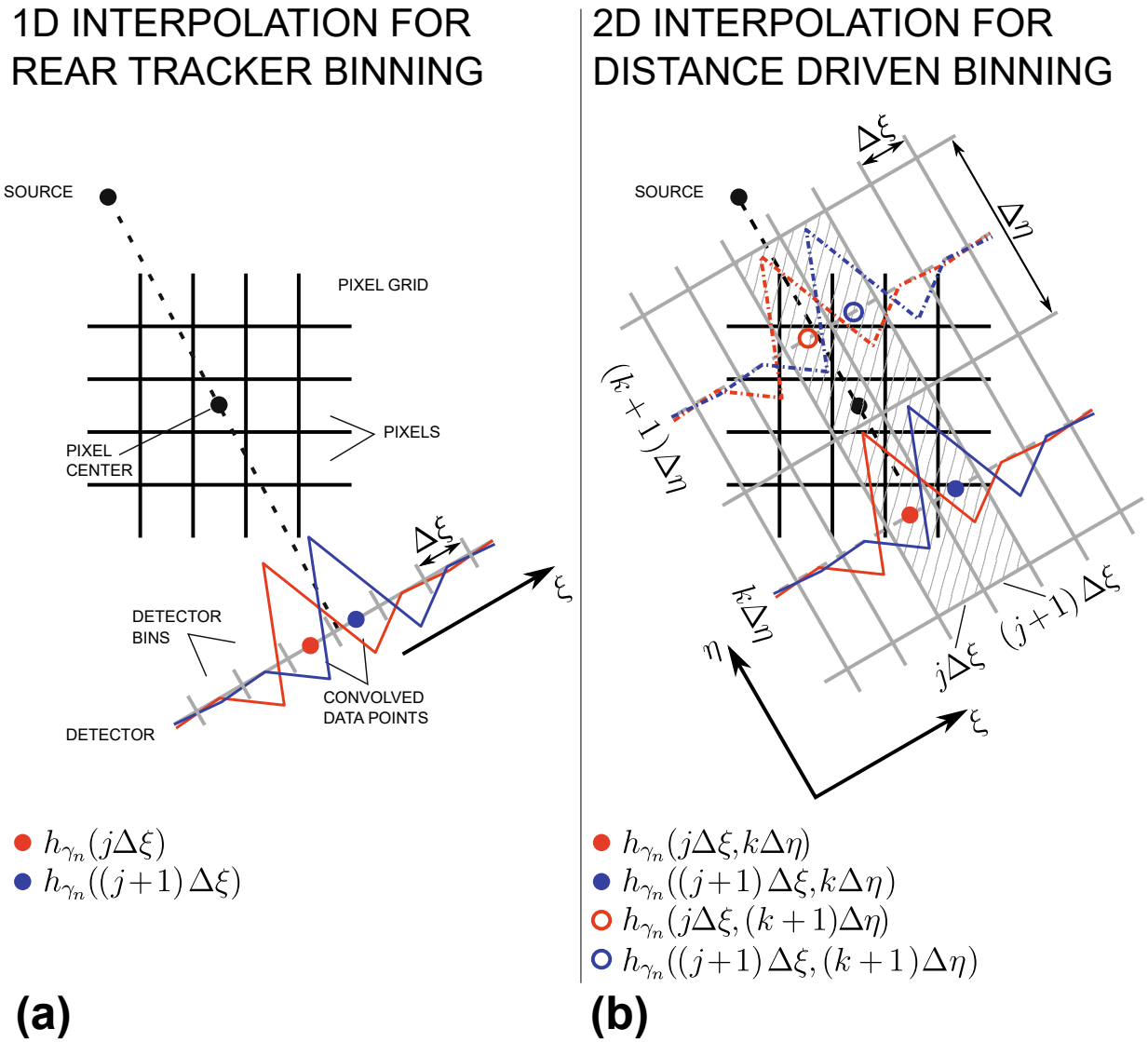

Figure A.3: Interpolation for a 2D image reconstruction **(a)** without and **(b)** with DDB. **(a)** For binning at the rear tracker, the value at the pixel center (black dot) requires a 1D interpolation (at the dashed line) of the convolved projection values (red and blue dot), which are a weighted linear combination of all projection values and thus mutually dependent due to the prior convolution with the convolution kernels (red and blue zigzag lines), as shown in figure A.1. **(b)** In the DDB case, a 2D interpolation is necessary, where projections, binned at different depths (i.e. $k\Delta\eta$ and $(k+1)\Delta\eta$), are involved. The four hatched pixels contribute to the value at the pixel center (black dot). The convolution is still only along ξ. The detector spacing $\Delta\xi$ and the depth spacing $\Delta\eta$ are only drawn different for visual clarity.

$$
\begin{aligned}
\mathrm{Var}\left[f\left(x_p, y_p\right)\right] = \left(\frac{\pi}{N_p}\Delta\xi\right)^2 \sum_{n=1}^{N_p} \{ \\
&[1-u]^2\,[1-v]^2\, V_{\gamma_n}(j\Delta\xi, k\Delta\eta) + u^2\,[1-v]^2\, V_{\gamma_n}((j+1)\Delta\xi, k\Delta\eta) \\
&+ [1-u]^2\, v^2 V_{\gamma_n}(j\Delta\xi, (k+1)\Delta\eta) + u^2 v^2 V_{\gamma_n}((j+1)\Delta\xi, (k+1)\Delta\eta) \\
&+ 2\cdot[1-u]\,u\,[1-v]^2\, C_{\gamma_n}(j\Delta\xi, (j+1)\Delta\xi, k\Delta\eta, k\Delta\eta) \\
&+ 2\cdot[1-u]^2\,[1-v]\,v C_{\gamma_n}(j\Delta\xi, j\Delta\xi, k\Delta\eta, (k+1)\Delta\eta) \\
&+ 2\cdot[1-u]\,u\,[1-v]\,v C_{\gamma_n}(j\Delta\xi, (j+1)\Delta\xi, k\Delta\eta, (k+1)\Delta\eta) \\
&+ 2\cdot[1-u]\,u\,[1-v]\,v C_{\gamma_n}((j+1)\Delta\xi, j\Delta\xi, k\Delta\eta, (k+1)\Delta\eta) \\
&+ 2\cdot u^2\,[1-v]\,v C_{\gamma_n}((j+1)\Delta\xi, (j+1)\Delta\xi, k\Delta\eta, (k+1)\Delta\eta) \\
&+ 2\cdot[1-u]\,uv^2 C_{\gamma_n}(j\Delta\xi, (j+1)\Delta\xi, (k+1)\Delta\eta, (k+1)\Delta\eta)\ \}\,,
\end{aligned}
\tag{A.12}
$$

where the equivalent expressions of V_{γ_n} and C_{γ_n} in two dimensions are

$$
V_{\gamma_n}(j\Delta\xi, k\Delta\eta) = \sum_{m=-D/2}^{D/2-1} g^2((j-m)\Delta\xi)\sigma^2_{\gamma_n}(m\Delta\xi, k\Delta\eta), \tag{A.13}
$$

$$
\begin{aligned}
C_{\gamma_n}(j\Delta\xi, j'\Delta\xi, k\Delta\eta, k'\Delta\eta) = \\
&\sum_{m=-D/2}^{D/2-1} g((j-m)\Delta\xi) g((j'-m)\Delta\xi) C_{\gamma_n}(m\Delta\xi, k\Delta\eta, k'\Delta\eta).
\end{aligned}
\tag{A.14}
$$

The covariance values between two data points from the same depth ($k = k'$) becomes again $C_{\gamma_n}(m\Delta\xi, k\Delta\eta, k\Delta\eta) = \sigma^2_{\gamma_n}(m\Delta\xi, k\Delta\eta)$, just like in equation 6.12. The remaining covariances are between projection values from adjacent depths, which we have not yet discussed. Note that projection values $p_{\gamma_n}(m\Delta\xi, k\Delta\eta)$ and $p_{\gamma_n}(m\Delta\xi, (k+1)\Delta\eta)$ are calculated from almost the same data set of protons, given that the pixel spacing in the η–direction ($\Delta\eta$) is sufficiently small. This is due to that fact that within $[k\Delta\eta, (k+1)\Delta\eta]$ only very few protons outside of $m\Delta\xi$ will scatter laterally into $m\Delta\xi$ and at the same time only very few protons within $m\Delta\xi$ will scatter to neighboring pixels. Thus projection values from any two neighboring depths are hardly different (below approximately 0.2 mm WEPL for the reconstruction of our simulation (see section 6.2.1), using $\Delta\eta = 1$mm) and can therefore be considered equal. We can thus assume that depth adjacent projection values are perfectly correlated while the two diagonal pixels in the bilinear interpolation have no correlation

$$
C_{\gamma_n}(m\Delta\xi, k\Delta\eta, k'\Delta\eta) = \delta_{k,k'}\sigma^2_{\gamma_n}(m\Delta\xi, k\Delta\eta). \tag{A.15}
$$

As there is now no difference between the projection values at $k\Delta\eta$ and $(k+1)\Delta\eta$, we can now make the replacement

$$C_{\gamma_n}(j\Delta\xi, j'\Delta\xi, k\Delta\eta, k'\Delta\eta) \to C_{\gamma_n}(j\Delta\xi, j'\Delta\xi, k\Delta\eta). \tag{A.16}$$

The noise reconstruction including DDB is then given by the expression of equation 6.21. It is similar to the noise reconstruction of the rear tracker binning (equation 6.15), as we neglected the covariance along η. The v–dependence in equation A.12 cancels under the approximation of equation A.16. The index $k = k(\eta_n)$ is still query point dependent, but as the projection values of two neighboring depths are considered to be equal, nearest neighbor or linear interpolation along η is sufficient for the noise map reconstruction.

A.0.3 Sinogram interpolation

We can estimate the function $F(\xi, \mathrm{IP})$ that maps the IPs to the corresponding energy loss along the straight line from the IP–hull intersection to some exit detector pixel $j\Delta\xi$ by interpolating across the Radon space, as the various straight proton paths, that contribute to one pixel, are line integrals from neighboring projections coming from different angles γ_n. The set of angles results from the IP–hull intersection coordinate (i.e. projecting the IPs onto the hull) and the exit detector pixel coordinate. Therefore every IP determines an angle $\gamma_{\mathrm{IP}}(j\Delta\xi)$, which is different for every $j\Delta\xi$. The set of the corresponding ξ values in the Radon space is determined by

$$\xi(j\Delta\xi, \mathrm{IP}) = x_{\mathrm{bin}} \cos\left(\gamma_{\mathrm{IP}}(j\Delta\xi)\right) + y_{\mathrm{bin}} \sin(\gamma_{\mathrm{IP}}(j\Delta\xi)), \tag{A.17}$$

where $(x_{\mathrm{bin}}, y_{\mathrm{bin}})$ is the coordinate of the exit detector bin $j\Delta\xi$ on the rear tracker in the image space (x, y). The determination of $F(j\Delta\xi, \mathrm{IP})$ is demonstrated in figure A.4. This process is similar to transforming parallel beam projections to fan beam projections, where $(x_{\mathrm{bin}}, y_{\mathrm{bin}})$ can be considered the source point. The set of angles is quite irregular though, depending on the object hull.

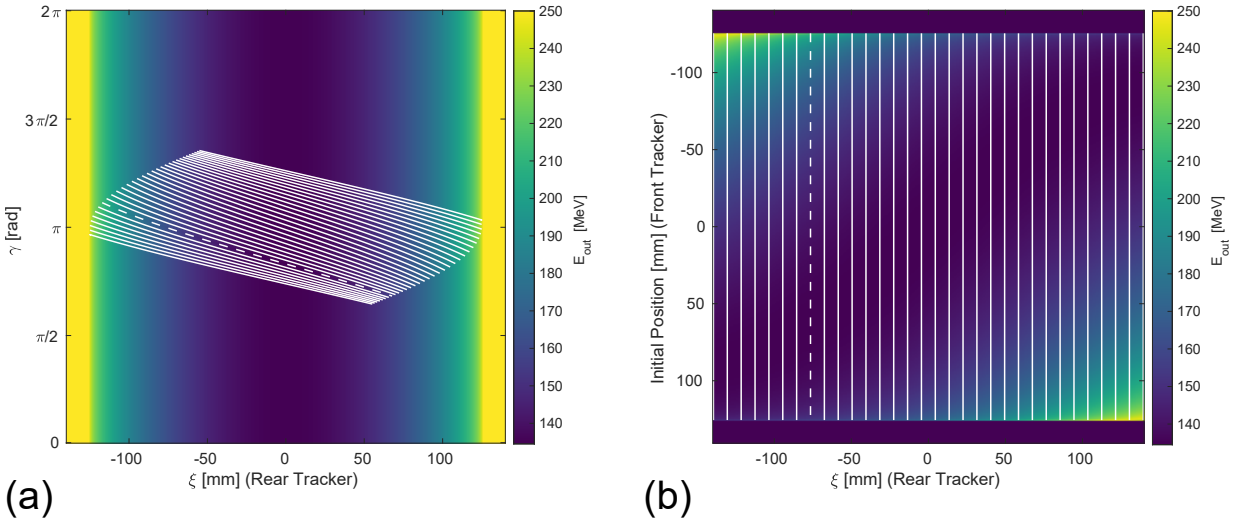

Figure A.4: Mapping of the Radon space in units of the exit energy **(a)** to $F(j\Delta\xi, \mathrm{IP})$ **(b)**. The data on the white lines in **(b)** was interpolated from the data shown in **(a)** along the corresponding curved lines, where the dashed lines in both figures belong together and are supposed to give orientation.

Appendix B

Appendix to "Image noise contributions in proton CT"

B.0.1 Derivation of the distorted energy deposit

The scaling factor S_n in equation 7.2 is fixed during the calibration procedure where for a degrader–free run the energy deposits of all stages are forced to be equal to values $E_n^{\mathrm{G4}} = \{25.25, 28.01, 32.76, 42.62, 67.71\}\,\mathrm{MeV}$ pre–calculated by GEANT4 (see details in Bashkirov et al. (2016)). Therefore we find

$$E_n^{\mathrm{G4}} = S_n \cdot \int_{R_n^i}^{R_n^f} \mathrm{d}x \, \frac{\mathrm{d}E/\mathrm{d}x}{1 + k_B \cdot \mathrm{d}E/\mathrm{d}x}, \tag{B.1}$$

where R_n^i and R_n^f are the proton's residual ranges at the entrance and at the exit of the stage, respectively. Now, let l_{stage} be the thickness of each scintillating stage and R_0 be the residual range in the detector material of a proton at the entrance of the detector without any degrader. Then we find for the degrader–free calibration run

$$R_n^i = R_0 - (n-1) \cdot l_{\mathrm{stage}}, \tag{B.2}$$

$$R_n^f = R_0 - n \cdot l_{\mathrm{stage}}. \tag{B.3}$$

For a proton that additionally passed through a WEPL during the calibration with the wedge–shaped phantom or during a measurement, we can calculate

$$R_n = R_n^i - \mathrm{WEPL}/\mathrm{RSP}_{\mathrm{det}}, \tag{B.4}$$

where $\mathrm{RSP}_{\mathrm{det}} = 1.038$ is the relative stopping power of the detector material (see Bashkirov

et al. (2016)). By using equations 7.2, B.1, B.2, B.3, B.4 this yields

$$E'_n(\mathrm{WEPL}) = E_n^{\mathrm{G4}} \cdot \frac{\int_{R_n^i - \mathrm{WEPL}/\mathrm{RSP}_{\mathrm{det}}}^{0} \mathrm{d}x \, \frac{\mathrm{d}E/\mathrm{d}x}{1 + k_B \cdot \mathrm{d}E/\mathrm{d}x}}{\int_{R_n^i}^{R_n^f} \mathrm{d}x \, \frac{\mathrm{d}E/\mathrm{d}x}{1 + k_B \cdot \mathrm{d}E/\mathrm{d}x}}, \tag{B.5}$$

where $E'_n(\mathrm{WEPL})$ is the expected energy measurement for a given Birks' factor k_B.

B.0.2 Noise contributions at a discontinuity

As a side study to investigations in section 7.2.8, and to demonstrate the impact of heterogeneities, we calculated standard deviation profiles for the steep edge of the calibration phantom with two bricks (upper part of the phantom as seen in figure 7.1). There, the WEPL rapidly drops from 101.8 mm to zero. The same scoring techniques as for the smooth part of the phantom were exploited. Figure B.1 shows a lateral profile of standard deviation for the steep edge of the phantom. In the vicinity of the discontinuity located at 104.5 mm from the isocenter, standard deviation is increased to 27 mm in the measurement, which is an increase by a factor of 8.1 compared to the homogeneous region. *WEPL scoring* standard deviation representing the scatter–only contribution is at 16 mm and thus more than half of the total standard deviation. Agreement between the measurement and *realistic scoring* is satisfactory.

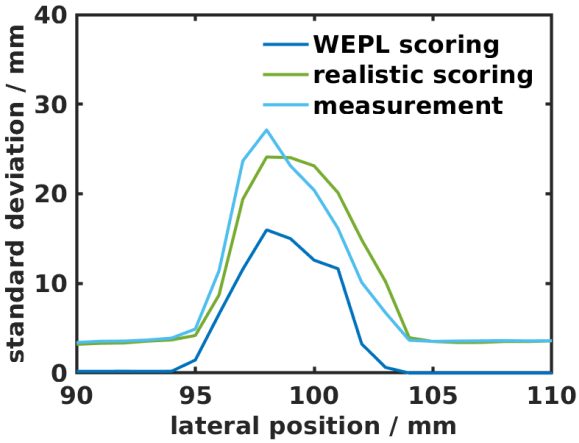

Figure B.1: Standard deviation profile along the lateral coordinate for a discontinuity of two bricks (ΔWEPL = 101.6 mm) and for front tracker binning. The discontinuity in WEPL is located at 104.5 mm from the isocenter. Note that the scale is different by a factor of 10 compared to figure 7.6 (a).

This shows that a discontinuity of 101.6 mm increased the WEPL standard deviation by a factor of 8.1. This distinct increase of noise was driven by the scatter–only component, which was negligible for the homogeneous phantom in figure 7.6, but contributed a large noise fraction at the heterogeneity. Therefore, heterogeneities are likely to dominate image noise in certain regions and need to be considered using a precise patient model. To achieve the same noise level at the edge as for the homogeneous wedge phantom, the dose would need to be increased by a factor of 65. However, this is an extreme case and neglects the smoothing effect of interpolation and filtering during the reconstruction. Moreover, this does not mean that the dose increase would be required considering the low diagnostic value of RSP at the skin. The Monte Carlo simulation modeled scattering correctly, as the resulting noise increase is captured well when compared to the measurement.

B.0.3 Full–volume RSP and noise histograms

As a sanity check, we calculated RSP and standard deviation histograms of the whole volume of the pediatric head phantom. For this, the two acquisitions of the superior and the inferior part of the phantom needed to be merged. The phantom's outer hull was determined per slice as described in section 7.2.9 and values outside of it were neglected. Slices close to the detector's upper and lower edges were also excluded, because noise increases dramatically there.

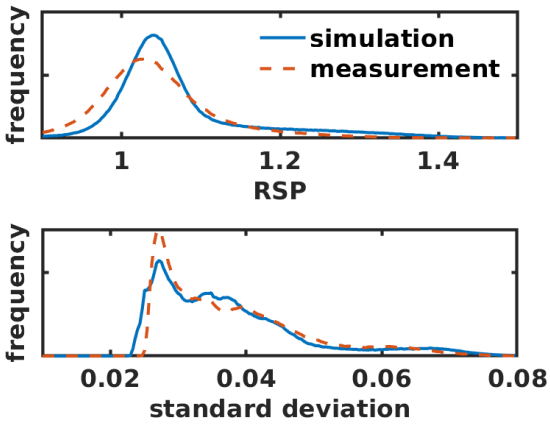

Figure B.2: RSP and standard deviation histograms for simulated and experimental data of all slices of the pediatric head phantom.

Figure B.2 shows histograms of RSP and standard deviation. While the RSP distribution showed a Gaussian shape around a mean value slightly above 1, the distribution of standard deviation values exhibited a broader structure. For simulated data, the standard deviation 5– to 95–percentile range was between 0.026 and 0.064 while for measurements it was between 0.026 and 0.061. This is an increase from the lower to the higher percentile value by a factor of 2.5 or 2.3, respectively. Note that the lower percentile value is close to the non–scatter contribution $\sigma_{\text{non-scatter}}$ in section 7.3.5.

In conclusion, over the complete volume of the pediatric head phantom, agreement of standard deviation histograms was comparable to the agreement of RSP histograms.

Appendix C

Appendix to "Biological uncertainties in proton treatment planning"

C.0.1 Dose and LET scoring

In the GEANT4 simulations, the reference physics list `QGSP_BIC_HP` was used and the dose in each voxel was calculated according to

$$D_w = \frac{\rho_m}{\rho_w} \frac{S_w}{S_m} D_m, \tag{C.1}$$

where D_m is the dose to material, ρ_m and ρ_w the mass density of the material in the voxel and the density of water (`G4_WATER`) respectively. The unrestricted stopping powers of water (S_w) and of the material (S_m) were computed at the proton energy by the GEANT4 function `ComputeTotalDEDX`.

The $\mathrm{LET_d}$ in voxel i (L_i) was the average over all steps S_n for all particles N

$$L_i = \frac{\sum_{n=1}^{N} \sum_{s=1}^{S_n} \frac{\varepsilon_{sn}^2}{l_{sn}}}{\sum_{n=1}^{N} \sum_{s=1}^{S_n} \varepsilon_{sn}}, \tag{C.2}$$

where ε is the energy deposition (`GetTotalEnergyDeposit`) and l_{sn} the length of step s of event n (`GetStepLength`). Since the dose–to–water concept was used for dose calculation, the energy deposition ε was scaled with (S_w/S_m), which is equivalent to equation C.1. While the energy deposition of any particle type was scored in dose calculation, only primary and secondary protons were taken into account for $\mathrm{LET_d}$ calculation (Grassberger & Paganetti 2011). In GEANT4, the incident lateral PB shape was modeled as a two–dimensional Gaussian with a FWHM equal to 9.4 mm and the initial energy spread was assumed to be normally distributed with a standard deviation equal to 0.5% of the nominal energy, which is a realistic value in modern proton therapy facilities (Pedroni

et al. 2005, Grevillot et al. 2010). In CERR, the weights of the PBs were determined by the minimization of a Chi–squared cost function

$$\chi^2 = \sum\nolimits_{i \in T} \left(D_i(\mathbf{w}) - D^T \right)^2 , \tag{C.3}$$

with the Matlab/2014a routine `fmincon` as described in (Schell & Wilkens 2010). D^T denotes the desired dose in structure T and $D_i(\mathbf{w})$ the dose in a voxel i for a current set of M spot weights $\mathbf{w} = (w_1, w_2, ..., w_M)$ (Wilkens & Oelfke 2005).

C.0.2 Biological effect optimization

In the presence of a set of M beam spots (PBs), the dose in a voxel i (D_i) can be expressed as

$$D_i(\mathbf{w}) = \sum_{j=1}^{M} w_j D_{ij}, \tag{C.4}$$

where the D_{ij} matrix contains the dose deposited for unit fluence by beam j in voxel i and $\mathbf{w}$ denotes the relative fluence weights of all M beam spots. Accordingly to the D_{ij} matrix the L_{ij} matrix is defined for $\mathrm{LET_d}$. Hence the total dose average LET for a set of M PBs in the voxel i can be calculated by

$$\begin{aligned} L_i(\mathbf{w}) &= \frac{1}{D_i} \sum_{j=1}^{M} w_j D_{ij} L_{ij}, \quad \text{for} \quad D_i > 0 \\ L_i(\mathbf{w}) &= 0, \quad \text{for} \quad D_i = 0 \end{aligned} \tag{C.5}$$

Instead of optimizing for a uniformly distributed RBE weighted dose (RWD), a desired biological effect is optimized. In the LQ model, the biological effect is the negative logarithm of the survival fraction (equation 10.1) $\epsilon = -ln(S) = \alpha D + \beta D^2$. Therefore the desired biological effect in a tissue T (ϵ^T) can directly be determined from clinical experience with photon irradiation (here, using the values from tables 10.1 and 10.2). A quadratic deviation cost function can be formulated as

$$F_\epsilon(\mathbf{w}) = \sum_{i \in T} \left(\epsilon_i(\mathbf{w}) - \epsilon^T \right)^2 , \tag{C.6}$$

using equation 10.2 to calculate the α_p and β_p parameter. Hence,

$$F_\epsilon(\mathbf{w}) = \sum_{i \in T} \left(\alpha_x^T D_i(\mathbf{w}) + q \beta_x^T L_i(\mathbf{w}) D_i(\mathbf{w}) + \beta_x^T D_i^2(\mathbf{w}) - \epsilon^T \right)^2 \tag{C.7}$$

and the first partial derivative

$$\begin{aligned} \frac{\partial F_\epsilon(\mathbf{w})}{\partial w_k} &= \sum_{i \in T} 2 \left(\alpha_x^T D_i(\mathbf{w}) + q \beta_x^T L_i(\mathbf{w}) D_i(\mathbf{w}) + \beta_x^T D_i^2(\mathbf{w}) - \epsilon^T \right) \\ &\cdot \left(\alpha_x^T D_{ik} + q \beta_x^T L_{ik} D_{ik} + 2 \beta_x^T D_i(\mathbf{w}) D_{ik} \right), \end{aligned} \tag{C.8}$$

which is necessary for the gradient based optimization. Cost functions for OARs and multiple targets can be defined in the same way, optionally multiplied with a penalty factor and added to a total cost function. Note that the α_x^T and β_x^T are tissue specific parameters.

Appendix D

Appendix to "GEANT4 hadronic models for prompt–gamma emission"

D.0.1 GEANT4 physics list

Table D.1: GEANT4 models used for the simulation of hadronic interactions.

Physical process	particle	GEANT4 process	GEANT4 model	Energy range
Elastic	p, d, t, ^{3}He, α GenericIon	G4HadronElasticProcess	G4HadronElastic	–
	n	G4HadronElasticProcess	G4NeutronHPElastic G4HadronElastic	≤ 20 MeV ≥ 19 MeV
Inelastic	p	G4ProtonInelasticProcess	G4BinaryCascade	–
	n	G4NeutronInelasticProcess	G4NeutronHPInelastic G4BinaryCascade	≤ 20 MeV ≥ 19 MeV
	d	G4DeuteronInelasticProcess	G4QMDReaction G4BinaryLightIonReaction	– –
	t	G4TritonInelasticProcess	G4QMDReaction G4BinaryLightIonReaction	– –
	^{3}He	G4IonInelasticProcess	G4QMDReaction G4BinaryLightIonReaction	– –
	α	G4AlphaInelasticProcess	G4QMDReaction G4BinaryLightIonReaction	– –
	GenericIon	G4IonInelasticProcess	G4QMDReaction G4BinaryLightIonReaction	– –
Fission	n	G4HadronFissionProcess	G4NeutronHPFission G4LFission	≤ 20 MeV ≥ 19 MeV
Capture	n	G4HadronCaptureProcess	G4NeutronHPCapture G4LCapture	≤ 20 MeV ≥ 19 MeV
Decay	GenericIon	G4Decay	G4RadioactiveDecay	–

D.0.2 GANIL fragmentation experiment

Table D.2: Angular distribution of production rates of charged fragments in count/ion/sr for a 5 mm thick PMMA target. Data (Braunn 2010) are compared with the QMD default value of $L = 2\,\mathrm{fm}^2$, as well as with $L = 1\,\mathrm{fm}^2$, $L = 0.8\,\mathrm{fm}^2$.

	$Z=1$				$Z=1$			
	Data	$L=2\,\mathrm{fm}^2$	$L=1\,\mathrm{fm}^2$	$L=0.8\,\mathrm{fm}^2$	Data	$L=2\,\mathrm{fm}^2$	$L=1\,\mathrm{fm}^2$	$L=0.8\,\mathrm{fm}^2$
7°	1.40×10^{-1}	1.89×10^{-1}	1.51×10^{-1}	1.58×10^{-1}	2.10×10^{-1}	1.52×10^{-1}	1.20×10^{-1}	1.20×10^{-1}
10°	7.50×10^{-2}	1.55×10^{-1}	1.22×10^{-1}	1.27×10^{-1}	1.10×10^{-1}	9.52×10^{-2}	6.20×10^{-2}	5.13×10^{-2}
16°	3.00×10^{-2}	8.75×10^{-2}	6.60×10^{-2}	7.42×10^{-2}	1.80×10^{-2}	2.29×10^{-2}	1.45×10^{-2}	1.39×10^{-2}
20°	3.00×10^{-2}	5.26×10^{-2}	4.56×10^{-2}	5.30×10^{-2}	9.80×10^{-3}	6.36×10^{-3}	6.33×10^{-3}	6.33×10^{-3}
30°	8.40×10^{-3}	2.43×10^{-2}	2.08×10^{-2}	2.17×10^{-2}	1.80×10^{-3}	8.95×10^{-4}	1.39×10^{-3}	1.30×10^{-3}
40°	9.40×10^{-3}	1.29×10^{-2}	1.12×10^{-2}	1.12×10^{-2}	6.40×10^{-4}	2.86×10^{-4}	4.48×10^{-4}	4.39×10^{-4}
60°	2.10×10^{-3}	3.38×10^{-3}	3.65×10^{-3}	3.80×10^{-3}	4.70×10^{-5}	6.97×10^{-5}	6.61×10^{-5}	6.42×10^{-5}
70°	1.20×10^{-3}	1.41×10^{-3}	1.86×10^{-3}	2.04×10^{-3}	3.00×10^{-5}	2.94×10^{-5}	2.57×10^{-5}	1.83×10^{-5}
	$Z=3$				$Z=4$			
	Data	$L=2\,\mathrm{fm}^2$	$L=1\,\mathrm{fm}^2$	$L=0.8\,\mathrm{fm}^2$	Data	$L=2\,\mathrm{fm}^2$	$L=1\,\mathrm{fm}^2$	$L=0.8\,\mathrm{fm}^2$
7°	4.20×10^{-2}	2.10×10^{-2}	2.02×10^{-2}	2.02×10^{-2}	2.10×10^{-2}	1.45×10^{-2}	2.01×10^{-2}	2.01×10^{-2}
10°	1.30×10^{-2}	1.26×10^{-2}	1.04×10^{-2}	1.09×10^{-2}	5.50×10^{-3}	8.86×10^{-3}	8.01×10^{-3}	7.26×10^{-3}
16°	8.10×10^{-4}	1.42×10^{-3}	1.86×10^{-3}	2.10×10^{-3}	2.30×10^{-4}	5.01×10^{-4}	9.19×10^{-4}	8.09×10^{-4}
20°	8.50×10^{-4}	1.10×10^{-4}	5.37×10^{-4}	5.37×10^{-4}	2.00×10^{-4}	4.00×10^{-5}	1.87×10^{-4}	1.87×10^{-4}
30°	4.60×10^{-5}	9.17×10^{-6}	4.77×10^{-5}	8.26×10^{-5}	8.20×10^{-6}	∅	2.75×10^{-5}	3.12×10^{-5}
	$Z=5$				$Z=6$			
	Data	$L=2\,\mathrm{fm}^2$	$L=1\,\mathrm{fm}^2$	$L=0.8\,\mathrm{fm}^2$	Data	$L=2\,\mathrm{fm}^2$	$L=1\,\mathrm{fm}^2$	$L=0.8\,\mathrm{fm}^2$
7°	2.50×10^{-2}	3.24×10^{-2}	2.92×10^{-2}	2.92×10^{-2}	3.40×10^{-2}	2.97×10^{-2}	2.20×10^{-2}	2.20×10^{-2}
10°	3.40×10^{-3}	1.20×10^{-2}	7.22×10^{-3}	5.08×10^{-3}	3.10×10^{-3}	8.53×10^{-3}	4.35×10^{-3}	$2,59\times10^{-3}$
16°	4.40×10^{-5}	1.28×10^{-4}	3.39×10^{-4}	3.76×10^{-4}	2.10×10^{-5}	5.50×10^{-5}	1.38×10^{-4}	1.49×10^{-4}
20°	3.20×10^{-5}	∅	3.20×10^{-5}	3.20×10^{-5}	1.00×10^{-5}	5.00×10^{-6}	2.17×10^{-5}	2.17×10^{-5}
30°	1.10×10^{-6}	∅	1.83×10^{-6}	8.00×10^{-6}	1.30×10^{-7}	∅	1.83×10^{-6}	∅

Table D.3: Angular distribution of production rates of charged fragments in count/ion/sr for a 2 cm thick PMMA target. Data (Braunn 2010) are compared with the QMD default value of $L = 2\,\mathrm{fm}^2$, as well as with $L = 1\,\mathrm{fm}^2$, $L = 0.8\,\mathrm{fm}^2$.

	$Z = 1$				$Z = 1$			
	Data	$L = 2\,\mathrm{fm}^2$	$L = 1\,\mathrm{fm}^2$	$L = 0.8\,\mathrm{fm}^2$	Data	$L = 2\,\mathrm{fm}^2$	$L = 1\,\mathrm{fm}^2$	$L = 0.8\,\mathrm{fm}^2$
7°	3.60×10^{-1}	3.94×10^{-1}	3.48×10^{-1}	3.81×10^{-1}	7.40×10^{-1}	3.53×10^{-1}	3.11×10^{-1}	3.08×10^{-1}
10°	2.90×10^{-1}	3.37×10^{-1}	2.94×10^{-1}	3.17×10^{-1}	4.10×10^{-1}	2.49×10^{-1}	1.75×10^{-1}	1.56×10^{-1}
16°	9.30×10^{-2}	2.19×10^{-1}	1.91×10^{-1}	2.06×10^{-1}	8.80×10^{-2}	1.01×10^{-1}	5.47×10^{-2}	4.51×10^{-2}
33°	4.00×10^{-2}	4.56×10^{-2}	4.51×10^{-2}	4.95×10^{-2}	5.10×10^{-3}	3.10×10^{-3}	1.92×10^{-3}	1.67×10^{-3}
40°	2.10×10^{-2}	2.41×10^{-2}	2.58×10^{-2}	2.77×10^{-2}	1.50×10^{-3}	8.57×10^{-4}	5.30×10^{-4}	4.61×10^{-4}
60°	2.40×10^{-3}	9.72×10^{-3}	5.09×10^{-3}	5.58×10^{-3}	5.20×10^{-5}	2.20×10^{-5}	3.20×10^{-5}	2.39×10^{-5}
	$Z = 3$				$Z = 4$			
	Data	$L = 2\,\mathrm{fm}^2$	$L = 1\,\mathrm{fm}^2$	$L = 0.8\,\mathrm{fm}^2$	Data	$L = 2\,\mathrm{fm}^2$	$L = 1\,\mathrm{fm}^2$	$L = 0.8\,\mathrm{fm}^2$
7°	1.40×10^{-1}	3.85×10^{-2}	4.41×10^{-2}	6.02×10^{-2}	5.70×10^{-2}	2.13×10^{-2}	4.85×10^{-2}	6.73×10^{-2}
10°	4.90×10^{-2}	2.99×10^{-2}	2.55×10^{-2}	3.08×10^{-2}	1.70×10^{-2}	1.60×10^{-2}	2.06×10^{-2}	2.26×10^{-2}
16°	4.60×10^{-3}	1.01×10^{-2}	6.93×10^{-3}	6.98×10^{-3}	1.20×10^{-3}	3.56×10^{-3}	3.12×10^{-3}	2.70×10^{-3}
	$Z = 5$				$Z = 6$			
	Data	$L = 2\,\mathrm{fm}^2$	$L = 1\,\mathrm{fm}^2$	$L = 0.8\,\mathrm{fm}^2$	Data	$L = 2\,\mathrm{fm}^2$	$L = 1\,\mathrm{fm}^2$	$L = 0.8\,\mathrm{fm}^2$
7°	7.20×10^{-2}	5.80×10^{-2}	8.10×10^{-2}	7.20×10^{-2}	1.00×10^{-1}	4.19×10^{-2}	4.64×10^{-2}	4.52×10^{-2}
10°	1.10×10^{-2}	3.21×10^{-2}	2.02×10^{-2}	1.45×10^{-2}	6.10×10^{-3}	5.29×10^{-3}	5.34×10^{-3}	5.33×10^{-3}
16°	8.50×10^{-4}	2.54×10^{-3}	1.20×10^{-3}	7.06×10^{-4}	5.40×10^{-4}	4.06×10^{-4}	3.94×10^{-4}	3.61×10^{-4}

Bibliography

Abraham, J. & Staffurth, J. (2016). Hormonal therapy for cancer, *Medicine (Baltimore)* **44**(1): 30–33.

Agodi, C., Bellini, F., Cirrone, G. A. P., Collamati, F., Cuttone, G., Lucia, E. D., Napoli, M. D., Domenico, A. D., Faccini, R., Ferroni, F., Fiore, S., Gauzzi, P., Iarocci, E., Marafini, M., Mattei, I., Paoloni, A., Patera, V., Piersanti, L., Romano, F., Sarti, A., Sciubba, A. & Voena, C. (2012). Precise measurement of prompt photon emission from 80 MeV/u carbon ion beam irradiation, *J. Instrum.* **7**(03): P03001–P03001.

Agostinelli, S., Allison, J., Amako, K., Apostolakis, J., Araujo, H., Arce, P., Asai, M., Axen, D., Banerjee, S., Barrand, G., Behner, F., Bellagamba, L., Boudreau, J., Broglia, L., Brunengo, A., Burkhardt, H., Chauvie, S., Chuma, J., Chytracek, R., Cooperman, G., Cosmo, G., Degtyarenko, P., DellAcqua, A., Depaola, G., Dietrich, D., Enami, R., Feliciello, A., Ferguson, C., Fesefeldt, H., Folger, G., Foppiano, F., Forti, A., Garelli, S., Giani, S., Giannitrapani, R., Gibin, D., Cadenas, J. G., González, I., Abril, G. G., Greeniaus, G., Greiner, W., Grichine, V., Grossheim, A., Guatelli, S., Gumplinger, P., Hamatsu, R., Hashimoto, K., Hasui, H., Heikkinen, A., Howard, A., Ivanchenko, V., Johnson, A., Jones, F., Kallenbach, J., Kanaya, N., Kawabata, M., Kawabata, Y., Kawaguti, M., Kelner, S., Kent, P., Kimura, A., Kodama, T., Kokoulin, R., Kossov, M., Kurashige, H., Lamanna, E., Lampén, T., Lara, V., Lefebure, V., Lei, F., Liendl, M., Lockman, W., Longo, F., Magni, S., Maire, M., Medernach, E., Minamimoto, K., de Freitas, P. M., Morita, Y., Murakami, K., Nagamatu, M., Nartallo, R., Nieminen, P., Nishimura, T., Ohtsubo, K., Okamura, M., ONeale, S., Oohata, Y., Paech, K., Perl, J., Pfeiffer, A., Pia, M., Ranjard, F., Rybin, A., Sadilov, S., Salvo, E. D., Santin, G., Sasaki, T., Savvas, N., Sawada, Y., Scherer, S., Sei, S., Sirotenko, V., Smith, D., Starkov, N., Stoecker, H., Sulkimo, J., Takahata, M., Tanaka, S., Tcherniaev, E., Tehrani, E. S., Tropeano, M., Truscott, P., Uno, H., Urban, L., Urban, P., Verderi, M., Walkden, A., Wander, W., Weber, H., Wellisch, J., Wenaus, T., Williams, D., Wright, D., Yamada, T., Yoshida, H. & Zschiesche, D. (2003). Geant4—a simulation toolkit, *Nucl. Instrum. Methods Phys. Res., Sect. A* **506**(3): 250–303.

Alaei, P. & Spezi, E. (2015). Imaging dose from cone beam computed tomography in radiation therapy., *Phys. Med.* **31**: 647–658.

Allison, J., Amako, K., Apostolakis, J., Araujo, H., Dubois, P. A., Asai, M., Barrand, G., Capra, R., Chauvie, S., Chytracek, R., Cirrone, G., Cooperman, G., Cosmo, G., Cuttone, G., Daquino, G., Donszelmann, M., Dressel, M., Folger, G., Foppiano, F., Generowicz, J., Grichine, V., Guatelli, S., Gumplinger, P., Heikkinen, A., Hrivnacova, I., Howard, A., Incerti, S., Ivanchenko, V., Johnson, T., Jones, F., Koi, T., Kokoulin, R., Kossov, M., Kurashige, H., Lara, V., Larsson, S., Lei, F., Link, O., Longo, F., Maire, M., Mantero, A., Mascialino, B., McLaren, I., Lorenzo, P. M., Minamimoto, K., Murakami, K., Nieminen, P., Pandola, L., Parlati, S., Peralta, L., Perl, J., Pfeiffer, A., Pia, M., Ribon, A., Rodrigues, P., Russo, G., Sadilov, S., Santin, G., Sasaki, T., Smith, D., Starkov, N., Tanaka, S., Tcherniaev, E., Tome, B., Trindade, A., Truscott, P., Urban, L., Verderi, M., Walkden, A., Wellisch, J., Williams, D., Wright, D. & Yoshida, H. (2006). Geant4 developments and applications, *IEEE Trans. Nucl. Sci.* **53**(1): 270–278.

Almeida, I. P., Schyns, L. E. J. R., Vaniqui, A., van der Heyden, B., Dedes, G., Resch, A. F., Kamp, F., Zindler, J. D., Parodi, K., Landry, G. & Verhaegen, F. (2018). Monte Carlo proton dose calculations using a radiotherapy specific dual-energy CT scanner for tissue segmentation and range assessment., *Phys. Med. Biol.* **63**: 115008.

Ambrosi, P., Borowski, M. & Iwatschenko, M. (2010). Considerations concerning the use of counting active personal dosemeters in pulsed fields of ionising radiation, *Radiat. Prot. Dosim.* **139**(4): 483–493.

Ammazzalorso, F., Jelen, U., Engenhart-Cabillic, R. & Schlegel, W. (2014). Dosimetric robustness against setup errors in charged particle radiotherapy of skull base tumors, *Radiat. Oncol.* **9**(1).

Ankerhold, U., Hupe, O. & Ambrosi, P. (2009). Deficiencies of active electronic radiation protection dosemeters in pulsed fields, *Radiat. Prot. Dosim.* **135**(3): 149–153.

Arbor, N., Dauvergne, D., Dedes, G., Létang, J. M., Parodi, K., Quiñones, C. T., Testa, E. & Rit, S. (2015). Monte Carlo comparison of x-ray and proton CT for range calculations of proton therapy beams., *Phys. Med. Biol.* **60**: 7585–7599.

Artikov, A., Budagov, J., Chirikov-Zorin, I., Chokheli, D., Lyablin, M., Bellettini, G., Menzione, A., Tokar, S., Giokaris, N. & Manousakis-Katsikakis, A. (2005). Properties of the Ukraine polystyrene-based plastic scintillator UPS 923a, *Nucl. Instrum. Methods Phys. Res., Sect. A* **555**(1-2): 125–131.

Bartolac, S., Graham, S., Siewerdsen, J. & Jaffray, D. (2011). Fluence field optimization for noise and dose objectives in CT., *Med. Phys.* **38 Suppl 1**: S2.

Bartolac, S. & Jaffray, D. (2013). Compensator models for fluence field modulated computed tomography., *Med. Phys.* **40**: 121909.

Bashkirov, V. A., Johnson, R. P., Sadrozinski, H. F.-W. & Schulte, R. W. (2016a). Development of proton computed tomography detectors for applications in hadron therapy, *Nucl. Instrum. Methods Phys. Res., Sect. A* **809**: 120–129.

Bashkirov, V. A., Schulte, R. W., Hurley, R. F., Johnson, R. P., Sadrozinski, H. F.-W., Zatserklyaniy, A., Plautz, T. & Giacometti, V. (2016). Novel scintillation detector design and performance for proton radiography and computed tomography., *Med. Phys.* **43**: 664–674.

Baudino, T. (2015). Targeted cancer therapy: The next generation of cancer treatment, *Curr. Drug Discov. Technol.* **12**(1): 3–20.

Bauer, J., Unholtz, D., Sommerer, F., Kurz, C., Haberer, T., Herfarth, K., Welzel, T., Combs, S. E., Debus, J. & Parodi, K. (2013). Implementation and initial clinical experience of offline PET/CT-based verification of scanned carbon ion treatment, *Radiother. Oncol.* **107**(2): 218–226.

Baumann, M., Krause, M., Overgaard, J., Debus, J., Bentzen, S. M., Daartz, J., Richter, C., Zips, D. & Bortfeld, T. (2016). Radiation oncology in the era of precision medicine, *Nat. Rev. Cancer* **16**(4): 234–249.

Bel, A., van Herk, M., Bartelink, H. & Lebesque, J. (1993). A verification procedure to improve patient set-up accuracy using portal images, *Radiother. Oncol.* **29**(2): 253–260.

Berger, M., Coursey, J., Zucker, M. & Chang, J. (2005). ESTAR, PSTAR, and ASTAR: Computer Programs for Calculating Stopping-Power and Range Tables for Electrons, Protons, and Helium Ions (version 1.2.3).

Berndt (2016). *DECT based tissue segmentation as input to Monte Carlo simulations for proton treatment verification using PET imaging*, Master's thesis, Department of Medical Physics, Ludwig–Maximilians–University of Munich, Garching, Germany.

Bernier, J., Hall, E. J. & Giaccia, A. (2004). Radiation oncology: a century of achievements, *Nat. Rev. Cancer* **4**(9): 737–747.

Bethe, H. (1930). Zur Theorie des Durchgangs schneller Korpuskularstrahlen durch Materie, *Ann. Phys-berlin.* **397**(3): 325–400.

BfUNnS (2018). Verordnung zum schutzvor der schädlichen wirkung ionisierender strahlung (strahlenschutzverordnung - strlschv, *Technical report*, Bundesministerium für Umwelt Naturschutz und nukleare Sicherheit.

Birks, J. B. (1951). Scintillations from Organic Crystals: Specific Fluorescence and Relative Response to Different Radiations, *Proc. Phys. Soc. Sect. A* **64**(10): 874–877.

Bohr, N. (1915). LX. on the decrease of velocity of swiftly moving electrified particles in passing through matter, *London, Edinburgh Dublin Philos. Mag. J. Sci.* **30**(178): 581–612.

Bopp, C. (2014). *The proton as a dosimetric and diagnostic probe*, PhD thesis, Université de Strasbourg.

Borne, F., Delacroix, D., Gel, J. M., Mass, D. & Amiranoff, F. (2002). Radiation protection for an ultra-high intensity laser, *Radiat. Prot. Dosim.* **102**(1): 61–70.

Bortfeld, T. (1997). An analytical approximation of the Bragg curve for therapeutic proton beams, *Med. Phys.* **24**(12): 2024–2033.

Bortfeld, T. (2006). IMRT: a review and preview, *Phys. Med. Biol.* **51**(13): R363–R379.

Bortfeld, T., Boyer, A. L., Schlegel, W., Kahler, D. L. & Waldron, T. J. (1994). Realization and verification of three-dimensional conformal radiotherapy with modulated fields, *Int. J. Radiat. Oncol.* **30**(4): 899–908.

Bourque, A. E., Carrier, J.-F. & Bouchard, H. (2014). A stoichiometric calibration method for dual energy computed tomography., *Phys. Med. Biol.* **59**: 2059–2088.

Bovik, A. (2009). *The Essential Guide to Image Processing*, ACADEMIC PR INC.

Bradt, H. L. & Peters, B. (1950). The heavy nuclei of the primary cosmic radiation, *Phys. Rev.* **77**(1): 54–70.

Bragg, W. H. & Kleeman, R. (1905). On the α particles of radium, and their loss of range in passing through various atoms and molecules, *London, Edinburgh Dublin Philos. Mag. J. Sci.* **10**(57): 318–340.

Brahme, A. (1987). Design principles and clinical possibilities with a new generation of radiation therapy equipment: A review, *Acta Oncol.* **26**(6): 403–412.

Braunn, B. (2010). *La fragmentation du 12C à 95 MeV par nucléon appliquée au domaine de la hadronthérapie. Etude expérimentale et simulations sur cibles épaisses de PMMA*, PhD thesis, Université de Caen.

Braunn, B., Labalme, M., Ban, G., Chevallier, M., Colin, J., Cussol, D., Dauvergne, D., Fontbonne, J., Haas, F., Guertin, A., Lebhertz, D., Foulher, F. L., Pautard, C., Ray, C., Rousseau, M., Salsac, M., Stuttge, L., Testa, E. & Testa, M. (2011). Nuclear reaction measurements of 95MeV/u 12C interactions on PMMA for hadrontherapy, *Nucl. Instrum. Methods Phys. Res. Sect. B Beam Interact. Mater. At.* **269**(22): 2676–2684.

Brousmiche, S., Souris, K., de Xivry, O., Lee, J., Macq, B. & Seco, J. (2015). Assessing how stochastic CT noise can lead to systematic proton range errors: TH-CD-BRA-04, *Med. Phys.* **42**(6).

Broyden, C. G. (1970). The Convergence of a Class of Double-rank Minimization Algorithms, *IMA J. Appl. Math.* **6**(3): 222–231.

Buzug, T. (2008). *Computed Tomography*, Springer-Verlag GmbH.

Bär, E., Lalonde, A., Zhang, R., Jee, K.-W., Yang, K., Sharp, G., Liu, B., Royle, G., Bouchard, H. & Lu, H.-M. (2018). Experimental validation of two dual-energy CT methods for proton therapy using heterogeneous tissue samples., *Med. Phys.* **45**: 48–59.

Böhlen, T., Cerutti, F., Chin, M., Fassò, A., Ferrari, A., Ortega, P., Mairani, A., Sala, P., Smirnov, G. & Vlachoudis, V. (2014). The FLUKA code: Developments and challenges for high energy and medical applications, *Nucl. Data Sheets* **120**: 211–214.

Carabe, A., España, S., Grassberger, C. & Paganetti, H. (2013). Clinical consequences of relative biological effectiveness variations in proton radiotherapy of the prostate, brain and liver., *Phys. Med. Biol.* **58**: 2103–2117.

Carabe, A., Moteabbed, M., Depauw, N., Schuemann, J. & Paganetti, H. (2012). Range uncertainty in proton therapy due to variable biological effectiveness., *Phys. Med. Biol.* **57**: 1159–1172.

Carabe-Fernandez, A., Dale, R. G. & Jones, B. (2007). The incorporation of the concept of minimum RBE (RBEmin) into the linear-quadratic model and the potential for improved radiobiological analysis of high-LET treatments., *Int. J. Radiat. Biol.* **83**: 27–39.

Casiraghi, M., Albertini, F. & Lomax, A. J. (2013). Advantages and limitations of the 'worst case scenario' approach in IMPT treatment planning, *Phys. Med. Biol.* **58**(5): 1323–1339.

Chan, T., Tan, P. W. & Tang, J. I. (2016). Proton therapy for early stage prostate cancer: is there a case?, *Onco. Targets Ther.* **Volume 9**: 5577–5586.

Chu, W. T., Ludewigt, B. A. & Renner, T. R. (1993). Instrumentation for treatment of cancer using proton and light-ion beams, *Rev. Sci. Instrum.* **64**(8): 2055–2122.

Chvetsov, A. V. & Paige, S. L. (2010). The influence of CT image noise on proton range calculation in radiotherapy planning, *Phys. Med. Biol.* **55**(6): N141.

Connell, P. P. & Hellman, S. (2009). Advances in radiotherapy and implications for the next century: A historical perspective, *Cancer Res.* **69**(2): 383–392.

Cooper, G. (2019). *The Cell*, Oxford University Press Inc.

Cormack, A. M. (1963). Representation of a function by its line integrals, with some radiological applications, *J. Appl. Phys.* **34**(9): 2722–2727.

Cortés-Giraldo, M. A. & Carabe, A. (2015). A critical study of different monte carlo scoring methods of dose average linear-energy-transfer maps calculated in voxelized geometries irradiated with clinical proton beams, *Phys. Med. Biol.* **60**(7): 2645–2669.

Coutrakon, G., Bashkirov, V., Hurley, F., Johnson, R., Rykalin, V., Sadrozinski, H. & Schulte, R. (2013). Design and construction of the 1st proton CT scanner, *AIP Conference Proceedings 1525, 327 (2013)*.

Daido, H., Nishiuchi, M. & Pirozhkov, A. S. (2012). Review of laser-driven ion sources and their applications, *Rep. Prog. Phys.* **75**(5): 056401.

Dasu, A. & Toma-Dasu, I. (2013). Impact of variable RBE on proton fractionation., *Med. Phys.* **40**: 011705.

Deasy, J. O., Blanco, A. I. & Clark, V. H. (2003). CERR: A computational environment for radiotherapy research, *Med. Phys.* **30**(5): 979–985.

Debus, J., Hug, E., Liebsch, N., OFarrel, D., Finkelstein, D., Efird, J. & Munzenrider, J. (1997). Brainstem tolerance to conformal radiotherapy of skull base tumors, *Int. J. Radiat. Oncol.* **39**(5): 967–975.

Dedes, G., De Angelis, L., Rit, S., Hansen, D., Belka, C., Bashkirov, V., Johnson, R. P., Coutrakon, G., Schubert, K. E., Schulte, R. W., Parodi, K. & Landry, G. (2017). Application of fluence field modulation to proton computed tomography for proton therapy imaging., *Phys. Med. Biol.* **62**: 6026–6043.

Dedes, G., Dickmann, J., Giacometti, V., Rit, S., Krah, N., Meyer, S., Bashkirov, V., Schulte, R., Johnson, R. P., Parodi, K. & Landry, G. (2020). The role of Monte Carlo simulation in understanding the performance of proton computed tomography, *Z. Med. Phys.* .

Dedes, G., Dickmann, J., Niepel, K., Wesp, P., Johnson, R. P., Pankuch, M., Bashkirov, V., Rit, S., Volz, L., Schulte, R. W., Landry, G. & Parodi, K. (2019). Experimental comparison of proton CT and dual energy x-ray CT for relative stopping power estimation in proton therapy., *Phys. Med. Biol.* **64**: 165002.

Dedes, G., Johnson, R. P., Pankuch, M., Detrich, N., Pols, W. M. A., Rit, S., Schulte, R. W., Parodi, K. & Landry, G. (2018). Experimental fluence-modulated proton computed tomography by pencil beam scanning., *Med. Phys.* **45**: 3287–3296.

Dedes, G., Pinto, M., Dauvergne, D., Freud, N., Krimmer, J., Létang, J. M., Ray, C. & Testa, E. (2014). Assessment and improvements of Geant4 hadronic models in the context of prompt-gamma hadrontherapy monitoring., *Phys. Med. Biol.* **59**: 1747–1772.

DeVita, V. T. & Chu, E. (2008). A history of cancer chemotherapy, *Cancer Res.* **68**(21): 8643–8653.

Dickmann, J., Rit, S., Pankuch, M., Johnson, R. P., Schulte, R. W., Parodi, K., Dedes, G. & Landry, G. (2020). An optimization algorithm for dose reduction with fluence-modulated proton CT, *Med. Phys.* **47**(4): 1895–1906.

Dickmann, J., Sarosiek, C., Rykalin, V., Pankuch, M., Coutrakon, G., Johnson, R. P., Bashkirov, V., Schulte, R. W., Parodi, K., Landry, G. & Dedes, G. (2021). Proof of concept image artifact reduction by energy-modulated proton computed tomography (EMpCT), *Phys. Med.* **81**: 237–244.

Dickmann, J., Sarosiek, C., Rykalin, V., Pankuch, M., Rit, S., Detrich, N., Coutrakon, G., Johnson, R. P., Schulte, R. W., Parodi, K., Landry, G. & Dedes, G. (2020a). Experimental realization of dynamic fluence field optimization for proton computed tomography, *Phys. Med. Biol.* **65**(19): 195001.

Dickmann, J., Wesp, P., Rädler, M., Rit, S., Pankuch, M., Johnson, R. P., Bashkirov, V., Schulte, R. W., Parodi, K., Landry, G. & Dedes, G. (2019). Prediction of image noise contributions in proton computed tomography and comparison to measurements, *Phys. Med. Biol.* **64**(14): 145016.

Durante, M. & Paganetti, H. (2016). Nuclear physics in particle therapy: a review., *Rep. Prog. Phys.* **79**: 096702.

Döpp, A., Hehn, L., Götzfried, J., Wenz, J., Gilljohann, M., Ding, H., Schindler, S., Pfeiffer, F. & Karsch, S. (2018). Quick x-ray microtomography using a laser-driven betatron source, *Optica* **5**(2): 199.

Engelsman, M., Schwarz, M. & Dong, L. (2013). Physics controversies in proton therapy, *Semin. Radiat. Oncol.* **23**(2): 88–96.

Enghardt, W., Crespo, P., Fiedler, F., Hinz, R., Parodi, K., Pawelke, J. & Pönisch, F. (2004). Charged hadron tumour therapy monitoring by means of PET, *Nucl. Instrum. Meth. A.* **525**(1-2): 284–288.

Englbrecht, F. S., Döpp, A., Hartmann, J., Lindner, F. H., Groß, M. L., Wirth, H.-F., Thirolf, P. G., Karsch, S., Schreiber, J., Parodi, K. & Dedes, G. (2020). Radiation protection modelling for 2.5 petawatt-laser production of ultrashort x-ray, proton and ion bunches: Monte carlo model of the munich CALA facility, *J. Radiol. Prot.* **40**(4): 1048–1073.

Esarey, E., Schroeder, C. B. & Leemans, W. P. (2009). Physics of laser-driven plasma-based electron accelerators, *Rev. Mod. Phys.* **81**(3): 1229–1285.

Esposito, M., Waltham, C., Taylor, J. T., Manger, S., Phoenix, B., Price, T., Poludniowski, G., Green, S., Evans, P. M., Allport, P. P., Manolopulos, S., Nieto-Camero, J., Symons, J. & Allinson, N. M. (2018). PRaVDA: The first solid-state system for proton computed tomography., *Phys. Med.* **55**: 149–154.

Eyges, L. (1948). Multiple Scattering with Energy Loss, *Phys. Rev.* **74**: 1534–1535.

Farace, P., Righetto, R. & Meijers, A. (2016). Pencil beam proton radiography using a multilayer ionization chamber., *Phys. Med. Biol.* **61**: 4078–4087.

Fass, L. (2008). Imaging and cancer: A review, *Mol. Oncol.* **2**(2): 115–152.

Fekete, C.-A. C., Doolan, P., Dias, M. F., Beaulieu, L. & Seco, J. (2015). Developing a phenomenological model of the proton trajectory within a heterogeneous medium required for proton imaging., *Phys. Med. Biol.* **60**: 5071–5082.

Fermi, E. (1940). The ionization loss of energy in gases and in condensed materials, *Phys. Rev.* **57**: 485–493.

Ferrari, A., Sala, P., Fasso, A. & Ranft, J. (2005). FLUKA: A multi-particle transport code, *Technical report*.

Folger, G., Ivanchenko, V. N. & Wellisch, J. P. (2004). The binary cascade, *Eur. Phys. J. A* **21**(3): 407–417.

Forster (2020).
URL: *www.forster-bau.de/sandwich-construction.html*

Fowler, J., Chappell, R. & Ritter, M. (2001). Is α/β for prostate tumors really low?, *Int. J. Radiat. Oncol.* **50**(4): 1021–1031.

Frandes, M., Zoglauer, A., Maxim, V. & Prost, R. (2010). A tracking Compton-scattering imaging system for hadron therapy monitoring, *IEEE Trans. Nucl. Sci.* **57**(1): 144–150.

Frese, M. C., Wilkens, J. J., Huber, P. E., Jensen, A. D., Oelfke, U. & Taheri-Kadkhoda, Z. (2011). Application of constant vs. variable relative biological effectiveness in treatment planning of intensity-modulated proton therapy, *Int. J. Radiat. Oncol.* **79**(1): 80–88.

Frey, K., Unholtz, D., Bauer, J., Debus, J., Min, C. H., Bortfeld, T., Paganetti, H. & Parodi, K. (2014). Automation and uncertainty analysis of a method for in-vivo range verification in particle therapy, *Phys. Med. Biol.* **59**(19): 5903–5919.

Friedrich, T., Scholz, U., Durante, M. & Scholz, M. (2014). RBE of ion beams in hypofractionated radiotherapy (SBRT), *Phys. Med.* **30**(5): 588–591.

Fry, D. W., Shersby-Harvie, R. B. R., Mullett, L. B. & Walkinshaw, W. (1948). A travelling-wave linear accelerator for 4-MeV. electrons, *Nature* **162**(4126): 859–861.

Fuji, H., Schneider, U., Ishida, Y., Konno, M., Yamashita, H., Kase, Y., Murayama, S., Onoe, T., Ogawa, H., Harada, H., Asakura, H. & Nishimura, T. (2013). Assessment of organ dose reduction and secondary cancer risk associated with the use of proton beam therapy and intensity modulated radiation therapy in treatment of neuroblastomas., *Radiat. Oncol.* **8**: 255.

Gaillard, S. A., Kluge, T., Flippo, K. A., Bussmann, M., Gall, B., Lockard, T., Geissel, M., Offermann, D. T., Schollmeier, M., Sentoku, Y. & Cowan, T. E. (2011). Increased laser-accelerated proton energies via direct laser-light-pressure acceleration of electrons in microcone targets, *Phys. Plasmas* **18**(5): 056710.

Gang, G. J., Mao, A., Wang, W., Siewerdsen, J. H., Mathews, A., Kawamoto, S., Levinson, R. & Stayman, J. W. (2019). Dynamic fluence field modulation in computed tomography using multiple aperture devices., *Phys. Med. Biol.* **64**: 105024.

Gang, G. J., Siewerdsen, J. H. & Stayman, J. W. (2017). Task-driven optimization of fluence field and regularization for model-based iterative reconstruction in computed tomography., *IEEE Trans. Med. Imaging* **36**: 2424–2435.

Gasques, L. R., Chamon, L. C., Silva, C. P., Pereira, D., Alvarez, M. A. G., Rossi, E. S., Likhachev, V. P., Carlson, B. V. & Conti, C. D. (2002). Determination of the 12C nuclear density through heavy-ion elastic scattering experiments, *Phys. Rev. C* **65**(4).

GEANT-Collaboration et al. (2016). *Physics reference manual*.

Giacometti, V., Bashkirov, V. A., Piersimoni, P., Guatelli, S., Plautz, T. E., Sadrozinski, H. F.-W., Johnson, R. P., Zatserklyaniy, A., Tessonnier, T., Parodi, K., Rosenfeld, A. B. & Schulte, R. W. (2017a). Software platform for simulation of a prototype proton CT scanner., *Med. Phys.* **44**: 1002–1016.

Giacometti, V., Guatelli, S., Bazalova-Carter, M., Rosenfeld, A. & Schulte, R. (2017b). Development of a high resolution voxelised head phantom for medical physics applications, *Phys. Med.* **33**: 182–188.

Giovannini, G., Böhlen, T., Cabal, G., Bauer, J., Tessonnier, T., Frey, K., Debus, J., Mairani, A. & Parodi, K. (2016). Variable RBE in proton therapy: comparison of different model predictions and their influence on clinical-like scenarios, *Radiat. Oncol.* **11**(1).

Golnik, C., Hueso-González, F., Müller, A., Dendooven, P., Enghardt, W., Fiedler, F., Kormoll, T., Roemer, K., Petzoldt, J., Wagner, A. & Pausch, G. (2014). Range assessment in particle therapy based on prompt γ-ray timing measurements, *Phys. Med. Biol.* **59**(18): 5399–5422.

Gonsalves, A. J., Nakamura, K., Daniels, J., Mao, H.-S., Benedetti, C., Schroeder, C. B., Tóth, C., van Tilborg, J., Mittelberger, D. E., Bulanov, S. S., Vay, J.-L., Geddes, C.

G. R., Esarey, E. & Leemans, W. P. (2015). Generation and pointing stabilization of multi-GeV electron beams from a laser plasma accelerator driven in a pre-formed plasma waveguide, *Phys. Plasmas* **22**(5): 056703.

Gore, J. C. & Tofts, P. S. (1978). Statistical limitations in computed tomography, *Phys. Med. Biol.* **23**(6): 1176–1182.

Gottschalk, B. (2010). On the scattering power of radiotherapy protons., *Med. Phys.* **37**: 352–367.

Gottschalk, B. (2012). Techniques of proton radiotherapy: Transport theory, *ArXiv* .

Gottschalk, B. (2018). Radiotherapy proton interactions in matter, *ArXiv* .

Gottschalk, B., Koehler, A., Schneider, R., Sisterson, J. & Wagner, M. (1993). Multiple Coulomb scattering of 160 MeV protons, *Nucl. Instrum. Methods Phys. Res. Sect. B Beam Interact. Mater. At.* **74**(4): 467–490.

Graham, S., Moseley, D., Siewerdsen, J. & Jaffray, D. (2007a). Compensators for dose and scatter management in cone-beam computed tomography, *Med. Phys.* **34**(7): 2691–2703.

Graham, S., Siewerdsen, J., Jaffray, D., Graham S A, S. J. H. & A, J. D. (2007b). Intensity-modulated fluence patterns for task-specific imaging in cone-beam CT, *Proc. Spie* **6510**(651003).

Grassberger, C. & Paganetti, H. (2011). Elevated LET components in clinical proton beams, *Phys. Med. Biol.* **56**(20): 6677–6691.

Grassberger, C., Trofimov, A., Lomax, A. & Paganetti, H. (2011). Variations in linear energy transfer within clinical proton therapy fields and the potential for biological treatment planning, *Int. J. Radiat. Oncol.* **80**(5): 1559–1566.

Grevillot, L., Frisson, T., Zahra, N., Bertrand, D., Stichelbaut, F., Freud, N. & Sarrut, D. (2010). Optimization of GEANT4 settings for proton pencil beam scanning simulations using GATE, *Nucl. Instrum. Methods Phys. Res., Sect. B* **268**(20): 3295–3305.

Grubbé, E. H. (1933). Priority in the therapeutic use of x-rays, *Radiology* **21**(2): 156–162.

Grün, R., Friedrich, T., Krämer, M., Zink, K., Durante, M., Engenhart-Cabillic, R. & Scholz, M. (2013). Physical and biological factors determining the effective proton range., *Med. Phys.* **40**: 111716.

Gueth, P., Dauvergne, D., Freud, N., Létang, J. M., Ray, C., Testa, E. & Sarrut, D. (2013). Machine learning-based patient specific prompt-gamma dose monitoring in proton therapy, *Phys. Med. Biol.* **58**(13): 4563–4577.

Gunderson, L. (2016). *Clinical radiation oncology*, Elsevier, Philadelphia, PA.

Gwosch, K., Hartmann, B., Jakubek, J., Granja, C., Soukup, P., Jäkel, O. & Martišíková, M. (2013). Non-invasive monitoring of therapeutic carbon ion beams in a homogeneous phantom by tracking of secondary ions, *Phys. Med. Biol.* **58**(11): 3755–3773.

Habs, D., Thirolf, P. G., Gross, M., Allinger, K., Bin, J., Henig, A., Kiefer, D., Ma, W. & Schreiber, J. (2010). Introducing the fission–fusion reaction process: using a laser-accelerated th beam to produce neutron-rich nuclei towards the n=126 waiting point of the r-process, *Appl. Phys. B* **103**(2): 471–484.

Han, D., Siebers, J. V. & Williamson, J. F. (2016). A linear, separable two-parameter model for dual energy CT imaging of proton stopping power computation., *Med. Phys.* **43**: 600.

Hansen, D. C., Bassler, N., Sørensen, T. S. & Seco, J. (2014). The image quality of ion computed tomography at clinical imaging dose levels, *Med. Phys.* **41**(11): 111908.

Hansen, D. C., Sangild Sørensen, T. & Rit, S. (2016). Fast reconstruction of low dose proton CT by sinogram interpolation., *Phys. Med. Biol.* **61**: 5868–5882.

Hansen, D. C., Seco, J., Sørensen, T. S., Petersen, J. B. B., Wildberger, J. E., Verhaegen, F. & Landry, G. (2015). A simulation study on proton computed tomography (CT) stopping power accuracy using dual energy CT scans as benchmark, *Acta Oncol.* **54**(9): 1638–1642.

Hansen, D., Sorensen, T. & Seco, J. (2013). WE-g-141-05: The image quality of ion computed tomography at clinical imaging dose levels, *Med. Phys.* **40**(6Part30): 508–508.

Hanson, K. M., Bradbury, J. N., Cannon, T. M., Hutson, R. L., Laubacher, D. B., Macek, R., Paciotti, M. A. & Taylor, C. A. (1977). Application of protons to computer tomography, Nuclear science symposium, San Francisco, CA, USA, 19 Oct 1977.

Harpen, M. D. (1999). A simple theorem relating noise and patient dose in computed tomography., *Med. Phys.* **26**: 2231–2234.

Hartnack, C., Puri, R. K., Aichelin, J., Konopka, J., Bass, S., Stöcker, H. & Greiner, W. (1998). Modelling the many-body dynamics of heavy ion collisions: Present status and future perspective, *Eur. Phys. J. A* **1**(2): 151–169.

Helmbrecht, S., Santiago, A., Enghardt, W., Kuess, P. & Fiedler, F. (2012). On the feasibility of automatic detection of range deviations from in-beam PET data, *Phys. Med. Biol.* **57**(5): 1387–1397.

Henriquet, P., Testa, E., Chevallier, M., Dauvergne, D., Dedes, G., Freud, N., Krimmer, J., Létang, J. M., Ray, C., Richard, M.-H. & Sauli, F. (2012). Interaction vertex imaging (IVI) for carbon ion therapy monitoring: a feasibility study., *Phys. Med. Biol.* **57**: 4655–4669.

Hogstrom, K. R. & Almond, P. R. (2006). Review of electron beam therapy physics, *Phys. Med. Biol.* **51**(13): R455–R489.

Hopfgartner, J., Stock, M., Knäusl, B. & Georg, D. (2012). Robustness of IMPT treatment plans with respect to inter-fractional set-up uncertainties: Impact of various beam arrangements for cranial targets, *Acta Oncol.* **52**(3): 570–579.

Hounsfield, G. N. (1973). Computerized transverse axial scanning (tomography): Part 1. description of system, *Brit. J. Radiol.* **46**(552): 1016–1022.

Hsi, W. C., Indelicato, D. J., Vargas, C., Duvvuri, S., Li, Z. & Palta, J. (2009). In vivo verification of proton beam path by using post-treatment PET/CT imaging, *Med. Phys.* **36**(9Part1): 4136–4146.

Hsieh, S. S. & Pelc, N. J. (2014). Control algorithms for dynamic attenuators., *Med. Phys.* **41**: 061907.

Huck, S. M., Fung, G. S. K., Parodi, K. & Stierstorfer, K. (2019). Technical note: Sheet-based dynamic beam attenuator - a novel concept for dynamic fluence field modulation in x-ray CT., *Med. Phys.* **46**: 5528–5537.

Huck, S. M., Parodi, K., Stierstorfer, K. & Fung, G. S. K. (2019). Optimized intensity modulation for a dynamic beam attenuator in x-ray computed tomography, *in* H. Bosmans, G.-H. Chen & T. G. Schmidt (eds), *Medical Imaging 2019: Physics of Medical Imaging*, SPIE.

Hudobivnik, N., Schwarz, F., Johnson, T., Agolli, L., Dedes, G., Tessonnier, T., Verhaegen, F., Thieke, C., Belka, C., Sommer, W. H., Parodi, K. & Landry, G. (2016). Comparison of proton therapy treatment planning for head tumors with a pencil beam algorithm on dual and single energy CT images., *Med. Phys.* **43**: 495.

Huesman, R., Gullberg, G., Greenberg, W. & Budinger, T. (1977). Users manual: Donner algorithms for reconstruction tomography, *Technical report*. (Berkely, CA: Lawrence Laboratory).

Huesman, R. H. (1984). A new fast algorithm for the evaluation of regions of interest and statistical uncertainty in computed tomography., *Phys. Med. Biol.* **29**: 543–552.

Huesman, R., Rosenfeld, A. & Solmitz, F. (1975). Comparison of heavy charged particles and x-rays for axial tomograpic scanning, *Technical Report* .

Hueso-González, F., Enghardt, W., Fiedler, F., Golnik, C., Janssens, G., Petzoldt, J., Prieels, D., Priegnitz, M., Römer, K. E., Smeets, J., Stappen, F. V., Wagner, A. & Pausch, G. (2015). First test of the prompt gamma ray timing method with heterogeneous targets at a clinical proton therapy facility, *Phys. Med. Biol.* **60**(16): 6247–6272.

Hünemohr, N., Krauss, B., Dinkel, J., Gillmann, C., Ackermann, B., Jäkel, O. & Greilich, S. (2013). Ion range estimation by using dual energy computed tomography, *Z Med Phys* **23**(4): 300–313.

Hünemohr, N., Krauss, B., Tremmel, C., Ackermann, B., Jäkel, O. & Greilich, S. (2013a). Experimental verification of ion stopping power prediction from dual energy CT data in tissue surrogates, *Phys. Med. Biol.* **59**(1): 83.

Hurley, R., Schulte, R., Bashkirov, V., Wroe, A., Ghebremedhin, A., Sadrozinski, H.-W., Rykalin, V., Coutrakon, G., Koss, P. & Patyal, B. (2012). Water-equivalent path length calibration of a prototype proton CT scanner, *Med. Phys.* **39**(5): 2438–2446.

ICRP (2007). The 2007 recommendations of the international commission on radiological protection. ICRP publication 103, *Annals of the ICRP* **37**(2-4): 9–34.

ICRU49 (1993). ICRU report 49, S.P. ranges for protons and alpha particles., *Technical report*.

ICRU78 (2007). *International Commission on Radiation Units and Measurements, I. ICRU report 78, Prescribing, recording, and reporting proton-beam therapy*, Oxford University Press.

Ivanchenko, V., Kadri, O., Maire, M. & Urban, L. (2010). Geant4 models for simulation of multiple scattering, *J. Phys. Conf. Ser.*, Vol. 219, IOP Publishing, p. 032045.

Jagt, T. Z., Breedveld, S., van Haveren, R., Heijmen, B. J. & Hoogeman, M. S. (2020). Online-adaptive versus robust IMPT for prostate cancer: How much can we gain?, *Radiother. Oncol.* **151**: 228–233.

Janssen, F. M. F. C., Landry, G., Cambraia Lopes, P., Dedes, G., Smeets, J., Schaart, D. R., Parodi, K. & Verhaegen, F. (2014). Factors influencing the accuracy of beam range estimation in proton therapy using prompt gamma emission., *Phys. Med. Biol.* **59**: 4427–4441.

Jarlskog, C. Z. & Paganetti, H. (2008). Physics settings for using the Geant4 toolkit in proton therapy, *IEEE Trans. Nucl. Sci.* **55**(3): 1018–1025.

Jermann, M. (2016). Particle therapy patient statistics (per end of 2015).
URL: *https://www.ptcog.ch/archive/patientstatistics/Patientstatistics-updateDec2015.pdf*

Johnson, R. P. (2017). Review of medical radiography and tomography with proton beams, *Rep. Prog. Phys.* **81**(1): 016701.

Johnson, R. P., Bashkirov, V., Coutrakon, G., Giacometti, V., Karbasi, P., Karonis, N., Ordoñez, C., Pankuch, M., Sadrozinski, H.-W., Schubert, K. et al. (2017). Results from a prototype proton-CT head scanner, *Physcs. Proc.* **90**: 209–214.

Johnson, R. P., Bashkirov, V., DeWitt, L., Giacometti, V., Hurley, R. F., Piersimoni, P., Plautz, T. E., Sadrozinski, H. F.-., Schubert, K., Schulte, R., Schultze, B. & Zatserklyaniy, A. (2016). A fast experimental scanner for proton CT: Technical performance and first experience with phantom scans, *IEEE Trans. Nucl. Sci.* **63**(1): 52–60.

Joiner, M. C. & van der Kogel, A. (2009). *Basic Clinical Radiobiology*, CRC Press.

Jones, D. & Wambersie, A. (2007). Radiation therapy with fast neutrons: A review, *Nucl. Instrum. Meth. A.* **580**(1): 522–525.

Joseph, P. M. (1982). An Improved Algorithm for Reprojecting Rays through Pixel Images, *IEEE Trans. Med. Imaging* **1**(3): 192–196.
URL: *http://ieeexplore.ieee.org/document/4307572/*

Jäkel, O. (2008). The relative biological effectiveness of proton and ion beams., *Z. Med. Phys.* **18**: 276–285.

Kachelrieß, M., Sourbelle, K. & Kalender, W. A. (2006). Empirical cupping correction: A first-order raw data precorrection for cone-beam computed tomography, *Med. Phys.* **33**(5): 1269–1274.

Kaiser, A., Eley, J. G., Onyeuku, N. E., Rice, S. R., Wright, C. C., McGovern, N. E., Sank, M., Zhu, M., Vujaskovic, Z., Simone, C. B. & Hussain, A. (2019). Proton therapy delivery and its clinical application in select solid tumor malignancies, *J. Visualized Exp.* (144).

Kak, A. & Slaney, M. (1988). *Principles of computerized tomographic imaging*, IEEE press.

Kalender, W. A. (2006). X-ray computed tomography, *Phys. Med. Biol.* **51**(13): R29–R43.

Kalender, W. A., Buchenau, S., Deak, P., Kellermeier, M., Langner, O., van Straten, M., Vollmar, S. & Wilharm, S. (2008). Technical approaches to the optimisation of CT., *Phys. Med.* **24**: 71–79.

Kamran, S. C., Light, J. O. & Efstathiou, J. A. (2019). Proton versus photon-based radiation therapy for prostate cancer: emerging evidence and considerations in the era of value-based cancer care, *Prostate Cancer P. D.* **22**(4): 509–521.

Kempe, J., Gudowska, I. & Brahme, A. (2006). Depth absorbed dose and LET distributions of therapeutic H1, He4, Li7, and C12 beams, *Med. Phys.* **34**(1): 183–192.

Kim, H., Pyo, H., Noh, J. M., Lee, W., Park, B., Park, H. Y. & Yoo, H. (2019). Preliminary result of definitive radiotherapy in patients with non-small cell lung cancer who have underlying idiopathic pulmonary fibrosis: comparison between x-ray and proton therapy., *Radiat. Oncol.* **14**: 19.

Kjellberg, R., Sweet, W., Preston, W. & Koehler, A. (1962). The Bragg peak of a proton beam in intracranial therapy of tumors, *T. Am. Neurol. Assoc.* **87**.

Knopf, A.-C. & Lomax, A. (2013). In vivo proton range verification: a review, *Phys. Med. Biol.* **58**(15): R131–R160.

Knopf, A., Parodi, K., Bortfeld, T., Shih, H. A. & Paganetti, H. (2009). Systematic analysis of biological and physical limitations of proton beam range verification with offline PET/CT scans, *Phys. Med. Biol.* **54**(14): 4477–4495.

Knopf, A., Parodi, K., Paganetti, H., Cascio, E., Bonab, A. & Bortfeld, T. (2008). Quantitative assessment of the physical potential of proton beam range verification with PET/CT, *Phys. Med. Biol.* **53**(15): 4137–4151.

Koi, T. (2010). New native QMD code in Geant4, *Proceedings of the MC2010 Monte Carlo Conference* .

Koivula, L., Wee, L. & Korhonen, J. (2016). Feasibility of MRI-only treatment planning for proton therapy in brain and prostate cancers: Dose calculation accuracy in substitute CT images., *Med. Phys.* **43**: 4634.

Kormoll, T., Fiedler, F., Schöne, S., Wüstemann, J., Zuber, K. & Enghardt, W. (2011). A compton imager for in-vivo dosimetry of proton beams—a design study, *Nucl. Instrum. Meth. A.* **626-627**: 114–119.

Kozlovsky, B., Murphy, R. J. & Ramaty, R. (2002). Nuclear deexcitation gamma-ray lines from accelerated particle interactions, *Astrophys. J. Suppl. Ser.* **141**(2): 523–541.

Krah, N., Khellaf, F., Létang, J. M., Rit, S. & Rinaldi, I. (2018). A comprehensive theoretical comparison of proton imaging set-ups in terms of spatial resolution, *Phys. Med. Biol.* **63**(13): 135013.

Krah, N., Testa, M., Brons, S., Jäkel, O., Parodi, K., Voss, B. & Rinaldi, I. (2015). An advanced image processing method to improve the spatial resolution of ion radiographies, *Phys. Med. Biol.* **60**(21): 8525.

Kuperman, V. Y. (2016). Cell kill by megavoltage protons with high LET, *Phys. Med. Biol.* **61**(14): 5183–5197.

Kurz, C., Dedes, G., Resch, A., Reiner, M., Ganswindt, U., Nijhuis, R., Thieke, C., Belka, C., Parodi, K. & Landry, G. (2015). Comparing cone-beam CT intensity correction methods for dose recalculation in adaptive intensity-modulated photon and proton therapy for head and neck cancer., *Acta Oncol.* **54**: 1651–1657.

Lalonde, A., Bär, E. & Bouchard, H. (2017). A Bayesian approach to solve proton stopping powers from noisy multi-energy CT data, *Med. Phys.* **44**(10): 5293–5302.

Landau, L. D. (1944). On the energy loss of fast particles by ionization, *J. Phys.* **8**(4): 201–205.
URL: *https://cds.cern.ch/record/216256*

Landry, G., Dedes, G., Zöllner, C., Handrack, J., Janssens, G., Orban de Xivry, J., Reiner, M., Paganelli, C., Riboldi, M., Kamp, F., Söhn, M., Wilkens, J. J., Baroni, G., Belka, C. & Parodi, K. (2015a). Phantom based evaluation of CT to CBCT image registration for proton therapy dose recalculation., *Phys. Med. Biol.* **60**: 595–613.

Landry, G., Gaudreault, M., van Elmpt, W., Wildberger, J. E. & Verhaegen, F. (2016). Improved dose calculation accuracy for low energy brachytherapy by optimizing dual energy CT imaging protocols for noise reduction using sinogram affirmed iterative reconstruction, *Z. Med. Phys.* **26**(1): 75–87.

Landry, G. & Hua, C.-H. (2018). Current state and future applications of radiological image guidance for particle therapy., *Med. Phys.* **45**: e1086–e1095.

Landry, G., Nijhuis, R., Dedes, G., Handrack, J., Thieke, C., Janssens, G., Orban de Xivry, J., Reiner, M., Kamp, F., Wilkens, J. J., Paganelli, C., Riboldi, M., Baroni, G., Ganswindt, U., Belka, C. & Parodi, K. (2015b). Investigating CT to CBCT image registration for head and neck proton therapy as a tool for daily dose recalculation., *Med. Phys.* **42**: 1354–1366.

Larsson, B., Leksell, L., Rexed, B., Sourander, P., Mair, W. & Andersson, B. (1958). The high-energy proton beam as a neurosurgical tool, *Cah. Rev. The.* **182**(4644): 1222–1223.

Leahy, R., Clackdoyle, R. & Noo, F. (2009). Computed tomography, *The Essential Guide to Image Processing*, Elsevier, pp. 741–776.

Lee, H. H., Li, B., Duan, X., Zhou, L., Jia, X. & Yang, M. (2019). Systematic analysis of the impact of imaging noise on dual-energy CT-based proton stopping-power-ratio estimation, *Med. Phys.* **46**: 2251–2263.

Lee, J., Kim, C., Min, B., Kwak, J., Park, S., Lee, S. B., Park, S. & Cho, S. (2015). Sparse-view proton computed tomography using modulated proton beams., *Med. Phys.* **42**: 1129–1137.

Lell, M. M. & Kachelrieß, M. (2020). Recent and upcoming technological developments in computed tomography: High speed, low dose, deep learning, multienergy., *Invest. Radiol.* **55**: 8–19.

Lestand, L., Montarou, G., Force, P. & Pauna, N. (2012). In-beam quality assurance using induced $\beta+$ ctivity in hadrontherapy: a preliminary physical requirements study using Geant4, *Phys. Med. Biol.* **57**(20): 6497–6518.

Li, T., Liang, Z., Singanallur, J. V., Satogata, T. J., Williams, D. C. & Schulte, R. W. (2006). Reconstruction for proton computed tomography by tracing proton trajectories: a Monte Carlo study., *Med. Phys.* **33**: 699–706.

Lindner, F. H., Bin, J. H., Englbrecht, F., Haffa, D., Bolton, P. R., Gao, Y., Hartmann, J., Hilz, P., Kreuzer, C., Ostermayr, T. M., Rösch, T. F., Speicher, M., Parodi, K., Thirolf, P. G. & Schreiber, J. (2018). A novel approach to electron data background treatment in an online wide-angle spectrometer for laser-accelerated ion and electron bunches, *Rev. Sci. Instrum.* **89**(1): 013301.

Lindner, F., Haffa, D., Bin, J., Englbrecht, F., Gao, Y., Gebhard, J., Hartmann, J., Hilz, P., Kreuzer, C., Lehrack, S., Ostermayr, T., Rösch, T., Speicher, M., Würl, M., Parodi, K., Schreiber, J. & Thirolf, P. (2017). Towards swift ion bunch acceleration by high-power laser pulses at the centre for advanced laser applications (CALA), *Nucl. Instrum. Methods Phys. Res. Sect. B Beam Interact. Mater. At.* **402**: 354–357.

Linz, U. & Alonso, J. (2016). Laser-driven ion accelerators for tumor therapy revisited, *Phys. Rev. Accel. Beams* **19**(12).

Lomax, A. J. (2008). Intensity modulated proton therapy and its sensitivity to treatment uncertainties 2: the potential effects of inter-fraction and inter-field motions, *Phys. Med. Biol.* **53**(4): 1043–1056.

Lomax, A. J., Bortfeld, T., Goitein, G., Debus, J., Dykstra, C., Tercier, P. A., Coucke, P. A. & Mirimanoff, R. O. (1999). A treatment planning inter-comparison of proton and intensity modulated photon radiotherapy., *Radiother. Oncol.* **51**: 257–271.

Lopes, P. C., Pinto, M., Simoes, H., Biegun, A. K., Dendooven, P., Oxley, D. C., Parodi, K., Schaart, D. R. & Crespo, P. (2012). Optimization of collimator designs for real-time proton range verification by measuring prompt gamma rays, *2012 IEEE Nuclear Science Symposium and Medical Imaging Conference Record (NSS/MIC)*, IEEE.

Lu, W., Tzoufras, M., Joshi, C., Tsung, F. S., Mori, W. B., Vieira, J., Fonseca, R. A. & Silva, L. O. (2007). Generating multi-GeV electron bunches using single stage laser wakefield acceleration in a 3D nonlinear regime, *Physical Review Special Topics - Accelerators and Beams* **10**(6).

Macchi, A., Borghesi, M. & Passoni, M. (2013). Ion acceleration by superintense laser-plasma interaction, *Rev. Mod. Phys.* **85**(2): 751–793.

Mackie, T. R., Holmes, T., Swerdloff, S., Reckwerdt, P., Deasy, J. O., Yang, J., Paliwal, B. & Kinsella, T. (1993). Tomotherapy: A new concept for the delivery of dynamic conformal radiotherapy, *Med. Phys.* **20**(6): 1709–1719.

Magallanes, H. L. (2017). *Low-dose ion-based transmission radiography and tomography for optimization of carbon ion-beam therapy*, PhD thesis, Faculty of Physics, Ludwig-Maximilians-Universität München.

Mail, N., Moseley, D. J., Siewerdsen, J. H. & Jaffray, D. A. (2009). The influence of bowtie filtration on cone-beam CT image quality., *Med. Phys.* **36**: 22–32.

Mairani, A., Böhlen, T. T., Schiavi, A., Tessonnier, T., Molinelli, S., Brons, S., Battistoni, G., Parodi, K. & Patera, V. (2013). A Monte Carlo-based treatment planning tool for proton therapy, *Phys. Med. Biol.* **58**(8): 2471–2490.

Mao, A., Gang, G. J., Shyr, W., Levinson, R., Siewerdsen, J. H., Kawamoto, S. & Webster Stayman, J. (2018). Dynamic fluence field modulation for miscentered patients in computed tomography., *J. Med. Imaging* **5**: 043501.

Martin, C. (2010). Methods for Evaluating Infinite Sums. Online available at `http://web.math.ucsb.edu/~cmart07/Evaluating%20Series.pdf`.

Maruyama, T., Niita, K., Oyamatsu, K., Maruyama, T., Chiba, S. & Iwamoto, A. (1998). Quantum molecular dynamics approach to the nuclear matter below the saturation density, *Phys. Rev. C* **57**(2): 655–665.

Maspero, M., van den Berg, C. A. T., Landry, G., Belka, C., Parodi, K., Seevinck, P. R., Raaymakers, B. W. & Kurz, C. (2017). Feasibility of MR-only proton dose calculations for prostate cancer radiotherapy using a commercial pseudo-CT generation method., *Phys. Med. Biol.* **62**: 9159–9176.

Mayo, C., Martel, M. K., Marks, L. B., Flickinger, J., Nam, J. & Kirkpatrick, J. (2010a). Radiation dose–volume effects of optic nerves and chiasm, *Int. J. Radiat. Oncol.* **76**(3): S28–S35.

Mayo, C., Yorke, E. & Merchant, T. E. (2010). Radiation associated brainstem injury, *Int. J. Radiat. Oncol.* **76**(3): S36–S41.

McCollough, C. H., Bruesewitz, M. R. & Kofler, J. M. (2006). CT dose reduction and dose management tools: overview of available options., *Radiographics : a review publication of the Radiological Society of North America, Inc* **26**: 503–512.

McGowan, S. E., Albertini, F., Thomas, S. J. & Lomax, A. J. (2015). Defining robustness protocols: a method to include and evaluate robustness in clinical plans, *Phys. Med. Biol.* **60**(7): 2671–2684.

McNamara, A. L., Schuemann, J. & Paganetti, H. (2015). A phenomenological relative biological effectiveness (RBE) model for proton therapy based on all published in vitro cell survival data., *Phys. Med. Biol.* **60**: 8399–8416.

Meyer, S., Gianoli, C., Magallanes, L., Kopp, B., Tessonnier, T., Landry, G., Dedes, G., Voss, B. & Parodi, K. (2017). Comparative Monte Carlo study on the performance of integration- and list-mode detector configurations for carbon ion computed tomography., *Phys. Med. Biol.* **62**: 1096–1112.

Meyer, S., Kamp, F., Tessonnier, T., Mairani, A., Belka, C., Carlson, D. J., Gianoli, C. & Parodi, K. (2019). Dosimetric accuracy and radiobiological implications of ion computed tomography for proton therapy treatment planning., *Phys. Med. Biol.* **64**: 125008.

Meyer, S., Magallanes, L., Kopp, B., Tessonnier, T., Landry, G., Dedes, G., Voss, B., Jaekel, O., Belka, C., Gianoli, C. et al. (2016). Tomographic imaging with carbon ion beams, *2016 IEEE Nuclear Science Symposium, Medical Imaging Conference and Room-Temperature Semiconductor Detector Workshop (NSS/MIC/RTSD)*, IEEE, pp. 1–5.

Miller, K. D., Nogueira, L., Mariotto, A. B., Rowland, J. H., Yabroff, K. R., Alfano, C. M., Jemal, A., Kramer, J. L. & Siegel, R. L. (2019). Cancer treatment and survivorship statistics, 2019, *CA: Cancer J. Clin.* **69**(5): 363–385.

Min, C.-H., Kim, C. H., Youn, M.-Y. & Kim, J.-W. (2006). Prompt gamma measurements for locating the dose falloff region in the proton therapy, *Appl. Phys. Lett.* **89**(18): 183517.

Min, C. H., Lee, H. R. & Kim, C. H. (2011). Two-dimensional prompt gamma measurement simulation for in vivo dose verification in proton therapy: A Monte Carlo study, *Nucl. Technol.* **175**(1): 11–15.

Miralbell, R., Lomax, A., Cella, L. & Schneider, U. (2002). Potential reduction of the incidence of radiation-induced second cancers by using proton beams in the treatment of pediatric tumors., *Int. J. Radiat. Oncol.* **54**: 824–829.

Möhler, C., Wohlfahrt, P., Richter, C. & Greilich, S. (2016). Range prediction for tissue mixtures based on dual-energy CT, *Phys. Med. Biol.* **61**(11): N268.

Molière, G. (1947). Theorie der Streuung schneller geladener Teilchen I - Einzelstreuung am abgeschirmten Coulomb-Feld, *Z. Naturforsch. A* **2**: 133–145.

Molière, G. (1948). Theorie der Streuung schneller geladener Teilchen II - Mehrfach- und Vielfachstreuung, *Z. Naturforsch. A* **3**: 78–97.

Moteabbed, M., España, S. & Paganetti, H. (2011). Monte Carlo patient study on the comparison of prompt gamma and PET imaging for range verification in proton therapy, *Phys. Med. Biol.* **56**(4): 1063–1082.

Mulkens, T. H., Bellinck, P., Baeyaert, M., Ghysen, D., Van Dijck, X., Mussen, E., Venstermans, C. & Termote, J.-L. (2005). Use of an automatic exposure control mechanism for dose optimization in multi-detector row CT examinations: clinical evaluation., *Radiology* **237**: 213–223.

Murphy, M. J., Balter, J., Balter, S., BenComo Jr, J. A., Das, I. J., Jiang, S. B., Ma, C.-M., Olivera, G. H., Rodebaugh, R. F., Ruchala, K. J. et al. (2007). The management of imaging dose during image-guided radiotherapy: report of the AAPM Task Group 75, *Med. Phys.* **34**(10): 4041–4063.

Möhler, C., Russ, T., Wohlfahrt, P., Elter, A., Runz, A., Richter, C. & Greilich, S. (2018). Experimental verification of stopping-power prediction from single- and dual-energy computed tomography in biological tissues., *Phys. Med. Biol.* **63**: 025001.

Nakajima, K., Iwata, H., Ogino, H., Hattori, Y., Hashimoto, S., Hayashi, K., Toshito, T., Akita, K., Iwana, M., Baba, F., Nakamae, K., Shibamoto, Y. & Mizoe, J. (2017). Clinical Outcomes of Image-Guided Proton Therapy for Stage I Non–small Cell Lung Cancer, *Int. J. Radiat. Oncol.* **99**(2): E483–E484.
URL: *https://linkinghub.elsevier.com/retrieve/pii/S0360301617328134*

Nelder, J. A. & Mead, R. (1965). A Simplex Method for Function Minimization, *Comput. J.* **7**(4): 308–313.
URL: *https://academic.oup.com/comjnl/article-lookup/doi/10.1093/comjnl/7.4.308*

Nenoff, L., Matter, M., Hedlund Lindmar, J., Weber, D. C., Lomax, A. J. & Albertini, F. (2019). Daily adaptive proton therapy - the key to innovative planning approaches for paranasal cancer treatments., *Acta Oncol.* **58**: 1423–1428.

Newhauser, W. D. & Durante, M. (2011). Assessing the risk of second malignancies after modern radiotherapy, *Nat. Rev. Cancer* **11**(6): 438–448.

Newhauser, W. D. & Zhang, R. (2015). The physics of proton therapy, *Phys. Med. Biol.* **60**(8): R155–R209.

Niita, K., Chiba, S., Maruyama, T., Maruyama, T., Takada, H., Fukahori, T., Nakahara, Y. & Iwamoto, A. (1995). Analysis of the (n,xN′) reactions by quantum molecular dynamics plus statistical decay model, *Phys. Rev. C* **52**(5): 2620–2635.

Nishio, T., Miyatake, A., Ogino, T., Nakagawa, K., Saijo, N. & Esumi, H. (2010). The development and clinical use of a beam ON-LINE PET system mounted on a rotating gantry port in proton therapy, *Int. J. Radiat. Oncol.* **76**(1): 277–286.

Otto, K. (2007). Volumetric modulated arc therapy: IMRT in a single gantry arc, *Med. Phys.* **35**(1): 310–317.

Paganetti, H. (2009). Dose to water versus dose to medium in proton beam therapy, *Phys. Med. Biol.* **54**(14): 4399–4421.

Paganetti, H. (2012). Range uncertainties in proton therapy and the role of Monte Carlo simulations., *Phys. Med. Biol.* **57**: R99–117.

Paganetti, H. (2014). Relative biological effectiveness (RBE) values for proton beam therapy. Variations as a function of biological endpoint, dose, and linear energy transfer., *Phys. Med. Biol.* **59**: R419–R472.

Paganetti, H. (2016). *Proton Beam Therapy*, IOP Publishing.

Paganetti, H., Niemierko, A., Ancukiewicz, M., Gerweck, L. E., Goitein, M., Loeffler, J. S. & Suit, H. D. (2002). Relative biological effectiveness (RBE) values for proton beam therapy, *Int. J. Radiat. Oncol.* **53**(2): 407–421.

Park, P. C., Zhu, X. R., Lee, A. K., Sahoo, N., Melancon, A. D., Zhang, L. & Dong, L. (2012). A beam-specific planning target volume (PTV) design for proton therapy to account for setup and range uncertainties, *Int. J. Radiat. Oncol.* **82**(2): e329–e336.

Park, Y.-K., Sharp, G. C., Phillips, J. & Winey, B. A. (2015). Proton dose calculation on scatter-corrected cbct image: Feasibility study for adaptive proton therapy, *Med. Phys.* **42**(8): 4449–4459.

Parodi, K., Enghardt, W. & Haberer, T. (2001). In-beam PET measurements of radioactivity induced by proton beams, *Phys. Med. Biol.* **47**(1): 21–36.

Parodi, K., Paganetti, H., Shih, H. A., Michaud, S., Loeffler, J. S., DeLaney, T. F., Liebsch, N. J., Munzenrider, J. E., Fischman, A. J., Knopf, A. & Bortfeld, T. (2007). Patient study of in vivo verification of beam delivery and range, using positron emission tomography and computed tomography imaging after proton therapy, *Int. J. Radiat. Oncol.* **68**(3): 920–934.

Parodi, K. & Polf, J. C. (2018). In vivo range verification in particle therapy, *Med. Phys.* **45**(11).

Paterson, R. & Parker, H. M. (1934). A dosage system for gamma ray therapy, *Brit. J. Radiol.* **7**(82): 592–632.

Pawelke, J., Enghardt, W., Haberer, T., Hasch, B., Hinz, R., Kramer, M., Lauckner, E. & Sobiella, M. (1997). In-beam PET imaging for the control of heavy-ion tumour therapy, *IEEE Trans. Nucl. Sci.* **44**(4): 1492–1498.

Payne, M. G. (1969). Energy straggling of heavy charged particles in thick absorbers, *Phys. Rev.* **185**(2): 611–623.

Pedroni, E., Scheib, S., Böhringer, T., Coray, A., Grossmann, M., Lin, S. & Lomax, A. (2005). Experimental characterization and physical modelling of the dose distribution of scanned proton pencil beams, *Phys. Med. Biol.* **50**(3): 541–561.

Penfold, S. N., Rosenfeld, A. B., Schulte, R. W. & Schubert, K. E. (2009). A more accurate reconstruction system matrix for quantitative proton computed tomography., *Med. Phys.* **36**: 4511–4518.

Penfold, S. N., Schulte, R. W., Censor, Y. & Rosenfeld, A. B. (2010). Total variation superiorization schemes in proton computed tomography image reconstruction., *Med. Phys.* **37**: 5887–5895.

Perris, A. & Zarris, G. (1989). Specific primary ionisation for electrons, protons and alpha particles incident on water, *Phys. Med. Biol.* **34**(8): 1113–1118.

Pettersen, H. E. S., Alme, J., Biegun, A., van den Brink, A., Chaar, M., Fehlker, D., Meric, I., Odland, O. H., Peitzmann, T., Rocco, E., Wang, H., Yang, S., Zhang, C. & Röhrich, D. (2016). Proton tracking in a high-granularity digital tracking calorimeter for proton CT purposes, *Nucl. Instr. Meth. Phys. Res. A 860 51-61* .

Piersimoni, P., Ramos-Méndez, J., Geoghegan, T., Bashkirov, V., Schulte, R. & Faddegon, B. (2017). The effect of beam purity and scanner complexity on proton CT accuracy, *Med. Phys.* **44**(1): 284–298.

Pinto, M., Bajard, M., Brons, S., Chevallier, M., Dauvergne, D., Dedes, G., De Rydt, M., Freud, N., Krimmer, J., La Tessa, C., Létang, J. M., Parodi, K., Pleskač, R., Prieels, D., Ray, C., Rinaldi, I., Roellinghoff, F., Schardt, D., Testa, E. & Testa, M. (2015). Absolute prompt-gamma yield measurements for ion beam therapy monitoring., *Phys. Med. Biol.* **60**: 565–594.

Pinto, M., Dauvergne, D., Freud, N., Krimmer, J., Letang, J. M., Ray, C., Roellinghoff, F. & Testa, E. (2014). Design optimisation of a TOF-based collimated camera prototype for online hadrontherapy monitoring, *Phys. Med. Biol.* **59**(24): 7653–7674.

Pisani, L., Lockman, D., Jaffray, D., Yan, D., Martinez, A. & Wong, J. (2000). Setup error in radiotherapy: on-line correction using electronic kilovoltage and megavoltage radiographs, *Int. J. Radiat. Oncol.* **47**(3): 825–839.

Podgorsak, E. B. (2016). *Radiation Physics for Medical Physicists*, Springer-Verlag GmbH.

Polf, J. C., Panthi, R., Mackin, D. S., McCleskey, M., Saastamoinen, A., Roeder, B. T. & Beddar, S. (2013). Measurement of characteristic prompt gamma rays emitted from oxygen and carbon in tissue-equivalent samples during proton beam irradiation, *Phys. Med. Biol.* **58**(17): 5821–5831.

Polf, J. C., Peterson, S., McCleskey, M., Roeder, B. T., Spiridon, A., Beddar, S. & Trache, L. (2009). Measurement and calculation of characteristic prompt gamma ray spectra emitted during proton irradiation, *Phys. Med. Biol.* **54**(22): N519–N527.

Poludniowski, G., Allinson, N. M. & Evans, P. M. (2014). Proton computed tomography reconstruction using a backprojection-then-filtering approach., *Phys. Med. Biol.* **59**: 7905–7918.

Poston, G. (2007). *Textbook of surgical oncology*, Informa Healthcare, London.

Priegnitz, M., Helmbrecht, S., Janssens, G., Perali, I., Smeets, J., Stappen, F. V., Sterpin, E. & Fiedler, F. (2015). Measurement of prompt gamma profiles in inhomogeneous targets with a knife-edge slit camera during proton irradiation, *Phys. Med. Biol.* **60**(12): 4849–4871.

Quiñones, C. T., Létang, J. M. & Rit, S. (2016). Filtered back-projection reconstruction for attenuation proton CT along most likely paths., *Phys. Med. Biol.* **61**: 3258–3278.

Radon, J. (1917). Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten, *Akad. Wiss.* **69**: 262–277.

Ramachandran, G. N. & Lakshminarayanan, A. V. (1971). Three-dimensional reconstruction from radiographs and electron micrographs: application of convolutions instead of Fourier transforms., *P. Natl. Acad. Sci. Usa.* **68**: 2236–2240.

Rank, C. M., Hünemohr, N., Nagel, A. M., Röthke, M. C., Jäkel, O. & Greilich, S. (2013). MRI-based simulation of treatment plans for ion radiotherapy in the brain region., *Radiother. Oncol.* **109**: 414–418.

Reichhart, L., Akimov, D. Y., Araújo, H., Barnes, E., Belov, V., Burenkov, A., Chepel, V., Currie, A., DeViveiros, L., Edwards, B. et al. (2012). Quenching factor for low-energy nuclear recoils in a plastic scintillator, *Phys. Rev. C* **85**(6): 065801.

Resch, A. F., Landry, G., Kamp, F., Cabal, G., Belka, C., Wilkens, J. J., Parodi, K. & Dedes, G. (2017). Quantification of the uncertainties of a biological model and their impact on variable RBE proton treatment plan optimization., *Phys. Med.* **36**: 91–102.

Richard, M.-H., Chevallier, M., Dauvergne, D., Freud, N., Henriquet, P., Foulher, F. L., Letang, J. M., Montarou, G., Ray, C., Roellinghoff, F., Testa, E., Testa, M. & Walenta, A. H. (2011). Design guidelines for a double scattering compton camera for prompt-γ imaging during ion beam therapy: A Monte Carlo simulation study, *IEEE Trans. Nucl. Sci.* **58**(1): 87–94.

Richter, C., Pausch, G., Barczyk, S., Priegnitz, M., Keitz, I., Thiele, J., Smeets, J., Stappen, F. V., Bombelli, L., Fiorini, C., Hotoiu, L., Perali, I., Prieels, D., Enghardt, W. & Baumann, M. (2016). First clinical application of a prompt gamma based in vivo proton range verification system, *Radiother. Oncol.* **118**(2): 232–237.

Rinaldi, I., Brons, S., Gordon, J., Panse, R., Voss, B., Jäkel, O. & Parodi, K. (2013). Experimental characterization of a prototype detector system for carbon ion radiography and tomography., *Phys. Med. Biol.* **58**: 413–427.

Rinaldi, I., Brons, S., Jäkel, O., Voss, B. & Parodi, K. (2014). Experimental investigations on carbon ion scanning radiography using a range telescope., *Phys. Med. Biol.* **59**: 3041–3057.

Rit, S., Dedes, G., Freud, N., Sarrut, D. & Létang, J. M. (2013). Filtered backprojection proton CT reconstruction along most likely paths., *Med. Phys.* **40**: 031103.

Rädler, M., Landry, G., Rit, S., Schulte, R. W., Parodi, K. & Dedes, G. (2018). Two-dimensional noise reconstruction in proton computed tomography using distance-driven filtered back-projection of simulated projections., *Phys. Med. Biol.* **63**: 215009.

Röntgen, W. C. (1898). Über eine neue Art von Strahlen, *Ann. Phys-berlin.* **300**(1): 12–17.

Sadrozinski, H. F.-W., Geoghegan, T., Harvey, E., Johnson, R. P., Plautz, T. E., Zatserklyaniy, A., Bashkirov, V., Hurley, R. F., Piersimoni, P., Schulte, R. W., Karbasi, P., Schubert, K. E., Schultze, B. & Giacometti, V. (2016). Operation of the preclinical head scanner for proton CT., *Nucl. Instr. Meth. Phys. Res. A* **831**: 394–399.

Saito, M. & Sagara, S. (2017a). A simple formulation for deriving effective atomic numbers via electron density calibration from dual-energy CT data in the human body, *Med. Phys.* **44**(6): 2293–2303.

Saito, M. & Sagara, S. (2017b). Simplified derivation of stopping power ratio in the human body from dual-energy CT data., *Med. Phys.* **44**: 4179–4187.

Sakorafas, G. H. & Safioleas, M. (2009). Breast cancer surgery: an historical narrative. part I. From prehistoric times to Renaissance, *Eur. J. Cancer Care* **18**(6): 530–544.

Sarosiek, C., DeJongh, E. A., Coutrakon, G., DeJongh, D. F., Duffin, K. L., Karonis, N. T., Ordoñez, C. E., Pankuch, M., Rykalin, V., Winans, J. R. & Welsh, J. S. (2021). Analysis of characteristics of images acquired with a prototype clinical proton radiography system, *Med. Phys.* **48**(5): 2271–2278.

Schell, S. & Wilkens, J. J. (2010). Advanced treatment planning methods for efficient radiation therapy with laser accelerated proton and ion beams, *Med. Phys.* **37**(10): 5330–5340.

Schippers, J. M. (2017). Beam-Transport Systems for Particle Therapy, *Proc. CAS-CERN Accel. Sch. Accel. Med. Appl.* **1**(CERN-2017-004-SP): 241–252.

Schmid, S., Landry, G., Thieke, C., Verhaegen, F., Ganswindt, U., Belka, C., Parodi, K. & Dedes, G. (2015). Monte Carlo study on the sensitivity of prompt gamma imaging to proton range variations due to interfractional changes in prostate cancer patients., *Phys. Med. Biol.* **60**: 9329–9347.

Schneider, U., Pedroni, E. & Lomax, A. (1996). The calibration of CT Hounsfield units for radiotherapy treatment planning., *Phys. Med. Biol.* **41**: 111–124.

Schneider, W., Bortfeld, T. & Schlegel, W. (2000). Correlation between CT numbers and tissue parameters needed for Monte Carlo simulations of clinical dose distributions., *Phys. Med. Biol.* **45**: 459–478.

Schreiber, J., Bolton, P. R. & Parodi, K. (2016). Invited review article: "hands-on" laser-driven ion acceleration: A primer for laser-driven source development and potential applications, *Rev. Sci. Instrum.* **87**(7): 071101.

Schulte, R. W., Bashkirov, V., Klock, M. C. L., Li, T., Wroe, A. J., Evseev, I., Williams, D. C. & Satogata, T. (2005). Density resolution of proton computed tomography., *Med. Phys.* **32**: 1035–1046.

Schulte, R. W., Penfold, S. N., Tafas, J. T. & Schubert, K. E. (2008). A maximum likelihood proton path formalism for application in proton computed tomography., *Med. Phys.* **35**: 4849–4856.

Schultze, B., Censor, Y., Karbasi, P., Schubert, K. E. & Schulte, R. W. (2019). An improved method of total variation superiorization applied to reconstruction in proton computed tomography, *IEEE Trans. Med. Imaging* **39**(2): 294–307.

Seltzer, S. (1993). Stopping-powers and range tables for electrons, protons, and helium ions, NIST Standard Reference Database 124.

Shäfer, D., van de Haar, P. & Grass, M. (2015). Fully 3D variance reconstruction for circular cone beam CT, *The 13th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, Fully3D, pp. 174–177.

Shinoda, H., Kanai, T. & Kohno, T. (2006). Application of heavy-ion CT., *Phys. Med. Biol.* **51**: 4073–4081.

Shunhavanich, P., Bennett, N. R., Hsieh, S. S. & Pelc, N. J. (2018). Implementation of a piecewise-linear dynamic attenuator, *Medical Imaging 2018: Physics of Medical Imaging*, Vol. 10573, International Society for Optics and Photonics, p. 105730T.

Shunhavanich, P., Bennett, N. R., Hsieh, S. S. & Pelc, N. J. (2019). Implementation of a piecewise-linear dynamic attenuator, *J. Med. Imaging* **6**(2): 023502.

Sigmund, P., Schinner, A. & Paul, H. (2009). Errata and addenda for ICRU report 73, stopping of ions heavier than helium, *J ICRU* **5**(1): 1–10.

Sihver, L. & Mancusi, D. (2009). Present status and validation of HIBRAC, *Radiat. Meas.* **44**(1): 38–46.

Skyrme, T. (1958). The effective nuclear potential, *Nucl. Phys.* **9**(4): 615–634.

Slater, J. D., Slater, J. M. & Wahlen, S. (1992). The potential for proton beam therapy in locally advanced carcinoma of the cervix, *Int. J. Radiat. Oncol.* **22**(2): 343–347.

Slater, J. M. (2007). Developing and understanding a hospital-based proton facility: Bringing physics into medicine, *Technol. Cancer Res. Treat.* **6**(4_suppl): 1–7.

Smeets, J., Roellinghoff, F., Prieels, D., Stichelbaut, F., Benilov, A., Busca, P., Fiorini, C., Peloso, R., Basilavecchia, M., Frizzi, T., Dehaes, J. C. & Dubus, A. (2012). Prompt gamma imaging with a slit camera for real-time range control in proton therapy, *Phys. Med. Biol.* **57**(11): 3371–3405.

Snavely, R. A., Key, M. H., Hatchett, S. P., Cowan, T. E., Roth, M., Phillips, T. W., Stoyer, M. A., Henry, E. A., Sangster, T. C., Singh, M. S., Wilks, S. C., MacKinnon, A., Offenberger, A., Pennington, D. M., Yasuike, K., Langdon, A. B., Lasinski, B. F., Johnson, J., Perry, M. D. & Campbell, E. M. (2000). Intense high-energy proton beams from petawatt-laser irradiation of solids, *Phys. Rev. Lett.* **85**(14): 2945–2948.

Stavrev, P., Stavreva, N., Ruggieri, R. & Nahum, A. (2015). On differences in radiosensitivity estimation: TCP experiments versus survival curves. a theoretical study, *Phys. Med. Biol.* **60**(15): N293–N299.

Stayman, J. W., Mathews, A., Zbijewski, W., Gang, G., Siewerdsen, J., Kawamoto, S., Blevis, I. & Levinson, R. (2016). Fluence-field modulated x-ray CT using multiple aperture devices., *Proceedings of SPIE–the International Society for Optical Engineering* **9783**.

Stichelbaut F, J. Y. (2003). Verification of the proton beam position in the patient by the prompt gamma rays emission., *39th PTCOG Meeting (San Francisco, CA, 2003)*.

Strickland, D. & Mourou, G. (1985). Compression of amplified chirped optical pulses, *Opt. Commun.* **56**(3): 219–221.

Suit, H. D., Goitein, M., Tepper, J., Koehler, A. M., Schmidt, R. A. & Schneider, R. (1975). Exploratory study of proton radiation therapy using large field techniques and fractionated dose schedules, *Ann. Ny. Acad. Sci.* **35**(6): 1646–1657.

Suit, H., DeLaney, T., Goldberg, S., Paganetti, H., Clasie, B., Gerweck, L., Niemierko, A., Hall, E., Flanz, J., Hallman, J. & Trofimov, A. (2010). Proton vs carbon ion beams in the definitive radiation treatment of cancer patients., *Radiother. Oncol.* **95**: 3–22.

Szczykutowicz, T. P., Hermus, J., Geurts, M. & Smilowitz, J. (2015). Realization of fluence field modulated CT on a clinical TomoTherapy megavoltage CT system., *Phys. Med. Biol.* **60**: 7245–7257.

Szczykutowicz, T. P. & Mistretta, C. A. (2013a). Design of a digital beam attenuation system for computed tomography: part I. System design and simulation framework., *Med. Phys.* **40**: 021905.

Szczykutowicz, T. P. & Mistretta, C. A. (2013b). Design of a digital beam attenuation system for computed tomography. part II. Performance study and initial results., *Med. Phys.* **40**: 021906.

Szczykutowicz, T. P. & Mistretta, C. A. (2014). Experimental realization of fluence field modulated CT using digital beam attenuation., *Phys. Med. Biol.* **59**: 1305–1326.

Sørensen, B. S., Overgaard, J. & Bassler, N. (2011). In vitro RBE-LET dependence for multiple particle types., *Acta Oncol.* **50**: 757–762.

Taasti, V. T., Bäumer, C., Dahlgren, C. V., Deisher, A. J., Ellerbrock, M., Free, J., Gora, J., Kozera, A., Lomax, A. J., De Marzi, L. et al. (2018). Inter-centre variability of CT-based stopping-power prediction in particle therapy: survey-based evaluation, *Phys. Imag. Radiat. Oncol.* **6**: 25–30.

Taasti, V. T., Michalak, G. J., Hansen, D. C., Deisher, A. J., Kruse, J. J., Krauss, B., Muren, L. P., Petersen, J. B. B. & McCollough, C. H. (2017). Validation of proton stopping power ratio estimation based on dual energy CT using fresh tissue samples., *Phys. Med. Biol.* **63**: 015012.

Taasti, V. T., Petersen, J. B. B., Muren, L. P., Thygesen, J. & Hansen, D. C. (2016). A robust empirical parametrization of proton stopping power using dual energy CT., *Med. Phys.* **43**: 5547.

Tajima, T. & Dawson, J. M. (1979). Laser electron accelerator, *Phys. Rev. Lett.* **43**(4): 267–270.

Takada, Y., Kondo, K., Marume, T., Nagayoshi, K., Okada, I. & Takikawa, K. (1988). Proton computed tomography with a 250 MeV pulsed beam, *Nucl. Instr. Meth. Phys. Res. A* **273**(1): 410–422.

Tanaka, S., Nishio, T., Matsushita, K., Tsuneda, M., Kabuki, S. & Uesaka, M. (2016). Development of proton CT imaging system using plastic scintillator and CCD camera., *Phys. Med. Biol.* **61**: 4156–4167.

Tanaka, S., Nishio, T., Tsuneda, M., Matsushita, K., Kabuki, S. & Uesaka, M. (2018). Improved proton CT imaging using a bismuth germanium oxide scintillator., *Phys. Med. Biol.* **63**: 035030.

Taylor, J., Poludniowski, G., Price, T., Waltham, C., Allport, P., Casse, G., Esposito, M., Evans, P., Green, S., Manger, S. et al. (2016). An experimental demonstration of a new type of proton computed tomography using a novel silicon tracking detector, *Med. Phys.* **43**(11): 6129–6136.

Telsemeyer, J., Jäkel, O. & Martišíková, M. (2012). Quantitative carbon ion beam radiography and tomography with a flat-panel detector., *Phys. Med. Biol.* **57**: 7957–7971.

Testa, E., Bajard, M., Chevallier, M., Dauvergne, D., Foulher, F. L., Freud, N., Létang, J.-M., Poizat, J.-C., Ray, C. & Testa, M. (2008). Monitoring the Bragg peak location of 73MeV/u carbon ions by means of prompt γ-ray measurements, *Appl. Phys. Lett.* **93**(9): 093506.

Testa, E., Bajard, M., Chevallier, M., Dauvergne, D., Foulher, F. L., Freud, N., Létang, J., Poizat, J., Ray, C. & Testa, M. (2009). Dose profile monitoring with carbon ions by means of prompt-gamma measurements, *Nucl. Instrum. Methods Phys. Res. Sect. B Beam Interact. Mater. At.* **267**(6): 993–996.

Testa, M., Bajard, M., Chevallier, M., Dauvergne, D., Freud, N., Henriquet, P., Karkar, S., Foulher, F. L., Létang, J. M., Plescak, R., Ray, C., Richard, M.-H., Schardt, D. & Testa, E. (2010). Real-time monitoring of the Bragg-peak position in ion therapy by means of single photon detection, *Radiat. Environ. Bioph.* **49**(3): 337–343.

Thing, R. S., Bernchou, U., Mainegra-Hing, E., Hansen, O. & Brink, C. (2016). Hounsfield unit recovery in clinical cone beam CT images of the thorax acquired for image guided radiation therapy, *Phys. Med. Biol.* **61**(15): 5781–5802.

Tian, L., Huang, Z., Janssens, G., Landry, G., Dedes, G., Kamp, F., Belka, C., Pinto, M. & Parodi, K. (2021). Accounting for prompt gamma emission and detection for range verification in proton therapy treatment planning, *Phys. Med. Biol.* **66**(5): 055005.

Tian, L., Landry, G., Dedes, G., Kamp, F., Pinto, M., Niepel, K., Belka, C. & Parodi, K. (2018). Toward a new treatment planning approach accounting for in vivo proton range verification., *Phys. Med. Biol.* **63**: 215025.

Tian, L., Landry, G., Dedes, G., Pinto, M., Kamp, F., Belka, C. & Parodi, K. (2020). A new treatment planning approach accounting for prompt gamma range verification and interfractional anatomical changes, *Phys. Med. Biol.* **65**(9): 095005.

Tian, Z., Jia, X., Yuan, K., Pan, T. & Jiang, S. B. (2011). Low-dose CT reconstruction via edge-preserving total variation regularization, *Phys. Med. Biol.* **56**(18): 5949.

Tobias, C. A., Roberts, J., Lawrence, J., Low-Beer, B., Anger, H., Born, J., McCombs, R. & Huggins, C. (1955). Irradiation hypophysectomy and related studies using 340-MeV protons and 190-MeV deuterons.

Tommasino, F. & Durante, M. (2015). Proton radiobiology., *Cancers* **7**: 353–381.

Tretyak, V. (2010). Semi-empirical calculation of quenching factors for ions in scintillators, *Astropart. Phys.* **33**(1): 40–53.

Triolo, V. A. (1965). Nineteeth century foundations of cancer research advances in tumor pathology, nomenclature and theories of oncogenesis, *Cancer Res.* **25**: 75–106.

Tschalär, C. (1968a). Straggling distributions of extremely large energy losses, *Nucl. Instrum. Methods* **64**(3): 237–243.

Tschalär, C. (1968b). Straggling distributions of large energy losses, *Nucl. Instrum. Methods* **61**(2): 141–156.

van de Water, S., Safai, S., Schippers, J. M., Weber, D. C. & Lomax, A. J. (2019). Towards FLASH proton therapy: the impact of treatment planning and machine characteristics on achievable dose rates, *Acta Oncol.* **58**(10): 1463–1469.

Vavilov, P. V. (1957). Ionization losses of high-energy heavy particles, *Soviet Phys. JETP* **5**.
URL: *https://www.osti.gov/biblio/4311507*

Veiga, C., Alshaikhi, J., Amos, R., Lourenço, A. M., Modat, M., Ourselin, S., Royle, G. & McClelland, J. R. (2015). Cone-beam computed tomography and deformable registration-based "dose of the day" calculations for adaptive proton therapy, *Int. J. Part. Ther.* **2**(2): 404–414.

Veiga, C., Janssens, G., Teng, C.-L., Baudier, T., Hotoiu, L., McClelland, J. R., Royle, G., Lin, L., Yin, L., Metz, J. et al. (2016). First clinical investigation of cone beam computed tomography and deformable registration for adaptive proton therapy for lung cancer, *Int. J. Radiat. Oncol.* **95**(1): 549–559.

Verburg, J. M. & Seco, J. (2014). Proton range verification through prompt gamma-ray spectroscopy, *Phys. Med. Biol.* **59**(23): 7089–7106.

Verburg, J. M., Shih, H. A. & Seco, J. (2012). Simulation of prompt gamma-ray emission during proton radiotherapy, *Phys. Med. Biol.* **57**(17): 5459–5472.

Verellen, D., Ridder, M. D., Linthout, N., Tournel, K., Soete, G. & Storme, G. (2007). Innovations in image-guided radiotherapy, *Nat. Rev. Cancer* **7**(12): 949–960.

Verma, V., Lin, S. H., Simone, C. B. et al. (2016). Clinical outcomes and toxicities of proton radiotherapy for gastrointestinal neoplasms: a systematic review, *J. Gastrointest. Oncol.* **7**(4): 644.

Volz, L., Collins-Fekete, C.-A., Piersimoni, P., Johnson, R. P., Bashkirov, V., Schulte, R. & Seco, J. (2017). Stopping power accuracy and achievable spatial resolution of helium ion imaging using a prototype particle ct detector system, *Curr. Dir. Biomed. Eng.* **3**(2): 401–404.

Volz, L., Piersimoni, P., Bashkirov, V. A., Brons, S., Collins-Fekete, C.-A., Johnson, R. P., Schulte, R. W. & Seco, J. (2018). The impact of secondary fragments on the image quality of helium ion imaging., *Phys. Med. Biol.* **63**: 195016.

Waldman, A. D., Fritz, J. M. & Lenardo, M. J. (2020). A guide to cancer immunotherapy: from T cell basic science to clinical practice, *Nat. Rev. Immunol.* **20**(11): 651–668.

Wang, W., Gang, G. J., Siewerdsen, J. H., Levinson, R., Kawamoto, S. & Stayman, J. W. (2019). Volume-of-interest imaging with dynamic fluence modulation using multiple aperture devices., *J. Med. Imaging* **6**: 033504.

Weber, D. C., Schneider, R., Goitein, G., Koch, T., Ares, C., Geismar, J. H., Schertler, A., Bolsi, A. & Hug, E. B. (2012). Spot scanning-based proton therapy for intracranial meningioma: long-term results from the Paul Scherrer Institute, *Int. J. Radiat. Oncol.* **83**(3): 865–871.

Wedenberg, M., Lind, B. K. & Hårdemark, B. (2013). A model for the relative biological effectiveness of protons: the tissue specific parameter α/β of photons is a predictor for the sensitivity to LET changes., *Acta Oncol.* **52**: 580–588.

Weisskopf, V. F. & Ewing, D. H. (1940). On the yield of nuclear reactions with heavy elements, *Phys. Rev.* **57**(6): 472–485.

Wenz, J., Döpp, A., Khrennikov, K., Schindler, S., Gilljohann, M. F., Ding, H., Götzfried, J., Buck, A., Xu, J., Heigoldt, M., Helml, W., Veisz, L. & Karsch, S. (2019). Dual-energy electron beams from a compact laser-driven accelerator, *Nat. Photonics* **13**(4): 263–269.

Wenz, J., Schleede, S., Khrennikov, K., Bech, M., Thibault, P., Heigoldt, M., Pfeiffer, F. & Karsch, S. (2015). Quantitative x-ray phase-contrast microtomography from a compact laser-driven betatron source, *Nat. Commun.* **6**(1).

WHO (2020). https://www.who.int/health-topics/cancertab=tab_1.

Wilkens, J. J. (2004). *Evaluation of Radiobiological Effects in Intensity Modulated Proton Therapy : New Strategies for Inverse Treatment Planning*, PhD thesis.

Wilkens, J. J. & Oelfke, U. (2003). Analytical linear energy transfer calculations for proton therapy, *Med. Phys.* **30**(5): 806–815.

Wilkens, J. J. & Oelfke, U. (2004). A phenomenological model for the relative biological effectiveness in therapeutic proton beams., *Phys. Med. Biol.* **49**: 2811–2825.

Wilkens, J. J. & Oelfke, U. (2005). Optimization of radiobiological effects in intensity modulated proton therapy., *Med. Phys.* **32**: 455–465.

Willemink, M. J. & Noël, P. B. (2018). The evolution of image reconstruction for CT—from filtered back projection to artificial intelligence, *Eur. Radiol.* **29**(5): 2185–2195.

Wilson, R. R. (1946). Radiological use of fast protons, *Radiology* **47**(5): 487–491.

Wohlfahrt, P., Möhler, C., Hietschold, V., Menkel, S., Greilich, S., Krause, M., Baumann, M., Enghardt, W. & Richter, C. (2017a). Clinical implementation of dual-energy CT for proton treatment planning on pseudo-monoenergetic CT scans, *Int. J. Radiat. Oncol.* **97**(2): 427–434.

Wohlfahrt, P., Möhler, C., Stützer, K., Greilich, S. & Richter, C. (2017b). Dual-energy CT based proton range prediction in head and pelvic tumor patients, *Radiother. Oncol.* **125**(3): 526–533.

Wouters, B. G., Lam, G. K., Oelfke, U., Gardey, K., Durand, R. E. & Skarsgard, L. D. (1996). Measurements of relative biological effectiveness of the 70 MeV proton beam at TRIUMF using Chinese hamster V79 cells and the high-precision cell sorter assay., *Radiat. Res.* **146**: 159–170.

Wunderlich, A. & Noo, F. (2008). Image covariance and lesion detectability in direct fan-beam x-ray computed tomography., *Phys. Med. Biol.* **53**: 2471–2493.

Xiao, Y., Reiff, J. E., Holmes, T., Holmes, T., Vargas, H. A., Akin, O., Hricak, H., Speer, T. W., Wazer, D. E., Rusu, I., Budach, V., Varnum, S. M., Sowa, M. B., Morgan, W. F., DeNittis, A. S., Price, R. A., Yeung, D., Palta, J., Gracely, E. J., Fisher, B. J., Daugherty, L. C., Reiff, J. E., Yu, Y., Doyle, L., Limbergen, E., Guerrieri, P., Montemaggi, P., Guerrieri, P., Montemaggi, P., Knowlton, C. A., Mackay, M. K., Limbergen, E., Guerrieri, P., Montemaggi, P., Guerrieri, P., Montemaggi, P., Brashears, J. H., Calvo, F. A., Calvo, F. A., Calvo, F. A., Calvo, F. A., González-Domingo, M., Usychkin, S., Dragun, A. E., Knowlton, C. A., Mackay, M. K., Guerrieri, P., Montemaggi, P., Rübe, C. E., Holmes, T., Kong, F.-M., Wang, J., Yu, Y., Doyle, L., Troicki, F. T., Poli, J. & Speer, T. W. (2013). Inverse treatment planning, *Encyclopedia of Radiation Oncology*, Springer Berlin Heidelberg, pp. 397–398.

Xie, Y., Ainsley, C., Yin, L., Zou, W., McDonough, J., Solberg, T. D., Lin, A. & Teo, B.-K. K. (2018). Ex vivo validation of a stoichiometric dual energy CT proton stopping power ratio calibration., *Phys. Med. Biol.* **63**: 055016.

Yang, M., Virshup, G., Clayton, J., Zhu, X., Mohan, R. & Dong, L. (2010). Theoretical variance analysis of single-and dual-energy computed tomography methods for calculating proton stopping power ratios of biological tissues, *Phys. Med. Biol.* **55**(5): 1343.

Yang, M., Zhu, X. R., Park, P. C., Titt, U., Mohan, R., Virshup, G., Clayton, J. E. & Dong, L. (2012). Comprehensive analysis of proton range uncertainties related to patient stopping-power-ratio estimation using the stoichiometric calibration., *Phys. Med. Biol.* **57**: 4095–4115.

Yu, C. X. (1995). Intensity-modulated arc therapy with dynamic multileaf collimation: an alternative to tomotherapy, *Phys. Med. Biol.* **40**(9): 1435–1449.

ZfKD (2016). Cancer in germany, *Technical report*.
URL: *https://www.krebsdaten.de/Krebs/EN/Content/Publications/Cancer_in_Germany/cancer_chapters_2015_2016/cancer_c00-c97_all_sites.pdf?__blob=publicationFile*

Zhang, Y. & Ning, R. (2008). Investigation of image noise in cone-beam CT imaging due to photon counting statistics with the Feldkamp algorithm by computer simulations, *J. X-ray Sci. Technol.* **16**: 143–158.

Zhu, X., España, S., Daartz, J., Liebsch, N., Ouyang, J., Paganetti, H., Bortfeld, T. R. & Fakhri, G. E. (2011). Monitoring proton radiation therapy with in-room PET imaging, *Phys. Med. Biol.* **56**(13): 4041–4057.

Acknowledgements

This manuscript may have a sole author, but in reality it is the outcome of the aspirations of a large number of people. I feel thankful and lucky that my scientific career has been so far quite the opposite of a lonely trip.

I am grateful to Dr. Denis Dauvergne and Dr. Étienne Testa for giving me the chance to make the transition to the field of Medical Physics and to work for almost three years in the outstandingly supportive and friendly environment of IPNL Lyon. A special mention goes to Dr. Simon Rit. It was just a hinge of fate that brought me to work alongside him in the topic of proton imaging, which turned out to become a long lasting cooperation and my main research focus during the past ten years.

I would like to express my gratitude to Prof. Katia Parodi for bringing me to the, back then newly founded, LMU Department of Medical Physics in Munich. Working next to a renown scientist like her, in a stimulating environment full of challenges and opportunities, transformed my scientific career beyond what I had ever imagined. It has been eight exciting years, during which I have learned a great deal about Medical Physics and I have matured and evolved as a scientist.

I am indebted to Prof. Reinhard Schulte, Prof. Robert Johnson and Dr. Mark Pankuch. It has been a great pleasure working with them. I cherish our collaboration and I always enjoy visiting them in Loma Linda, Santa Cruz or Chicago.

Throughout the past eleven years I had the privilege to work together with a large number of bright and diligent students. Many thanks to Jonas Hoppmann, Yuki Asano, Susanne Schmidt, Alexander Wiselsperger, Andreas Resch, Ludovica de Angelis, Martin Rädler, Philipp Wesp, Erika Kollitz, Liheng Tian, Hubertus Drosten, Aaron Schulz, Stefanie Götz and Franz Englbrecht. Last but not least, I want to thank Jannis Dickmann, who with his ingenuity and his dedication to his PhD project played a pivotal role in advancing proton imaging research activities of our international collaboration.

I could not possibly overstate the contribution of Prof. Guillaume Landry to the work presented in this manuscript. His scientific rigor, enthusiasm and unparalleled motivation have been decisive factors in the success of many of the projects presented here. Moreover, I can attest that sharing the workload of my scientific endeavors with a true friend made it all the more fun.

I would like to thank once again my parents. They are probably unaware of the full extent of it, but their incessant – for more than four decades – support and belief in my capabilities have had a far reaching impact on my actions and decisions.

Lena, my dear wife and companion in life, has been next to me in all my important moments. It is hard to even imagine someone else sacrificing so much so that I could follow my path. But above all, I thank her for creating with me our own family.

This manuscript is dedicated to my two sons, Prokopios (Προκόπης) and Dimitrios (Δημήτρης). If they ever read this manuscript I want them to remember that my professional life, although often demanding and time consuming, was eclipsed by the goal of being a good father. I wholeheartedly wish to both of them that in their lives they outshine any of my successes and achievements.