UNIVERSITY OF CALIFORNIA
SANTA CRUZ

AN EVALUATION OF A PROTOTYPE PROTON CT SCANNER

A dissertation submitted in partial satisfaction of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

in

PHYSICS

by

Tia Elizabeth Plautz

June 2016

The Dissertation of Tia Elizabeth Plautz
is approved:

Professor Robert P. Johnson, Chair

Professor Emeritus Hartmut F.-W. Sadrozinski

Professor Jason Nielsen

Tyrus Miller
Vice Provost and Dean of Graduate Studies

Contents

List of Figures

List of Tables

Abstract

An Evaluation of a Prototype Proton CT Scanner

by

Tia Elizabeth Plautz

Since the 1990s, the number of clinical proton therapy facilities around the world has been growing exponentially. Because of this, and the lack of imaging support for proton therapy in the treatment room, a renewed interest in proton radiography and computed tomography (CT) has emerged. This imaging modality was largely abandoned in the 1970s and '80s in favor of the already successful x-ray CT, for reasons including long acquisition times and inadequate spatial resolution. Protons are particularly useful for radiotherapy because of their finite range in matter and their favorable energy profile which facilitates greater conformality than other radiotherapies; however, in order to realize the full potential of proton radiotherapy, the range of protons in the patient must be precisely known.

Presently, proton therapy treatment planning is accomplished by taking x-ray CTs of the patient and converting each voxel into proton relative stopping power with respect to water (RSP) via a stoichiometrically-acquired calibration curve. However, since there is no unique relationship between Hounsfield values and RSP, this procedure has inherent uncertainties of a few percent in the proton range, requiring additional distal uncertainty margins in proton treatment plans. In contrast to x-ray CT, proton CT measures the RSP of an object directly, eliminating the need for Hounsfield-value-to-RSP conversion.

In the prototype proton CT scanner that we have developed, a low-intensity beam of 200 MeV protons traverses a patient entirely and stops in a downstream energy/range detector. The entry and exit vectors of each proton are measured in

order to determine a most-likely path of the proton through the object, and the response of the energy/range detector is converted to the water-equivalent path length of each proton in the object. These measurements are made at many angles between 0° and 360° in order to reconstruct a three-dimensional map of proton RSP in the object.

In this thesis, I present the prototype proton CT scanner that our collaboration has recently developed. I discuss the performance of the proton tracking system and data-acquisition system, as well as the performance of the energy/range detector in terms of speed, efficiency, and accuracy. In addition, I describe a digital model of this system for Monte Carlo simulation and the process by which it was validated. This discussion is followed by a proof-of-concept study on proton radiography in which two distinct methods for radiography are compared. Finally, I describe the process by which CT images are reconstructed, and evaluate the quality of the images reconstructed from data collected using the prototype scanner in terms of the fidelity of the reconstructed RSP values, and the spatial resolution.

Acknowledgments

I would first like to say thank you to my advisors, Prof. Robert Johnson and Prof. Hartmut Sadrozinski for your support in the pursuit of my degree. Thank you for your guidance over these past five years, and for providing me the environment to discover my interests. To the pCT collaboration, especially Dr. Reinhard Schulte, Dr. Andriy Zatserklyaniy, Valentina Giacometti, Prof. Vladimir Bashkirov, and Ford Hurley. Thank you for your guidance, mentorship and friendship, alike.

There have been many people along the way who also deserve recognition. To my Mom and Dad; thank you for raising me to be headstrong and independent, for nurturing my interest in science and math from a young age, and for teaching me that girls can do anything they set their minds to. To my high school science and math teachers, especially Dr. Linda Burke, Eric Tucker and Tom Rizzotti; thank you not only for cultivating my knowledge of math and science but for supporting me and providing a refuge for me at a very difficult time in my life. To Adrian Liu; thank you for being a continued friend and mentor – I honestly am not sure I could have accomplished this without you. To my undergraduate advisor, Prof. Yuval Grossman; thank you for your advice in my undergraduate years – I may not have applied to graduate school if not for your support. And to my undergraduate research advisor Prof. Don Bilderback; thank you for believing in me.

Finally, I thank my best friend and partner, Matt Wittmann, for his unconditional love. You are my rock and my star and you challenge me to be better every day. I am so lucky that I get to adventure through life and the world with you.

This dissertation includes reprints of text and figures from the following materials previously published or submitted for publication:

- Tia E. Plautz, et al. [9 authors]. "Evaluation of spatial resolution for a prototype proton CT scanner." Submitted. DATE.

- Tia Plautz, et al. [16 authors]. "200 MeV proton radiography studies with a hand phantom using a prototype proton CT scanner." In: *IEEE Transactions on Medical Imaging*. 33.4 (2014): 875-881.

- Robert P. Johnson, V. A. Bashkirov, ... T. E. Plautz, et al. [12 authors]. "A fast experimental scanner for proton CT: Technical performance and first experience with phantom scans." In: *IEEE Transactions on Nuclear Science*. 63.1 (2015): 5260.

- V. A. Bashkirov, R. W. Schulte, ... T. Plautz, et al. [8 authors]. "Novel scintillation detector design and performance for proton radiography and computed tomography." In: *Medical Physics*. 43.2 (2016): 664674.

Chapter 1

Introduction

1.1 Motivation for Proton Computed Tomography

The primary motivation for proton imaging arises from the usefulness of proton radiotherapy for treatment of various cancers. As protons travel through matter, they lose energy due to electrical interactions with other particles. At high energies, protons radiate very little, but as they slow down they radiate more. In the last 2% of their range, as they slow to a stop, they release all of their residual energy in a very localized area – a phenonmenon called the Bragg peak. By varying the beam energy, radiation oncologists can pinpoint treatment areas to within a few millimeters when the composition of the object is known.

Compared with x-rays, protons are capable of doing less damage to normal tissues, due to the low levels of radiation released proximal to the Bragg peak, and the steep falloff at the distal edge of the Bragg peak. As illustrated in fig. 1.1, MeV x-rays give off the largest amount of radiation near the proximal edge of their range in the patient, and gradually fall off toward the distal edge. In contrast,

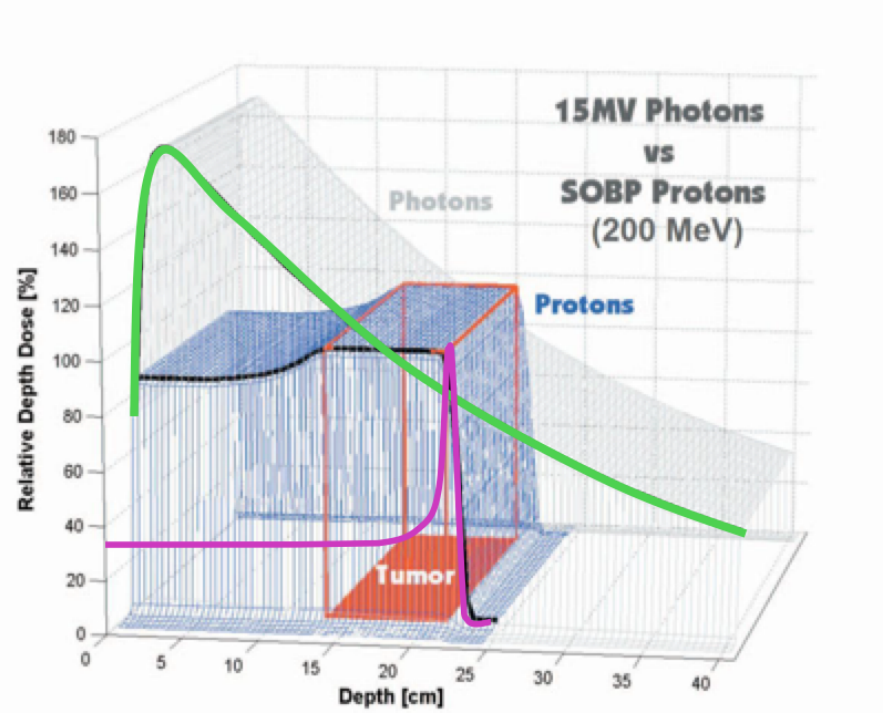


Figure 1.1: Comparison of relative radiation dose for 15 MeV photons (green curve) and 200 MeV protons (magenta curve). The width of the proton Bragg peak is exaggerated for clarity. The plot is normalized to the dose that is desired at the site of the tumor (orange box). The spread-out Bragg peak, produced by a modulated proton beam is colored black.

protons give off only a small fraction of their maximum dose at the proximal edge, and deliver no radiation past the distal edge of the Bragg peak. In addition, by modulating the energy of the proton beam, highly conformal dose distributions within the patient can be achieved, since lowering the energy of the beam reduces the range of protons in the material, and therefore, shifts the Bragg peak closer to the entry point. Likewise, increasing the energy increases the range of protons in the material.

Today, protons are generally accepted as the gold standard of radiotherapy treatment for certain types of cancers such as head and neck tumors, especially those at the base of the skull, and in pediatric cases, since these patients tend to

be more radiosensitive. In addition, proton radiotherapy is also becoming a more popular treatment for prostate and lung cancers.

Indeed, protons could be well-suited for any number of other localized cancers as well, however, in order to fully realize the potential of the fundamental properties of proton beams (and other, heavier ion beams) the distribution of proton stopping power within the patient must be well known prior to treatment. Presently, this information is obtained from pre-treatment x-ray CT scans of the patient. X-ray CT scans are reconstructed in terms of Hounsfield units which describe the relative attenuation of photons after they have passed through the patient. These reconstructions are then converted into maps of proton stopping power via empirically derived calibration curves [1]. However, the CT number and proton stopping power describe two different physical properties of the material, and although they are both dominated by the electron density of the material, the different dependencies of protons and photons on Z, and the Z/A ratio result in a non-unique relationship between Hounsfield values and proton stopping power. That is, materials with the same Hounsfield value may have different stopping powers, and vice versa. This feature is especially prominent near RSP=1 since most human tissue is composed primarily of water. This means that errors in proton range are consistently as high as 2% of the nominal proton range, or even higher in regions containing tissue, bone and air interfaces. In a region containing air cavities, such as the lungs or rectum, this could be the difference between sparing or irradiating an entire organ.

A 2012 study by Yang et al. [2] analyzed five categories of errors that contribute to uncertainties in the measurements of stopping power. These categories included 1) uncertainties in patient CT imaging, 2) uncertainties in the parameterized stoichiometric formula to calculate theoretical CT numbers, 3) Uncertainties

due to deviation of actual human body tissue from ICRU standard tissue, 4) Uncertainties in mean excitation energies, 5) Uncertainty due to energy dependence of [stopping power ratio] not accounted for by the dose algorithm. The study found that the uncertainties in lung, soft tissue and bone to be 5.0%, 1.6% and 2.4% respectively. In different tissue types the contribution of each category of error varied considerably, however in soft-tissue and bone these uncertainties were largely dominated by the variation in human tissue composition, as well as by errors in the measured Hounsfield values during CT scanning, neither of which would contribute to a calibration in proton CT[1]. A survey at the 2012 annual meeting of American Association of Physicists in Medicine showed that 33% of attendees polled said that range uncertainties are the main obstacle to making proton therapy mainstream [3].

Thus, the goal of proton imaging is to eliminate the conversion step entirely, and instead to measure *directly* the three-dimensional map of proton stopping power within the patient by converting proton residual energy into water-equivalent path length. Simulations and first experimental results have shown that by using a proton CT imaging system, one may be able to reduce this range uncertainty to about 1% or less without increasing the dose to the patient [4].

1.2 Historical Note

The use of energetic protons for therapeutic purposes was first proposed in a 1946 paper by Robert Wilson of the Harvard Cyclotron Laboratory (HCL) [5], best known for his work on the Manhattan project during World War II. Wilson proposed that proton beams would be advantageous for treating internal tumors

[1]In fact, none of categories 2-4 contribute in a proton CT calibration.

because of the favorable relationship between deposited dose and depth in the absorbing material. The first treatment of patients with protons was carried out at particle accelerator laboratories, notably Lawrence Berkeley Laboratory (LBL) [6] in 1955 and HCL soon after in 1961. The group at LBL initially used a high-energy "shoot through" technique for treatment before the advantage of stopping the proton within the tumor was realized by the Boston group.

The first high-energy hospital-based proton treatment facility was installed at Loma Linda University Medical Center (LLUMC) in Loma Linda, CA in 1990 [7][2]. Since then, a rapidly increasing number of hospitals in the U.S. and around the world have followed suit. As of January 2016 there are 18 proton centers in North America, 21 in Europe, the U.K. and Russia, 17 in Asia and 1 in South Africa [10]. In addition there are many more in the planning and construction stages.

The first mention of charged particle CT as an alternative to x-ray CT was by Allan Cormack in his 1963 paper on representing a function by its line integrals [11], however the problem was first tackled in 1968 by Koehler, who published the first proton radiograph in the journal *Science* [12]. Koehler imaged a stack of parallel-sided aluminum absorbers at the HCL beamline using proton fluence as the mechanism for contrast. Although other works on proton radiography followed, this factor limited the application to thin samples [13–15].

In 1972 Goitein used alpha particle data measured by Lyman at LBL to reconstruct a tomographic image [16]. The data were composed of 41 translations and 19 view angles of an elongated phantom, which he reconstructed using an iterative relaxation technique. These reconstructions are the first examples of transmission tomography using charged particles, and also of the energy loss of individual parti-

[2]The first proton facility was actually installed in Clatterbridge, England in 1989, but it is specialized for cancers of the eye and has a maximum energy of only 60 MeV [8, 9].

cles being utilized for contrast. Three years later, a study by Crowe et al. showed that alpha particle CT had a dose advantage over x-ray CT [17].

Cormack and Koehler joined forces in 1976 to construct the first prototype proton CT system at HCL [18]. The system used a proton-integrating technique (individual protons were not tracked) and a detector system composed of a scintillator crystal interfaced to photomultiplier tubes in order to measure the residual energy of protons traversing a phantom. From the residual energies, water-equivalent paths were inferred and Cormack's line integral theory was used to reconstruct the images. Although no reconstructed image was included in the publication, it was shown that density differences of 0.5% could be easily distinguished in the data.

Hanson et al. spearheaded the next effort in proton CT at Los Alamos Laboratory in New Mexico in the late 1970's and early '80's [19]. Hanson used a 240 MeV beam to scan a 30 cm diameter phantom and a 192 MeV beam to scan a 19 cm phantom. Both phantoms were submerged in a water bath and contained cylindrical inserts of varying density. Protons that passed through the phantom without significant deflection were measured, individually, by a position-sensitive proportional chamber, two scintillation counters and a hyperpure germanium detector which measured their residual energy. The results of their 1981 study [20] indicated a dose efficiency of nearly an order of magnitude in proton CT as compared with x-ray CT, however the spatial resolution in the proton images was less than half that of the spatial resolution in the x-ray images. A year later Hanson et al. performed a study using human tissue samples obtained from the New Mexico Medical Examiner's office and compared proton CT scans of the samples to x-ray CT scans. Notably, they found that high-density artifacts present in x-ray CT scans were not evident in proton CT, which they suggested made the technique

particularly desirable for scanning inhomogeneous materials. Unfortunately, their results were fairly consistent with their previous experiments; although proton imaging seemed to have a dose advantage over x-rays, spatial resolution was far reduced. In the conclusions of this study, Hanson et al. writes: "If the only advantage of the proton technique is better dose utilisation, the anticipated extra expense of implementing charged particle CT may not be justified for widespread routine diagnostic studies" [21].

Perhaps it was this statement that led to the abandonment of the topic for more than a decade, but with the popularity of proton radiotherapy on the rise, and the construction of new hospital-based radiotherapy facilities in the 1990s, it became clear to some that the dose-advantage of proton CT was indeed *n*ot the only advantage of the technique.

In the mid-'90s, Schneider and Pedroni of the Paul Scherrer Institute (PSI, Switzerland), developed a system for proton radiography and published an important article, which described a method using proton radiography in order to examine the accuracy of proton range prediction using x-ray CT [22, 23]. The study showed that standard stoichiometric calibration using tissue-equivalent samples was inadequate, resulting in very large range differences when compared with a proton radiograph. Further, they showed that these deviations could be significantly reduced by using organ-specific calibrations. The primary limitation of their study was the speed of data acquisition which was only about 1 kHz[3].

The University of California, Santa Cruz and Loma Linda University collaboration formed in 2003 after a group of high-energy and detector physicists met at Brookhaven National Laboratory to discuss the possibility of proton CT hardware development for clinical applications. Over the years, other groups have

[3]Ten years later, using a faster system, Schneider et al. published the first image of a live canine subject [24].

been included in this collaboration, especially Northern Illinois University (Illinois), California State University, San Bernardino (California), and now Baylor University (Texas). The first generation prototype was completed in 2010, and in 2011, the collaboration received funding to build a Phase-II system. It is the development and testing of this system that will be the subject of this thesis.

There are now many groups around the world studying proton imaging besides the UCSC/LLU/Baylor collaboration including groups at Northern Illinois University (United States), Heidelberg University (Germany), and at Niigata University (Japan), and collaborations such as PRIMA (PRoton IMAging) in Italy, and the Proton Radiotherapy Verification and Dosimetry Applications (PRaVDA) consortium in London, UK, as well as others [25].

1.3 Challenges of Proton Imaging

Although the benefits of proton CT could be vast, developing a clinical proton CT system is wrought with challenges. These challenges include both the physical limitations of the system, such as multiple Coulomb scattering (MCS) and other stochastic processes, as well as the mechanical difficulty of building a data acquisition system fast enough to be useful in the clinical environment.

MCS is the physical process that scatters charged particles from their incident direction without changing their total momentum. It is a stochastic process that involves many individual elastic interactions between a charged particle and the nuclei of the material through which it is traveling. Although each nuclear interaction produces a complex distribution of scattering angles, the overall result, when combining many of these processes is a fairly Gaussian distribution that is simple to model.

In the Gaussian approximation of MCS, the width of the distribution is given by

$$\sigma_\theta \propto \sqrt{\frac{x}{X_0}} \tag{1.1}$$

where X_0 is the characteristic "radiation length" of the material, which is defined as the mean distance over which a high energy electron loses all but 1/e of its energy. The angle of scattering is highly correlated with scattering in the x- and y-directions (assuming the particle is initially traveling along z) as well. The width of the σ_x and σ_y distributions are given by

$$\sigma_x = \sigma_y = \frac{1}{\sqrt{3}} x \sigma_\theta. \tag{1.2}$$

However, due to relativistic effects, the quantity σ_θ is also inversely proportional to the quantity βp where β is the relativistic velocity of the particle compared with the speed of light, c, and p is the momentum of the particle. An accurate model for MCS of protons in a material must take energy loss into account. For high energy physicists this correction is often unnecessary since the energy lost compared to the total energy negligible, but this is not necessarily the case in proton CT, where it is possible for the proton to lose a large fraction of its initial energy before exiting the phantom. Therefore we modify the above equations to account for energy loss:

$$\sigma_\theta^2 = \Theta_0^2 \int_0^z \frac{1}{\beta(z')^2 p(z')^2} \frac{\mathrm{d}z'}{X_0} \tag{1.3}$$

$$\sigma_y^2 = \Theta_0^2 \int_0^z \frac{(z-z')^2}{\beta(z')^2 p(z')^2} \frac{\mathrm{d}z'}{X_0} \tag{1.4}$$

where $\Theta_0 = 13.6$ MeV/c is a universal constant, and z' is the depth of the particle in the material [26].

MCS is roughly Gaussian for the small deflection angles we see in proton CT, and therefore the Gaussian (Lynch-Dahl) approximation is sufficient for the

central 98% of the projected angular distribution with the width given by

$$\theta_0 = \frac{13.6\text{MeV}}{\beta cp} z \sqrt{\frac{x}{X_0}} \left[1 + 0.038 \log \left(\frac{x}{X_0} \right) \right] \tag{1.5}$$

where x is the depth of the particle in the material and z is the charge of the particle. This equation comes from a fit to the Moliere distribution for singly charged particles with $\beta = 1$ for all Z and is accurate to 11% or better for $10^{-3} < x/X_0 < 100$ [27].

MCS is the physical limitation that puts a lower boundary on achievable spatial resolution. Although it can be modeled relatively well, its stochastic nature makes it impossible to know the exact path protons took through the object. Previous studies using simulation have shown that even the most likely path of a particle through a material can differ from the true path by several millimeters on average near the center of the phantom [28]. Heavier ions such as helium, carbon and oxygen are less susceptible to MCS due to their higher mass, but pose another problem since their higher masses make these particles harder to accelerate to the energies required for imaging. Higher energy protons also scatter less since they have less time to interact with nearby nuclei but lose less energy passing through a phantom thereby reducing image contrast. The optimal energy for proton imaging is yet unknown, but until devices can be tested, it is expected to be between 200 and 250 MeV [26].

Since protons lose kinetic energy in discrete steps through stochastic ionization processes, a monoenergetic beam entering a phantom shows an energy spread after passing through a homogeneous medium of a particular thickness. As a result, straggling within the phantom and the range detector limits the precision of the energy loss measurement. For energy losses not exceeding 20% of the initial energy, this distribution can be well described by a Gaussian function, with a variance

according to Bohr's theory of

$$\sigma_B^2 = \eta_e K \int_0^d \frac{1 - \frac{1}{2}\beta^2 E(E_{\text{in}}, x)}{1 - \beta^2 E(E_{\text{in}}, x)} \, \mathrm{d}x \tag{1.6}$$

where η_e is the electron density, $K = 170$ MeV/cm, d is the depth of the particle in the material and $E(E_{\text{in}}, x)$ is the mean energy of protons of incident energy E_{in} after traversing a path of length, x through the material. However, for energy losses larger than 20% of the initial energy, Bohr's theory is inaccurate and the distribution becomes skewed towards high energy losses. For this case the theory of Tschalär is more accurate [29, 30]. These effects determine the inherent physical limit for density resolution in proton imaging.

In addition to energy straggling, many protons undergo inelastic scattering events. When this happens, the primary proton transfers energy to a secondary particle, usually another proton. When these inelastic events occur after the proton has lost a significant fraction of its initial energy, the secondary particle may stop in the phantom and make unwanted contributions to patient dose. On the other hand, if the primary proton has a large amount of energy left, sufficient energy may be transfered to the secondary particle that it scatters out of the phantom and may be detected by the apparatus. In this case, these protons must be eliminated before image reconstruction, and typically can be identified because they have significantly lower energy than other protons passing through the same voxels [31].

The development of a clinical scanner also requires that a full scan can be obtained as quickly as possible. State-of-the-art x-ray CT scanners can acquire a full scan in seconds, however the requirement for individual particle tracking in proton CT imposes a limit on how quickly data can be acquired. It would be reasonable to ask a patient to sit still for a scan lasting 6 minutes, however a

shorter amount of time would be better still. Tracking and measuring the residual energy of individual particles requires fast sensors and fast data acquisition and management. In Hanson's 1979 paper he wrote: "In the present discussion we will concentrate on the feasibility of scanning a patient in 10 s with a proton beam. The objective would be to accumulate 10^8 events with which to make a CT reconstruction...At first sight the data handling problems associated with a 10 MHz data rate appear formidable. However, upon closer inspection, these problems are found to be soluble with present-day technology with only a modest amount of multiplexing and parallel processing" [19]. Indeed, this statement was optimistic as achieving a 10 MHz data rate is still problematic, even with nearly four additional decades of technology.

1.4 Economic Implications of Proton Therapy and Imaging

It is now widely accepted that the physical properties of protons give them an advantage over conventional, high-energy x-rays as a source of radiation for the treatment of tumors. As such, proton therapy has bourgeoned over the past 25 years and will likely accelerate its growth in the coming years, though the question remains among many whether or not proton radiotherapy is cost-effective.

An oft-cited, 2003 paper by Goitein and Jermann studied the relative costs of x-ray and proton radiation therapies. The authors exhaustively compared the costs of building and operating a proton facility compared with an intensity-modulated radiation therapy (IMRT) center[4] including construction, operation and treatment costs. The study found that the cost of construction alone for a

[4]IMRT using x-rays is considered the state-of-the-art x-ray treatment currently available.

proton facility was roughly 62.5 million EUR[5] compared with 16.8 million EUR for IMRT. This budget is largely dominated by the cost of equipment for proton therapy, which requires the construction of a synchrotron or cyclotron, as well as gantries, all of which are much larger than state-of-the-art electron linacs. After considering the costs of personel, equipment, building operation, business costs and other fixed and variable expenses, Goitein and Jermann estimated the cost per fraction[6] to be roughly 1,025 EUR for proton therapy compared with 425 EUR for IMRT, a ratio of 2.4:1. Adjusting for improvement in the efficiency of proton therapy, over the course of 1-2 decades, the authors suggested this ratio may improve to 2.1:1 or better, thus they conclude that proton therapy is likely to continue to cost more than IMRT of comparable sophistication by a factor of two for the foreseeable future, largely dominated by the cost of building and maintaining the facility [32].

Despite the obvious need for studies investigating the cost-effectiveness of proton therapy, there have been very few to date. A literature review by Lodge et al. indicates that there are too few randomized controlled trials (RCTs) to draw firm conclusions about the cost-effectiveness of proton (and other hadron) therapies. The authors report that though several studies have been done, few of them draw firm conclusions as to whether proton therapy is superior to IMRT, except in the cases of ocular tumors and skull base tumors, in which proton therapy has been conclusively found to be superior to conventional x-rays [33]. Lundkvist et al. have published several studies on cost-effectiveness of proton therapies including one regarding childhood medulloblastoma and one breast cancer. They found by comparing the lifetime expenses and quality-adjusted life years (QALYs) of

[5]The model facility contains two treatment rooms with gantries. Additional gantries, fixed beamlines and maintenance costs are not included in this estimate

[6]Patients receive anywhere from 5 to 40 fractions over the course of treatment.

childhood medulloblastoma survivors who underwent proton therapy with those who underwent conventional x-ray therapy, that those who were appropriately selected to receive protons, save on the order of 20,000 EUR over the course of their lives and gain 0.68 QALYs, mostly due to reductions in IQ loss and reductions in growth hormone (thyroid) deficiencies [34]. The authors' study on treating breast cancers with protons similarly indicated that proton therapy can be cost-effective if appropriate risk groups are selected for the therapy [35].

The expense of the infrastructure is certainly an obstacle for the health-care system in terms of resource allocation, and very much depends on the country in which the facility will be built. Some countries, such as the United States pass along the expense of the business to their patients, resulting in an ethical dilemma of whether a patient should be asked to pay twice as much for a treatment that has not been proven to be at least twice as effective. Other countries forgive the cost of constructing a facility, and therefore, the government must be convinced that the expense will result in increased benefit to the public.

Goitein and Jermann suggest that compared with x-ray technology, which is well established and part of a highly competitive market, proton technology has a larger scope for improvement. However, in order for large-scale growth in the field to occur two things must happen: first, proton CT must be proven to be superior to IMRT in more applications. Second, it must be made cheaper.

To address the first point, more studies need to be done to compare the effectiveness of proton therapy with conventional x-ray therapies. Concerns have arisen regarding the ethicality of RCTs when, from a theoretical point of view, one is certainly better than the other. As a result, other means of assessment may be required [36]. The primary concern, at this point in time, is to reduce range uncertainty, since this is the primary obstacle in the way of realizing the full

potential of charged particle therapy. This is the primary goal of proton imaging.

To address the second issue, some have suggested that since the gantry is the largest component of the system, requiring a huge of amount of infrastructure and equipment, perhaps the need for many gantries can be eliminated. The radiosurgery beamline at Massachusetts General Hospital, for example, uses a fixed beamline and instead manipulates the orientation of the patient [37]. This would certainly simplify the required infrastructure needed for proton therapy, however could be uncomfortable for the patient, and is not feasible in all cases.

1.5 Chapter Summary

Although it is now widely accepted that the physical properties of protons give them an advantage over conventional x-ray therapy, there still remains the question as to whether proton therapy is a cost-effective procedure. Proton therapy has been shown to be superior for treating certain types of cancers such as ocular tumors, pediatric cancers and head and neck cancers, however a lack of controlled studies has led to inconclusiveness regarding its outcomes for other types of cancers. This question is largely relevant due to the expense required for building and operating a proton therapy facility.

Since proton imaging requires the same equipment, personnel and facilities as proton therapy, many of the concerns associated with proton therapy are directly applicable to proton imaging as well, and certainly pose one of the biggest obstacles toward developing the technology further. It is largely because of these facts that our group has chosen to focus on developing a proton CT scanner for application in the less-controversial cases of head and neck tumors and pediatric cancers.

Since x-ray CT technology is so well established and understood, there is a lot of resistance in imaging communities to develop proton imaging technology. Some believe that dual-energy CT, an imaging modality which utilizes both keV and MeV x-rays for improved accuracy and contrast, will provide sufficiently robust calibrations. Still, proton imaging would require little additional infrastructure to an already operating proton therapy facility and therefore to others remains a strongly viable and exciting option. Indeed the potential longevity of proton imaging is largely dependent on the ongoing success of proton therapy treatment, and likewise, proton imaging could be the long sought-after solution to the range-uncertainty problem.

Chapter 2

Principles of Proton Imaging

2.1 Introduction

The theory of straight-ray CT is presently well understood. Conveniently, many of the concepts established for straight-ray CT carry directly over to proton CT, with a few key differences. In proton CT we define two coordinate systems: an image system, $(x\,, y\,, z)$, which is attached to the phantom or patient and adheres to digital imaging and communication (DICOM) standards, and a detector coordinate system, $(t\,, u\,, v)$, that is fixed to the detector, which rotates relative to the phantom. The relationship between these two coordinate systems is illustrated in fig. 2.1. The t and v coordinates define the horizontal and vertical axes, respectively, of the particle detectors, while the beam is defined to be in the u-direction. The origins of both coordinate systems are overlaid at the $u = 0$ plane.

For tomographic reconstruction, a reconstruction volume is defined to be a cylindrical space that is large enough to contain the entirety of the phantom, but not so large as to contain too many empty voxels. Image reconstruction is accomplished first by determining the boundary of the object being imaged, and

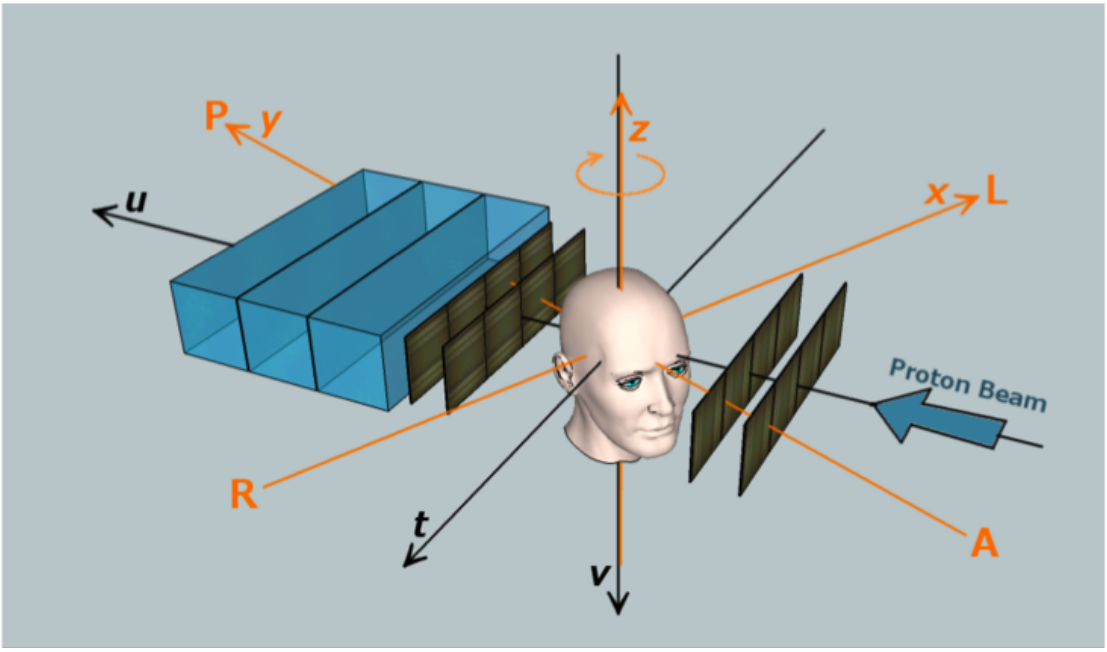


Figure 2.1: The image coordinates, $(x\,,y\,,z)$ and the detector coordinates, $(t\,,u\,,v)$ intersect at one another's origins at the center of the reconstruction volume.

then by determining the most likely path of each proton through the object.

A key distinction between x-ray imaging and proton imaging are the units of measurement. In x-ray imaging, an image is constructed based on the relative attenuation of x-rays after they have passed through the object. The attenuation is directly measured by x-ray detectors, whereas in proton CT, the quantity of interest is the water equivalent path length (WEPL) of each proton. Along with the path information, WEPL can then be used to reconstruct a 3-dimensional map of the relative stopping power (RSP) of the proton with respect to water within the object (section 2.2).

With a good estimate of the path of each proton, as well as its WEPL, iterative reconstruction methods (section 2.6) can be applied to the data in order to invert the resulting system of equations in order to obtain an image vector. In this chapter, I will discuss, in detail, these fundamental concepts of proton imaging

and the reconstruction of proton CT data.

2.2 Relative Stopping Power

Protons used for imaging have sufficient energy to traverse the phantom and deposit the Bragg peak in a downstream range detector or calorimeter. At 200 MeV, the projected range of protons in water is 25.9 cm according to the NIST PSTAR database [38], which is sufficient to traverse an adult human skull.

The mean energy loss per unit path length is also known as the stopping power and it is well described by Bethe-Bloch theory. The Bethe-Bloch equation for a proton is:

$$-\frac{\mathrm{d}E}{\mathrm{d}x}(E,\boldsymbol{r}) = K\frac{Z}{A}\frac{1}{\beta^2}\left[\frac{1}{2}\log\left(\frac{2m_e c^2\beta^2\gamma^2 T_{\max}}{I^2} - \beta^2\right)\right]. \tag{2.1}$$

Here, β is the relativistic velocity of the particle with respect to the speed of light (c), γ is the relativistic correction factor, and m_e is the mass of an electron. K is a constant defined as

$$K = 4\pi N_{\mathrm{A}} r_e^2 c^2 \tag{2.2}$$

where N_{A} is Avogadro's number and r_e is the orbital radius of the electron. The constant K evaluates to 0.307075 MeV·cm^2. Z and A are the atomic number and mass of the absorber, respectively and the ratio can be approximated for an arbitrary absorber composed of i elements using

$$\left(\frac{Z}{A}\right)_{\mathrm{mat}} = \sum_i \left(w_i \frac{Z_i}{A_i}\right) \tag{2.3}$$

where w_i is the fractional weight of the i'th element. I is the ionization energy (or mean excitation energy) of the material, which can be approximated for an

arbitrary material composed of i elements using

$$\langle I \rangle = \exp\left(\frac{\sum_i w_i (Z_i/A_i) \log I_i}{\sum_i w_i (Z_i/A_i)}\right) \tag{2.4}$$

where I_i is the ionization energy of the i'th element [39]. Finally $T_{\max}$ is the maximum amount of energy that can be transfered from the proton to the absorber in a single collision, and is given by $T_{\max} = 2m_e c^2 \beta^2 \gamma^2$ to first order.

It is convenient, for proton imaging, to write the Bethe-Bloch formula in the following form:

$$S(I(\boldsymbol{r}), E(\boldsymbol{r})) = K \frac{Z}{A} \frac{1}{\beta^2(E)} \left[\log \frac{2m_e c^2 \beta^2(E)}{I(\boldsymbol{r})(1-\beta^2(E))} - \beta^2(E)\right] \tag{2.5}$$

where $S(I(\boldsymbol{r}), E(\boldsymbol{r}))$ represents the stopping power of a proton in some material. The stopping power of a proton is strongly energy- (and therefore position-) dependent, which poses a significant problem when the materials through which the proton is traversing are unknown. However, the *relative* stopping power of the proton with respect to water (RSP) is fairly energy-independent. It is convenient to express this ratio as:

$$\mathrm{RSP_m}(E) = \eta_e \left[\frac{\log\left(\frac{2m_e c^2}{I_\mathrm{m}} \frac{\beta(E)^2}{1-\beta(E)^2}\right) - \beta(E)^2}{\log\left(\frac{2m_e c^2}{I_\mathrm{w}} \frac{\beta(E)^2}{1-\beta(E)^2}\right) - \beta(E)^2}\right] \tag{2.6}$$

where I_m is the ionization energy of the material given by eq. (2.4) and I_w is the ionization energy of water which is 75 eV. η_e is the relative electron density which is given by

$$\eta_e = \frac{\rho_\mathrm{m} \sum_i \left(w_i \frac{Z_i}{A_i}\right)}{\rho_\mathrm{w} \left(\frac{Z}{A}\right)_\mathrm{w}} \tag{2.7}$$

where ρ_m is the density of the material, ρ_w is the density of water which is 1.00 g/cm^3 and $\left(\frac{Z}{A}\right)_\mathrm{w} = 0.55509$.

2.3 Mean Energy Loss and Water Equivalent Path Length

In proton radiography and CT, the energy loss or residual range of protons traversing an object must be converted into WEPL, which is the integral of the object's RSP along the path of the proton. Mathematically this is expressed:

$$\text{WEPL} = \int_{\ell} \varrho \, \mathrm{d}\ell \tag{2.8}$$

where here we use ϱ to represent the RSP and ℓ is the path of the proton through the object.

For proton energies in the range between 30 and 250 MeV, the variation of RSP with proton energy is negligible. For example, for brain tissue (as defined by the ICRP) the difference of the RSP at 30 and 200 MeV is only 0.07%. For practical purposes, therefore RSP is considered to be independent of proton energy.

If the residual energy of the proton is measured after it exits the phantom, then an estimate of the WEPL can be obtained by numerically integrating the reciprocal of the stopping power

$$\text{WEPL} = \int_{E_{\text{out}}}^{E_{\text{in}}} \frac{\mathrm{d}E}{S(I,E)}. \tag{2.9}$$

A direct calibration between residual energy and WEPL can be obtained using a calibration procedure that correlates energy loss with the water equivalent thickness (WET) of a series of degraders with varying WET placed in the proton scanner, such as one described by Hurley et al. [40].

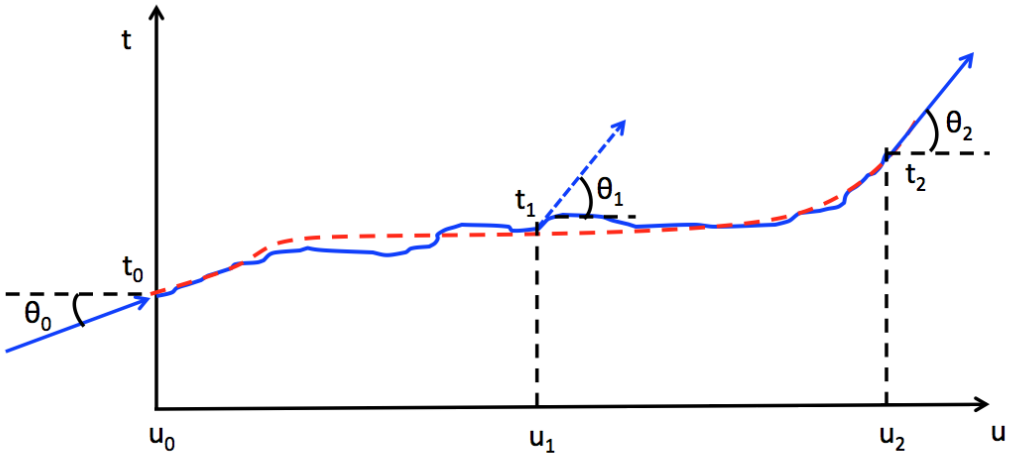


Figure 2.2: Scattering geometry in the t-u plane. The solid blue curve represents the (exaggerated) true path of a particle while the dashed red curve describes the most likely path given the entry and exit vectors.

2.4 Most Likely Path

Unlike x-rays, which travel in straight lines through an object being imaged, protons undergo multiple Coulomb scattering (MCS). Therefore the paths of protons through an arbitrary material are not well characterized and must be approximated using a "most likely path" (MLP). The resulting MLP will asymptotically approach the known entering and exiting proton trajectories for points approaching the corresponding surfaces.

The MLP (fig. 2.2) of a proton traveling in water is well described. The assumption that the proton is traveling in water is valid since most human tissues differ from water by only very small amounts. Early formalisms for MLP such as those by Schneider and Pedroni [41] and by D.C. Williams [26] required evaluating complicated ratios, polynomials and integrals. This model was validated experimentally in the study by Bruzzi et al. [42]. The more recent, compact

matrix-based formalism, developed by Schulte et al. [28], utilizes Bayesian statistics to determine transverse displacement and the direction of maximum likelihood at any intermediate depth in water, given known coordinates and angles of entry and exit to the object. I describe this method below.

Finding the MLP is a maximum likelihood problem that can be solved within the Bayesian framework: given the known entry vector, find the interior vector at a given depth that maximizes the posterior likelihood. A right-handed coordinate system $(t\,,u\,,v)$ in the reference frame of the detector system is defined such that t describes the horizontal coordinate, v the vertical coordinate, and the proton beam is incident in the u-direction. Since scattering in the horizontal and vertical directions are independent processes, we can confine our discussion to two dimensions, for example in the t-u plane. In this case, the location and direction of a proton at any depth u_1 is given by the vector

$$\boldsymbol{y}_1 = \begin{pmatrix} t_1 \\ \theta_1 \end{pmatrix}. \tag{2.10}$$

There is a prior likelihood of finding the proton with vector $\boldsymbol{y}_1$ at depth u_1 given the the information of the proton tracking before it enters the reconstruction volume, $L(\boldsymbol{y}_1|\text{entry data})$, a likelihood of finding the proton with available exit information given $\boldsymbol{y}_1$ at depth u_1, $L(\text{exit data}\,|\boldsymbol{y}_1)$, and a posterior likelihood that the proton had parameters $\boldsymbol{y}_1$ at depth u_1 given the observed exit information, $L(\boldsymbol{y}_1|\text{exit data})$. Then according to Bayes' theorem [43] the prior and posterior likelihood are related by:

$$L(\boldsymbol{y}_1|\text{exit data}) = L(\boldsymbol{y}_1|\text{entry data})L(\text{exit data}\,|\boldsymbol{y}_1). \tag{2.11}$$

The most likely location of the proton at a depth u_1 can be found by obtaining

the vector $\boldsymbol{y}_1$ for which the posterior likelihood is a maximum, i.e.

$$\boldsymbol{\nabla} L(\boldsymbol{y}_1|\text{exit data}) = \begin{pmatrix} \partial t_1 \\ \partial \theta_1 \end{pmatrix} L(\boldsymbol{y}_1|\text{exit data})\Big|_{\boldsymbol{y}_1 = y_{\text{MLP}}} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{2.12}$$

The Gaussian approximation of generalized Fermi-Eyges theory of MCS [44], which is sufficient for the derivation of the MLP, is applied. Assuming that a proton enters the phantom at $u = 0$ with zero lateral displacement and zero angle, the prior likelihood density function of vector $\boldsymbol{y}_1$ given the entry information can be expressed as:

$$L\left(\boldsymbol{y}_1 \middle| \boldsymbol{y}_0 = \mathbf{0}\right) = \exp\left(-\frac{1}{2}\boldsymbol{y}_1^T \boldsymbol{\Sigma}_1^{-1} \boldsymbol{y}_1\right) \tag{2.13}$$

where $\boldsymbol{\Sigma}_1^{-1}$ is the inverse of the symmetric positive definite scattering matrix

$$\boldsymbol{\Sigma}_1 = \begin{pmatrix} \sigma_{t_1}^2 & \sigma_{t_1\theta_1}^2 \\ \sigma_{t_1\theta_1}^2 & \sigma_{\theta_1}^2 \end{pmatrix}, \tag{2.14}$$

the elements of which correspond to the variances and covariances of t_1 and θ_1. The elements of the scattering matrix can be obtained using:

$$\sigma_{t_1}^2(u_0, u_1) = E_0^2 \left(1 + 0.038 \log \frac{u_1 - u_0}{X_0}\right)^2 \int_{u_0}^{u_1} \frac{(u_1 - u)^2}{\beta(u)^2 p(u)^2} \frac{\mathrm{d}u}{X_0} \tag{2.15}$$

$$\sigma_{\theta_1}^2(u_0, u_1) = E_0^2 \left(1 + 0.038 \log \frac{u_1 - u_0}{X_0}\right)^2 \int_{u_0}^{u_1} \frac{1}{\beta(u)^2 p(u)^2} \frac{\mathrm{d}u}{X_0} \tag{2.16}$$

$$\sigma_{t_1\theta_1}^2(u_0, u_1) = E_0^2 \left(1 + 0.038 \log \frac{u_1 - u_0}{X_0}\right)^2 \int_{u_0}^{u_1} \frac{u_1 - u}{\beta(u)^2 p(u)^2} \frac{\mathrm{d}u}{X_0} \tag{2.17}$$

which are given by MCS theory. In eqs. (2.15) to (2.17), $E_0 = 13.6$ MeV, and we assume $X_0 = 36.1$ cm, the radiation length of water. We also assume that the coordinates $(t\,, u)$ and angle of entry and exit from the phantom have been measured, where the subscript "0" will represent the entry information while the subscript "2" will represent the exit information.

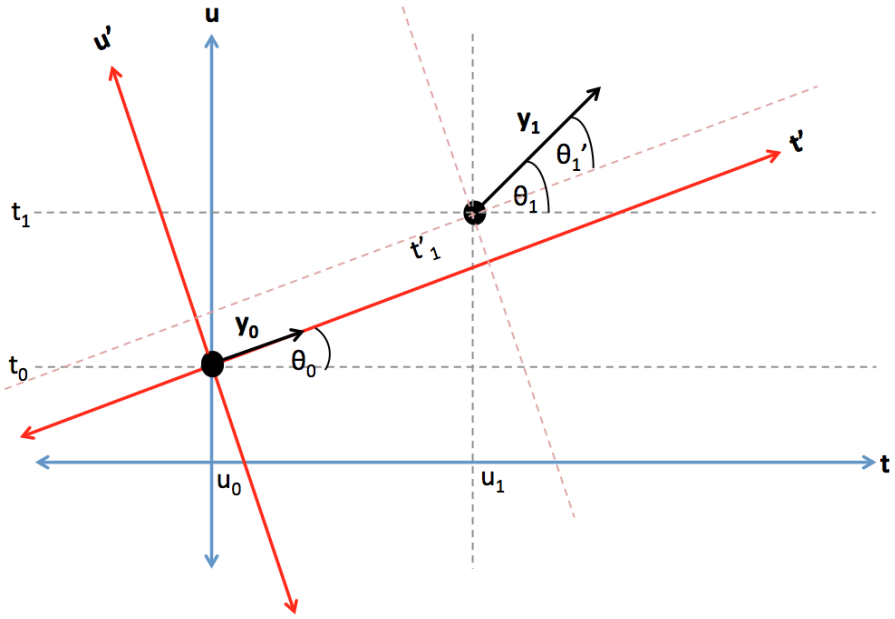


Figure 2.3: The elements of the vector $\boldsymbol{y}_1$ must be transformed from the coordinate system $(t\,, u)$ into $(t'\,, u')$ in order to apply eq. (2.13).

In order to use eq. (2.13), the coordinates of the vector $\boldsymbol{y}_1$ must be transformed as shown in fig. 2.3 such that

$$\theta_1' = \theta_1 - \theta_0 \tag{2.18}$$

$$t_1' = \cos(\theta_0)\left(t_1 - t_0 - \tan\theta_0(u_1 - u_0)\right). \tag{2.19}$$

Applying the small angle approximation, which is valid for MCS, eq. (2.19) becomes simply:

$$t_1' = t_1 - t_0 - \theta_0(u_1 - u_0). \tag{2.20}$$

This can be written in compact matrix form as:

$$\boldsymbol{y}_1' = \boldsymbol{y}_1 - \boldsymbol{R}_0\boldsymbol{y}_0 \tag{2.21}$$

where $\boldsymbol{R}_0$ is the transformation matrix defined as:

$$\boldsymbol{R}_0 = \begin{pmatrix} 1 & u_1 - u_0 \\ 0 & 1 \end{pmatrix}. \tag{2.22}$$

The prior likelihood of $\boldsymbol{y}_1$ given $\boldsymbol{y}_0$ (eq. (2.13)) is now

$$L(\boldsymbol{y}_1|\boldsymbol{y}_0) = \exp\left(-\frac{1}{2}\left(\boldsymbol{y}_1^T - \boldsymbol{y}_0^T\boldsymbol{R}_0^T\right)\boldsymbol{\Sigma}_1^{-1}\left(\boldsymbol{y}_1 - \boldsymbol{R}_0\boldsymbol{y}_0\right)\right) \tag{2.23}$$

where in order to obtain the first term in parenthesis we have applied the properties of the transpose:

$$(\boldsymbol{A}+\boldsymbol{B})^T = \boldsymbol{A}^T + \boldsymbol{B}^T \tag{2.24}$$

$$(\boldsymbol{A}\boldsymbol{B})^T = \boldsymbol{B}^T\boldsymbol{A}^T. \tag{2.25}$$

It is now straightforward to obtain the likelihood of the exit vector $\boldsymbol{y}_2$ at depth u_2 given $\boldsymbol{y}_1$ at depth u_1 by letting $\boldsymbol{y}_0 \to \boldsymbol{y}_1$ and $\boldsymbol{y}_1 \to \boldsymbol{y}_2$ in the eqs. (2.15) - (2.23) such that the prior likelihood of $\boldsymbol{y}_2$ given $\boldsymbol{y}_1$ is

$$L(\boldsymbol{y}_2|\boldsymbol{y}_1) = \exp\left(-\frac{1}{2}\left(\boldsymbol{y}_2^T - \boldsymbol{y}_1^T\boldsymbol{R}_1^T\right)\boldsymbol{\Sigma}_2^{-1}\left(\boldsymbol{y}_2 - \boldsymbol{R}_1\boldsymbol{y}_1\right)\right). \tag{2.26}$$

Finally, we can obtain the posterior likelihood by combining eqs. (2.23) and (2.26) as in eq. (2.11). This yields a posterior likelihood,

$$L(\boldsymbol{y}_1|\boldsymbol{y}_2) = \exp\left(-\frac{1}{2}\left(\left(\boldsymbol{y}_1^T - \boldsymbol{y}_0^T\boldsymbol{R}_0^T\right)\boldsymbol{\Sigma}_1^{-1}\left(\boldsymbol{y}_1 - \boldsymbol{R}_0\boldsymbol{y}_0\right) + \left(\boldsymbol{y}_2^T - \boldsymbol{y}_1^T\boldsymbol{R}_1^T\right)\boldsymbol{\Sigma}_2^{-1}\left(\boldsymbol{y}_2 - \boldsymbol{R}_1\boldsymbol{y}_1\right)\right)\right). \tag{2.27}$$

Maximizing the posterior likelihood requires minimizing the exponent in eq. (2.27) which we will call χ^2. Expanding eq. (2.27) gives

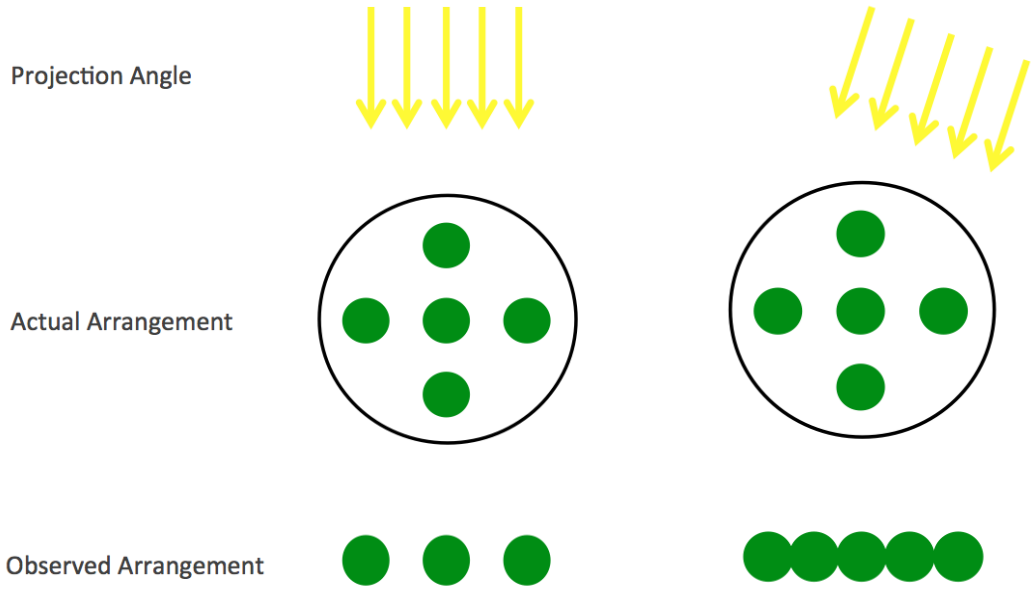
$$\begin{aligned}\chi^2 = &\left(\boldsymbol{y}_1^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{y}_1 - 2\boldsymbol{y}_0^T\boldsymbol{R}_0^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{y}_1 + \boldsymbol{y}_0^T\boldsymbol{R}_0^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{R}_0\boldsymbol{y}_0\right)\\ &+ \left(\boldsymbol{y}_2^T\boldsymbol{\Sigma}_2^{-1}\boldsymbol{y}_2 - 2\boldsymbol{y}_1^T\boldsymbol{R}_1^T\boldsymbol{\Sigma}_2^{-1}\boldsymbol{y}_2 + \boldsymbol{y}_1^T\boldsymbol{R}_1^T\boldsymbol{\Sigma}_2^{-1}\boldsymbol{R}_1\boldsymbol{y}_2\right).\end{aligned} \tag{2.28}$$

Differentiating this with respect to the elements of $\boldsymbol{y}_1$, and setting the result to zero gives us the MLP vector at any value of u

$$\boldsymbol{y}_{\mathrm{MLP}} = \left(\boldsymbol{\Sigma}_1^{-1} + \boldsymbol{R}_1^T\boldsymbol{\Sigma}_2^{-1}\boldsymbol{R}_1\right)^{-1}\left(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{R}_0\boldsymbol{y}_0 + \boldsymbol{R}_1^T\boldsymbol{\Sigma}_2^{-1}\boldsymbol{y}_2\right). \tag{2.29}$$

The error matrix ε_{ij} can also be calculated from χ^2 where

$$\varepsilon_{ij} = 2\left(\frac{\partial^2\chi^2}{\partial t_1\partial\theta_1}\right)^{-1} = \left(\boldsymbol{\Sigma}_1^{-1} + \boldsymbol{R}_1^T\boldsymbol{\Sigma}_2^{-1}\boldsymbol{R}_1\right)^{-1} \tag{2.30}$$



(a) Initial view (b) Rotated view

Figure 2.4: In (a), only three spheres are visible from the observation angle. By rotating the observation angle slightly with respect to the object as in (b), all five spheres become visible.

and the spatial resolution at any u-coordinate along the MLP is given by the square root of the element $\varepsilon_{t_1 t_1}$.

2.5 Filtered Backprojection

The filtered backprojection (FBP) is an image reconstruction technique commonly used in x-ray CT (and other straight-ray tomography) reconstruction. The FBP concept can be understood by a simple thought experiment [45]. Consider a semi-transparent object for which you are trying to guess the internal structure. The object is composed of a transparent cylinder and contains five spheres arranged as shown in fig. 2.4. The object cannot be viewed from the top or bottom but

only from the side. If viewed in the orientation shown in fig. 2.4a, only three objects are visible, but if observed at a slightly different angle, as in fig. 2.4b, all five objects are visible. One can then imagine that if the object is viewed from a large number of angles, one can get a relatively good idea of the internal structure of the object. The FBP carries out this same process in a mathematical fashion. Each viewing angle is a different "projection" of the object. In x-ray tomography individual projections are rendered in terms of the amount of x-ray attenuation. In proton CT each projection is rendered in terms of WET.

The mathematics of the FBP are derived from the Fourier slice theorem, which states that "the Fourier transform of a parallel projection of an object $f(x, y)$ obtained at angle θ equals a line in a 2D Fourier transform of $f(x, y)$." Mathematically speaking we define a projection of an object at angle θ:

$$p(t, \theta) = \int_{-\infty}^{\infty} \mathrm{d}x \int_{-\infty}^{\infty} \mathrm{d}y \, f(x, y) \delta(x \cos\theta + y \sin\theta - t) \tag{2.31}$$

where we have used the delta function to kill all the values except those along the line integrals for a given t, as in fig. 2.5. In this geometry, the standard rotation matrix gives us the coordinate transformation relations

$$\begin{aligned} t &= x \cos\theta + y \sin\theta \\ u &= -x \sin\theta + y \cos\theta. \end{aligned} \tag{2.32}$$

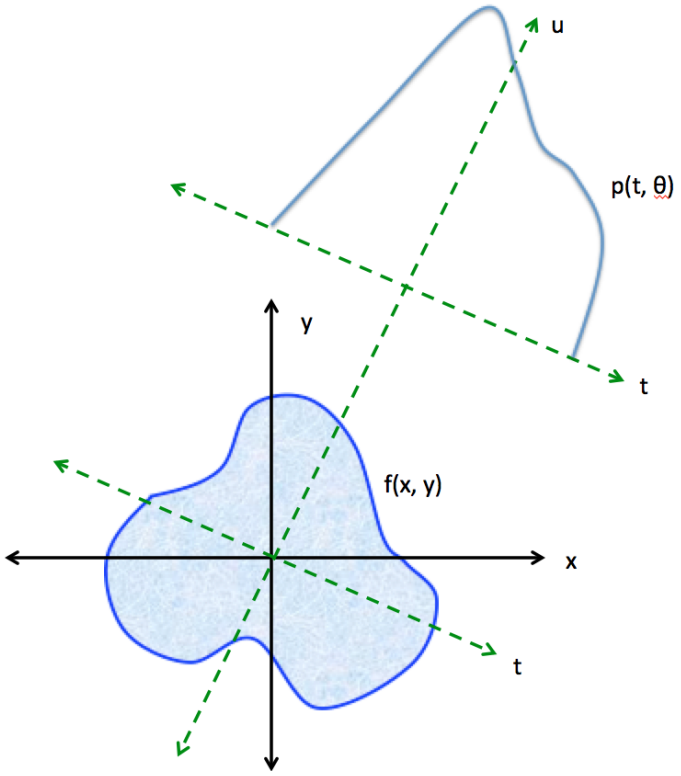


Figure 2.5: A one dimensional projection $p(t,\theta)$ of a two dimensional object, $f(x,y)$ is created by taking the values of the line integrals along a direction of projection, θ.

To begin, we take the Fourier transform of the projection function:

$$\begin{aligned} P(\omega,\theta) &= \mathscr{F}\{p(t,\theta)\} \\ &= \int_{-\infty}^{\infty} \mathrm{d}t\, \mathrm{e}^{-i\omega t} p(t,\theta) \\ &= \int_{-\infty}^{\infty} \mathrm{d}t\, \mathrm{e}^{-i\omega t} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{d}x\,\mathrm{d}y\, f(x,y)\,\delta(x\cos\theta + y\sin\theta - t) \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{d}x\,\mathrm{d}y\, f(x,y) \int_{-\infty}^{\infty} \mathrm{d}t\, \mathrm{e}^{i\omega t}\delta(x\cos\theta + y\sin\theta - t) \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{d}x\,\mathrm{d}y\, f(x,y)\, \mathrm{e}^{-i\omega(x\cos\theta + y\sin\theta)} \end{aligned} \tag{2.33}$$

We can now show that this is equivalent to the 2D Fourier transform of $f(x,y)$

which is given by:

$$\mathscr{F}\{f(x,y)\} = F(v,w) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{d}x\,\mathrm{d}y\,\mathrm{e}^{-i(vx+wy)} f(x,y). \tag{2.34}$$

Letting $v = \omega\cos\theta$ and $w = \omega\sin\theta$, eq. (2.33) and eq. (2.34) are trivially equivalent,

$$F(\omega\cos\theta, \omega\sin\theta) = P(\omega,\theta), \tag{2.35}$$

and therefore the Fourier slice theorem holds true.

To derive the FBP, we begin with the inverse Fourier transform in polar coordinates:

$$f(x,y) = \int_0^{2\pi}\int_0^{\infty} F(\omega\cos\theta, \omega\sin\theta)\,\mathrm{e}^{i\omega(x\cos\theta + y\sin\theta)}\omega\,\mathrm{d}\omega\,\mathrm{d}\theta. \tag{2.36}$$

Given the coordinate transform relation in eq. (2.32), and the result of eq. (2.35), we can replace $x\cos\theta + y\sin\theta$ with t, and $F(\omega\cos\theta, \omega\sin\theta)$ with $P(\omega,\theta)$ resulting in

$$f(x,y) = \int_0^{2\pi}\int_0^{\infty} \mathrm{e}^{i\omega t} P(\omega,\theta)\,\omega\,\mathrm{d}\omega\,\mathrm{d}\theta. \tag{2.37}$$

It is useful to prove that

$$p(t,\theta) = p(-t,\theta+\pi) \tag{2.38}$$

where

$$\begin{aligned} p(-t,\theta+\pi) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{d}x\,\mathrm{d}y\, f(x,y)\,\delta(x\cos(\theta+\pi) + y\sin(\theta+\pi) - (-t)) \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{d}x\,\mathrm{d}y\, f(x,y)\,\delta(-x\cos\theta - y\sin\theta + t). \end{aligned} \tag{2.39}$$

This further simplifies to

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{d}x\,\mathrm{d}y\, f(x,y)\,\delta(-(x\cos\theta + y\sin\theta - t)) \tag{2.40}$$

but the delta function is even, $\delta(x) = \delta(-x)$, and therefore eq. (2.40) is equal to $p(t, \theta)$. Given this result we also have that

$$P(\omega, \theta) = P(-\omega, \theta + \pi) \tag{2.41}$$

which we can use to write eq. (2.37) in a symmetric fashion,

$$f(x, y) = \int_0^{\pi} \int_0^{\infty} \mathrm{e}^{i\omega t} P(\omega, \theta)\, \omega \, \mathrm{d}\omega \, \mathrm{d}\theta + \int_0^{\pi} \int_0^{\infty} \mathrm{e}^{-i\omega t} P(-\omega, \theta + \pi)\, \omega \, \mathrm{d}\omega \, \mathrm{d}\theta, \tag{2.42}$$

which reduces to the final expression for the FBP:

$$f(x, y) = \int_0^{\pi} \int_0^{\infty} \mathrm{e}^{i\omega t} |\omega| P(\omega, \theta)\ \mathrm{d}\omega \, \mathrm{d}\theta, \tag{2.43}$$

where $|\omega|$ is the ramp function "filter." If we let $Q(\omega, \theta) = |\omega| P(\omega, \theta)$, then

$$f(x, y) = \int_0^{\pi} \int_0^{\infty} Q(\omega, \theta)\, \mathrm{e}^{i\omega t} \, \mathrm{d}\omega \, \mathrm{d}\theta \tag{2.44}$$

$$= \int_0^{\pi} q(t, \theta) \, \mathrm{d}\theta \tag{2.45}$$

where $q(t, \theta)$ is the inverse Fourier transform of the filtered projection $|\omega| P(\omega, \theta)$ [46].

The concept of backprojection is that each value along the function $p(t, \theta)$ is "painted" uniformly along the straight-line path of the ray through the object. We can examine the contribution of each filtered projection to the reconstructed image by understanding that $t = x \cos\theta + y \sin\theta$ represents a straight line that overlaps the ray path that produces the projection. Then, with a parallel beam, one can imagine that in an object with a cylindrical geometry, the more radially central areas of the object are intersected by a larger number of rays, thus giving the central region a heavier weight. This effect is offset by multiplying the Fourier transform of the projections by a filter function that is weighted less in the center and greater at the edges. For N projections, evenly spaced over π radians, the ramp filter $\pi|\omega|/N$ conserves the volume of the object over the summation of projections.

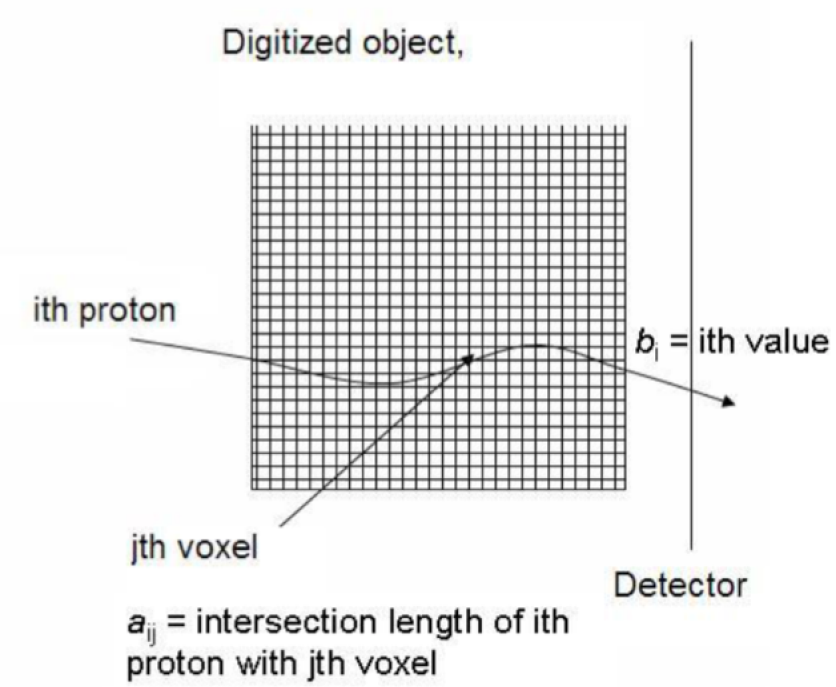


Figure 2.6: Image reconstruction is achieved by solving a system of linear equations $\boldsymbol{A}_{ij}\boldsymbol{x}_j = \boldsymbol{b}_j$.

2.6 Iterative Reconstruction

Iterative projection reconstruction algorithms are currently the preferred reconstruction method in proton CT. Iterative methods are capable of handling non-linear paths, and are capable of incorporating *a priori* knowledge of the phantom.

Image reconstruction is achieved by solving a very large, sparse system of linear equations $\boldsymbol{A}_{ij}\boldsymbol{x}_j = \boldsymbol{b}_j$, where the $\boldsymbol{b}_j$'s are the WEPLs corresponding to the energy lost by the i'th proton along its path, the $\boldsymbol{x}_j$'s are the unknown image vector, and the matrix components, a_{ij} correspond to the length of the intersection of the i'th proton history with the j'th voxel as in fig. 2.6. The physical measurement of each proton equals the total path length multiplied by the RSP along MLP.

A simplified example of iterative reconstruction is given in fig. 2.7. Consider

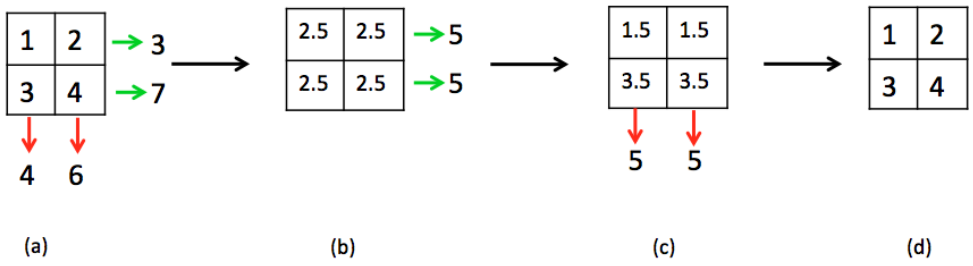


Figure 2.7: Illustration of iterative reconstruction concept. (a) the original values in the 4x4 image space and line integrals at 0 and 90 degrees. (b) The initial guess is a constant distribution. (c) An updated estimation of the object and line integrals. (d) The final estimate of object and line integrals reconstructs the original object.

the simple object shown in fig. 2.7(a). The line integrals along the projection at 0 and 90 degrees sum to 10. Given no previous knowledge of the object, it is reasonable for the initial guess of the object to be a constant distribution, as shown in (b). Taking the line integrals at 0 degrees again, one finds that the top row is overestimated by 2 and the bottom row is underestimated b 2, so in the next iteration, (c), the difference is split by adding 1 to each voxel in the bottom row and subtracting one from each voxel in the top row. Taking the line integrals at 90 degrees, once more, one finds that now the left column is overestimated by 1 and the right underestimated by 1. Once again splitting the difference, the original object is recovered in (d) [45].

In this simple example the original object was recovered in very few iterations (of course, this example could have easily been solved by direct matrix inversion), but when the image space grows, more iterations are required to recover the original object. If one considers a very large image space of size N, each row of the matrix $\boldsymbol{A}$ is a hyperplane in an N-dimensional space. If a unique solution to the system of equations exists, the solution is the point at which all the hyperplanes

intersect. In order to find the point of intersection, an initial guess must be made, and then projected onto each adjacent hyperplane, sequentially, until the difference between the initial solution vector, and the updated vector is null [47]. Mathematically this is represented by:

$$\boldsymbol{x}_i = \boldsymbol{x}_{i-1} - \frac{\boldsymbol{x}_{i-1} \cdot \boldsymbol{a}_i - b_i}{\boldsymbol{a}_i \cdot \boldsymbol{a}_i} \boldsymbol{a}_i \tag{2.46}$$

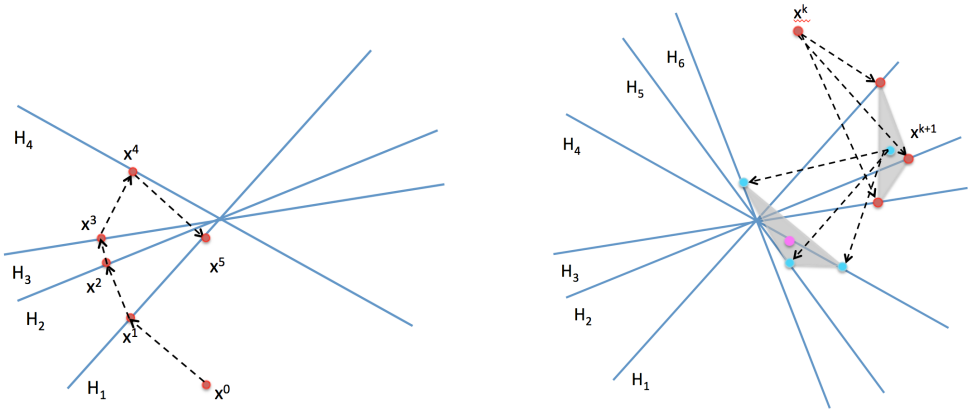
where $\boldsymbol{a}_i$ is the vector that describes the i'th row of the $\boldsymbol{A}$ matrix.

The simplest iterative algorithm is the fully-sequential algebraic reconstruction technique (ART) which is given by:

$$x^{k+1} = x^k + \lambda_k \frac{b_{i(k)} - \left\langle a^{i(k)}, x^k \right\rangle}{||a^{i(k)}||^2} a^{i(k)} \tag{2.47}$$

where k is the iteration number and λ is the relaxation parameter. The interpretation of eq. (2.47) is that given the current iterate x^k, and the hyperplane $H_{i(k)}$ defined by the $i(k)$'th row of the matrix equation, the new iterate, $k+1$ lies on the line through x^k that is perpendicular to $H_{i(k)}$. When the relaxation parameter is 1, x^{k+1} is the orthogonal projection of x^k onto $H_{i(k)}$. Setting the relaxation parameter to $\lambda < 1$, prevents overstepping of the correct solution by allowing the next iterate to have a value along the line segment connecting x^k with its orthogonal projection onto $H_{i(k)}$ [48].

Iterative methods such as fully-sequential ART have been shown to be highly accurate, but are computationally intensive since forward projections must be performed repeatedly [49]. Furthermore, reconstruction time is primarily dependent on the speed of the processing unit and the size of the data set. In proton CT, we modify the sequential ART algorithm to be block-iterative. A schematic depicting the difference between a purely sequential algorithm and a block iterative algorithm are shown in fig. 2.8. Block iterative techniques balance the accuracy



(a) ART algorithm (b) Block-iterative algorithm

Figure 2.8: (a) A fully sequential algorithm updates the image by projecting directly onto hyperplanes. (b) A block-iterative algorithm groups hyperplanes together and projects onto the subspace, significantly decreasing computation time.

achieved by sequential ART, with faster convergence times, since calculations within a block can be implemented in parallel.

In block-iterative reconstruction techniques, hyperplanes are grouped together into blocks. The number of blocks, their sizes and the assignments of hyperplanes to blocks can change dynamically from cycle to cycle as long as the block is representative of the entire data set. Block-iterative reconstruction is discussed in detail in [48–50].

2.7 Chapter Summary

We have now discussed the concepts fundamental to proton imaging including relative stopping power, water equivalent path length and most likely path, as well as image reconstruction methods including the filtered backprojection and iterative projection algorithms. In proton CT, maps of relative stopping power within an object are reconstructed directly from WEPL values of individual pro-

tons. This is accomplished by solving a large, sparse linear system of equations $\boldsymbol{A}_{ij}\boldsymbol{x}_j = \boldsymbol{b}_i$. In order to construct the $\boldsymbol{A}$ matrix, the most likely path of each proton must be determined using the information regarding the entry and exit of each proton from the phantom. The MLP is essentially an optimization problem in the Bayesian framework in which one is trying to maximize the posterior likelihood. The most likely path is essential for accurate reconstruction of the object, since the $\boldsymbol{A}$ matrix is constructed of binary values indicating which voxels were intersected by each proton.

Historically, filtered backprojection algorithms have been very good at solving image reconstruction problems for straight-ray tomography, however for proton CT, more versatile methods must be employed in order to incorporate curved proton paths and *a priori* information into the reconstruction. Iterative techniques can be extremely accurate, but can be computationally very expensive, and so an appropriate algorithm that balances both accuracy and speed must be selected. In subsequent chapters, especially chapter 6, we will discuss the implementation of these algorithms in proton CT.

Chapter 3

The Prototype Proton CT Scanner

3.1 Introduction

Although the the history of proton radiography and tomography goes back half a century, no clinical proton CT scanner presently exists. Modern technologies for particle detectors and computation now hold great promise to realize a system that can rapidly generate measurements of integrated relative stopping power (RSP).

The proton CT collaboration between UC Santa Cruz (UCSC), Santa Cruz, CA, and Loma Linda University Medical Center (LLUMC), Loma Linda, CA, has developed a sequence of increasingly faster proton CT scanners. The Phase 0 scanner, a small-scale prototype, was developed at LLUMC at the beginning of 2007. The system consisted of four 400 µm-thick, single-sided silicon strip detectors (SSD) with a 6.4×6.4 cm^2 sensitive area. These detectors were left over from previous, unrelated experimental work. Two silicon tiles were mounted with their strips orthogonal to one another to make a 2-D-sensitive tracking layer.

Because only four tiles were available, the system was composed of a single tracking plane upstream from the phantom and a single tracking plane downstream from the phantom, making it possible to measure proton position entering and exiting the phantom but not angle. A 4 cm-thick, thallium-doped caesium iodide crystal (CSI(Tl)) with a sensitive area of 6.4×6.4 cm^2 to match the sensitive area of the detectors was used as a calorimeter. Initial measurements of a small, 40 mm-diameter, 30 mm tall, acrylic cylindrical phantom containing high-contrast inserts were made at 100 MeV.

The Phase I prototype (2010) was composed of 16 SSD tiles and a segmented calorimeter with 18 CsI crystals. This system was composed of two particle telescopes, one upstream and one downstream from the phantom, with a sensitive area of approximately 8.9×17.8 cm^2, so that both the entering and exiting position *and* angle of each proton could be measured. This system operated at roughly 10 kHz, requiring upwards of 20 hours for data to be obtained for a full scan consisting of 90 projections at 4° intervals. While the results showed promise (see chapter 5), the slow data rate limited the clinical usefulness of this system.

Most recently, the pCT collaboration has built and successfully operated a Phase-II scanner which operates at >1 MHz. At this rate, it is possible to scan half of a human head in under 10 minutes. This scanner is the main subject of this thesis.

3.2 Experimental Setup

The Phase-II proton CT scanner (fig. 3.1) is comprised of two detector systems, the silicon tracking unit and the 5-stage scintillator, which functions as a hybrid energy/range detector. The silicon tracking detector is composed of two particle

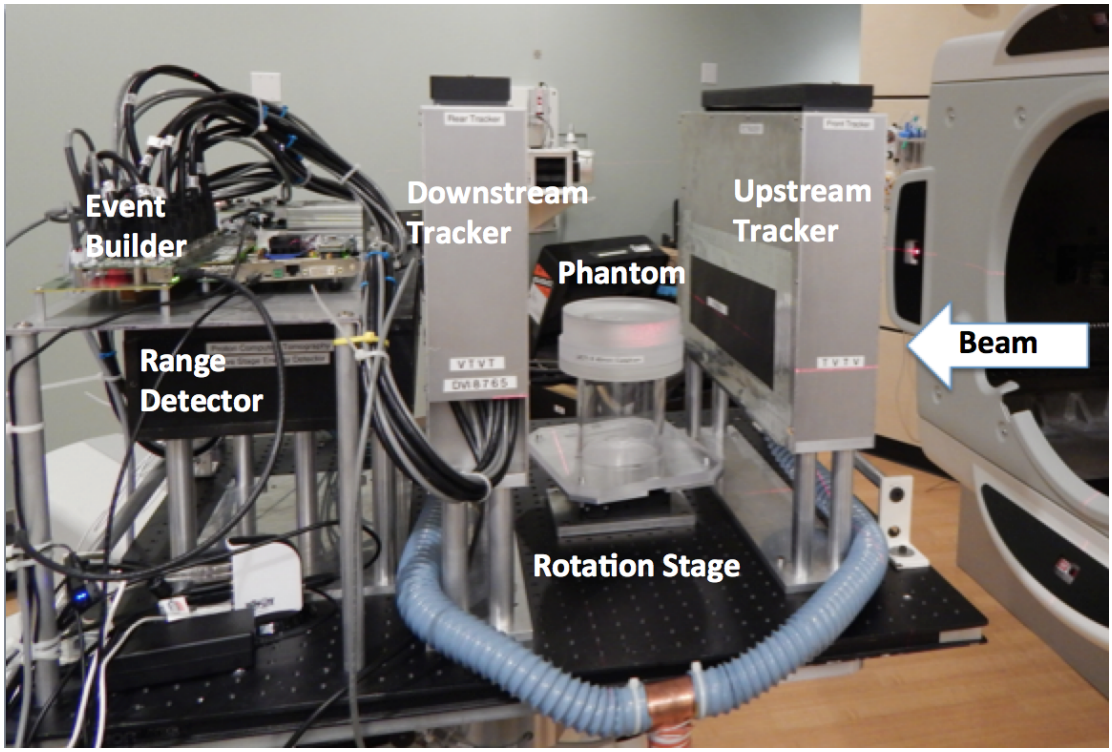


Figure 3.1: Photograph of the Phase-II proton CT head scanner in a proton beam line at the Northwestern Medicine Chicago Proton Center.

telescopes, one upstream and one downstream from the phantom. Each particle telescope is composed of sixteen SSDs arranged in 4 layers (two horizontal-sensitive and two vertical-sensitive) such that each layer is composed of four 1-D-sensitive tiles. The total sensitive area of the tracker is therefore roughly 8.9×35.2 cm^2. The range detector is composed of five stages of UPS-923A polystyrene-based scintillator, such that each stage a proton entirely passes through contributes directly to the proton's measured range, and the residual energy in the stage in which the proton stops can be converted into WEPL via calibration. Experiments have been performed both on the research beamline of the proton synchrotron at LLUMC as well as the beamline of the IBA proton cyclotron of the Northwestern Medicine Chicago Proton Center (NMCPC).

At LLUMC, a beam spread across the scanner was obtained by scattering the

beam through a 2 mm-thick lead foil immediately upon exit from the beam pipe, about 3 meters upstream from the scanner. Since protons must be measured individually, and since there is no segmentation in the multistage scintillator to distinguish between two protons in the same RF bucket, the intensity of the beam was set to a very low value so that in one RF bucket (118 ns) the average number of protons is much less than one.

At NMCPC, the 230 MeV proton beam is degraded to 200 MeV near the source and a field of interest is defined by the beam operator. The beam is spread out to an area of roughly 4×4 cm^2 area by a degrader followed by momentum and aperture slits and beamline optics. Wobbler magnets scan the beam across the field of interest.

Protons enter the scanner through the particle telescope upstream from the phantom, pass through the phantom and exit through the telescope downstream from the phantom, then pass into the multistage scintillator through which the proton travels until it loses all of its energy. The residual energy is transformed via calibration into WEPL.

The phantom is mounted on a rotation stage in between the two particle telescopes. During a scan the rotation stage rotates either in 2 or 4 degree intervals, measuring several millions of events per projection angle, or continuously for a prescribed amount of time.

3.3 Tracking Detector

The tracking detector is based upon well-established SSD technology, which is also used by several other contemporary efforts in proton CT [51–53]. SSDs are nearly ideal candidates for proton CT, and their relatively high cost per square centime-

ter compared with plastic scintillators is offset by high performance, reliability, stability and ease of assembly. SSDs are nearly 100% efficient for charged particle detection with close to zero noise occupancy, have inherently fine spatial resolution (defined by strip pitch), and have simple calibration which is stable over periods of many years. In addition they are easy to assemble and lack any hazardous materials or properties which would be unfavorable in a clinical environment [54].

Two layers of SSDs must be superposed in order to obtain 2-D sensitivity. For vertical sensitivity (V-detector), the strips are horizontal, while for horizontal sensitivity (T-detector) the strips are vertical. Although the parameters, such as thickness (400 µm) and strip pitch (228 µm) of the SSDs were optimized for the Fermi-LAT, they work well in the proton CT application (see chapter 7 for a more detailed discussion).

In order to minimize the gap between the sensitive areas of adjacent SSDs, the SSD edges were resawed very close to the guard ring such that the gap between adjacent sensors is only about 0.6 mm. These gaps are staggered from one layer to another in order to reduce the frequency of protons passing through more than one gap. Although this technique increased leakage current, there was no impact on SSD noise. An alternative approach is to overlap, rather than trim the sensors, which was done in the Phase-I prototype, however this led to ring artifacts in the reconstruction due to incorrectly calibrated WEPLs, since protons passing through the overlap passed through a larger-than-typical WET. Since there is little redundancy in the tracking, one-hit per layer (8 hits total) is required for complete track reconstruction. Proton tracks passing through one gap can typically be recovered by a process described in section 3.7, however if a proton passes through more than one gap in the same (upstream or downstream) telescope, that track cannot, in general, be recovered.

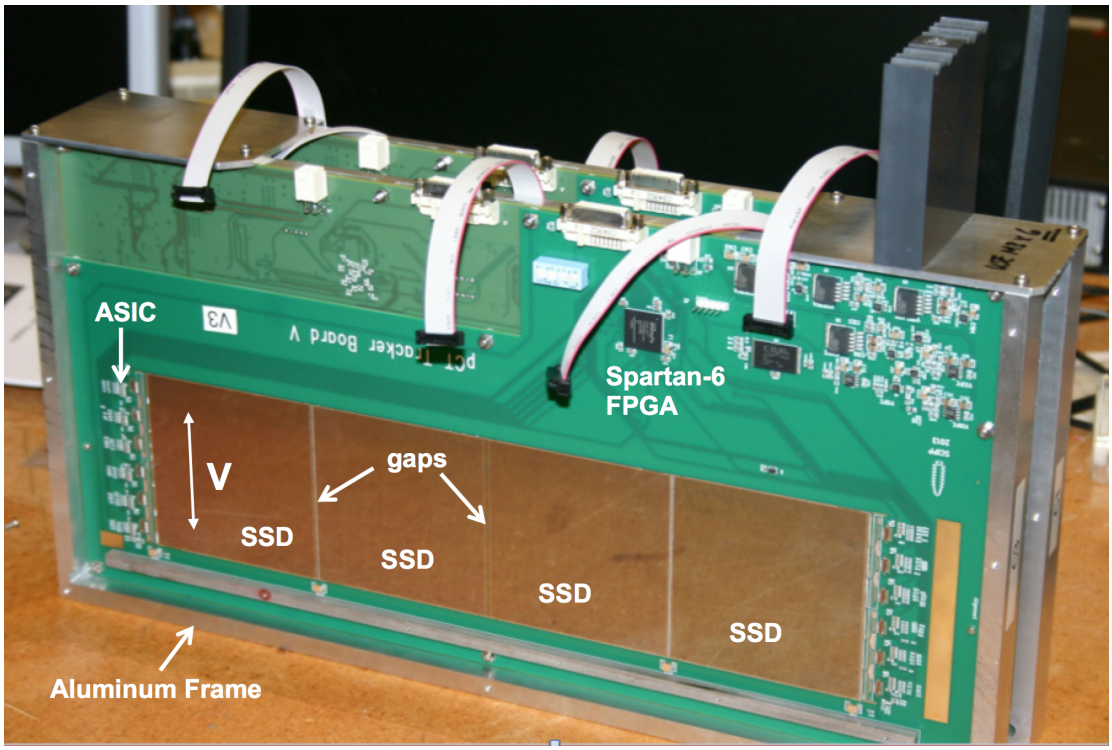


Figure 3.2: One of the two tracking modules with two V layers and two T layers, removed from its aluminum enclosure. A V tracking layer is visible. Strips on pairs of SSDs are wire bonded together and read out by six ASICs on either end of the board. Each V board has 12 ASICS and each T board has 24 ASICS. The loose cables visible are for programming the six Xilinx Spartan-6 FPGAs, one per V board and two per T board.

The SSDs are assembled on printed circuit boards in sets of four as shown in fig. 3.2. Each printed circuit board is cut out in the sensitive area of the detector and holds both a T-layer and V-layer to create a 2-D-sensitive tracking plane. The boards for the upstream and downstream plane in each particle telescope are aligned by pins 5 cm apart in the u direction, in an aluminum casing to better than 100 μm accuracy. This entire structure is contained in a light-tight aluminum cassette which contains "windows" at the sensitive area of the detectors. These windows are covered with 50 μm-thick blackened aluminum foil.

An analysis of 200 MeV data taken with no phantom installed at a data rate

of ∼1 MHz shows that the efficiency of each tracker layer is between 99.2% and 99.5%. When gap regions were excluded, the efficiency was found to be about 0.4% higher, indicating that most of the inefficiency can be explained by gaps in between the sensors. However, at very high data rates, with a fairly narrow beam, such as at the IBA beamline at NMCPC, efficiency is reduced due to pile-up of signals in the amplifiers. This effect can be minimized by spreading out the beam, so that the probability of a single strip being hit many consecutive times, before the previous signal attenuates, is reduced.

This efficiency is insensitive to threshold and timing settings, allowing the scanner to operate at successive beam tests scheduled weeks or months apart without any modifications to the hardware settings. This is because the electronic noise in the sensors is very low compared to the signal of a 200 MeV proton (at least 40 times smaller), and the rates of noise pulses above the typical ASIC threshold setting were on the order of one in one million per strip per trigger.

3.4 Tracker Readout

The front end of the data acquisition for the tracking detector is based on a custom integrated circuit (ASIC). A unique CMOS chip was designed for application in the proton CT scanner.

Data Acquisition Architecture

The custom ASIC is required to digitize the data and send it out rapidly while keeping front-end amplifiers active at all times. Each chip handles 64 channels, (one channel per strip) including logic for control, calibration, triggering, buffering and zero suppression, and outputs a formatted cluster list for each trigger to

corresponding field-programmable gate arrays (FPGAs). Twelve dozen ASICs are required for the system, corresponding to 6 ASICs per $\frac{1}{4}$-T-layer or per $\frac{1}{2}$-V-layer. Fourteen (twelve for the tracker + two for the energy detector) Spartan-6 FPGAs handle the data flow as well as control of twelve ASICs. Each tracker board is connected serially to the event builder, communication to which is made by low voltage differential signalling (LVDS).

The design of the ASIC is based on that of the Fermi-LAT with several application-specific changes, mostly related to speed, to accomodate the high data rate required for proton CT. Some important modifications include a digital 1-shot on each channel to define a short 150 ns trigger window followed by a FIFO for trigger latency, gain selection to optimize dynamic range for higher ionization of protons exiting versus entering the phantom, four parallel event processors that suppress data from all channels without signals and build cluster lists, and a dedicated output link, so that all chips can read in parallel instead of sequentially, as well as others. For a detailed description see Johnson et al. [55].

Data Flow

Figure 3.3 gives an overview of the contents of the ASIC. Each of 64 channels is associated with its own charge-sensitive amplifiers. If the discriminator determines a hit, the data is passed to a clock-synchronous edge detector which limits the signal to 150 ns (3 clock cycles). The pulse is sent on to the data masks, which can be programmed to mask out noisy or dead channels and then writes the data to a 64-wide by 32-deep hit buffer which is writing continuously. If a trigger acknowledge is received from the range detector, a "read" command is issued and the data is passed into one of four 64-bit event buffers, which bundle the data for readout logic processing. At this point the bitstream is reduced to a list of hit

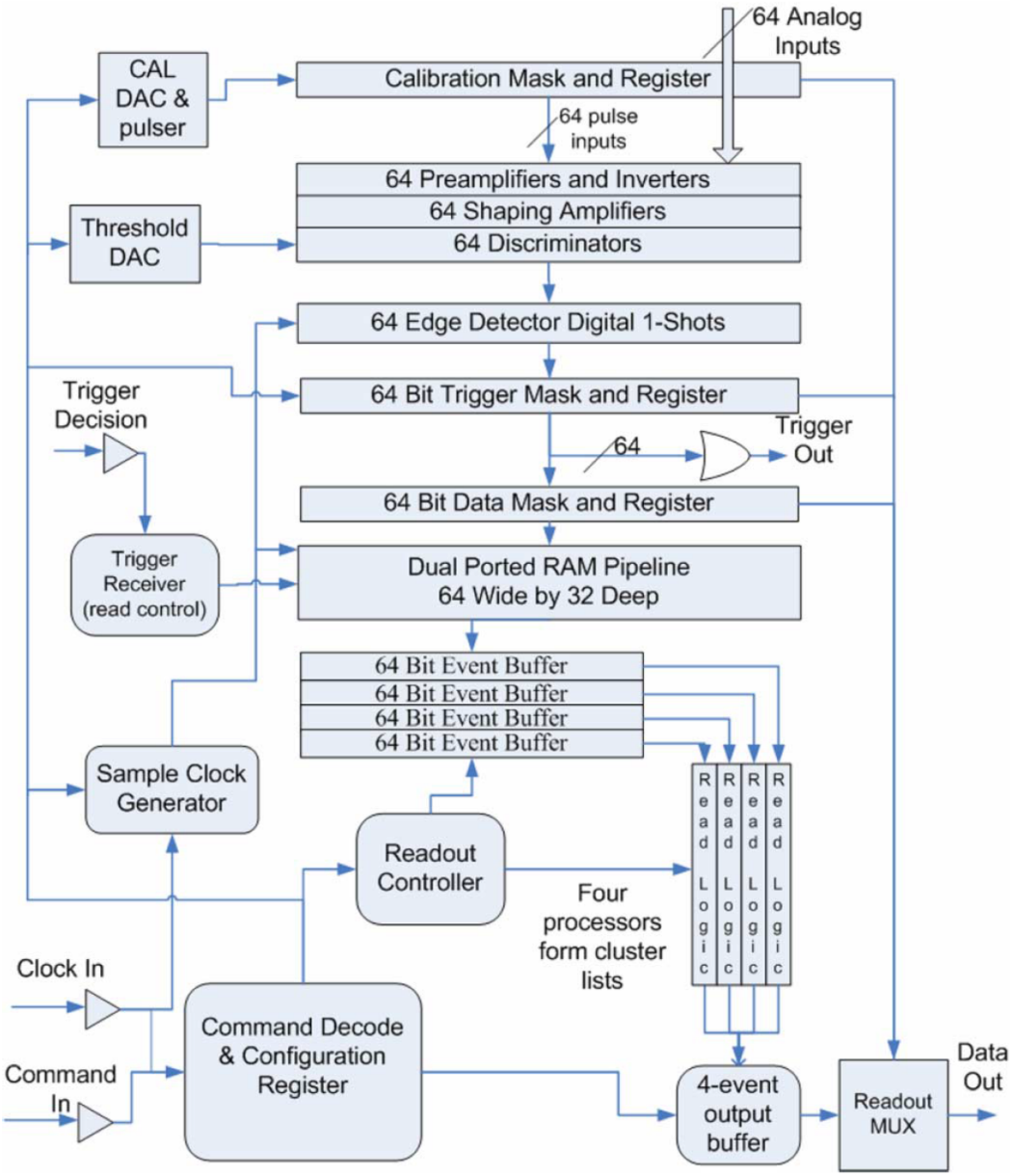


Figure 3.3: Simplified block diagram of the ASIC.

clusters, which are defined by the size of the cluster (the total number of strips in the cluster) and the address of the first strip in the cluster. The maximum theoretical number of clusters possible is thirty-two, if every other strip on the detector is hit, however, the chips are capable of buffering only ten clusters[1]. Again, four lists (corresponding to four events) can be buffered at a time.

3.5 Multistage Scintillator

In the current system, CsI crystals used in the Phase-I system were replaced with plastic scintillators for several reasons. The primary reason is that the signal decay rate in CsI is much too slow for the data rate required for proton CT and so we needed to choose a material with a faster response. Secondly, each of the 18 CsI crystals needed to be calibrated separately, since different gains in each crystal resulted in a non-uniform response and distinct image artifacts. Finally, in a calorimeter, the WEPL resolution of the detector is physically limited by range straggling, which is defined as the standard deviation of the Gaussian fit to the range distribution. This is approximately 1.1% of the proton's range. As a result, the uncertainty in WEPL increases with increasing WEPL which leads to image artifacts and degraded image quality. Furthermore, in order to approach the range straggling limit, the energy resolution of the detector needs to be better than 1%.

A multistage scintillator (MSS) combines properties of both a calorimeter and a range counter. It consists of a relatively small number of stages (we used $n = 5$) and therefore its response is defined by both the position of the stage in which the proton stops and the total energy that is deposited in the stage. Figure 3.4

[1]This is still far more than what is experimentally likely. Typically we expect one cluster of length one or possibly two if a proton passes between two strips

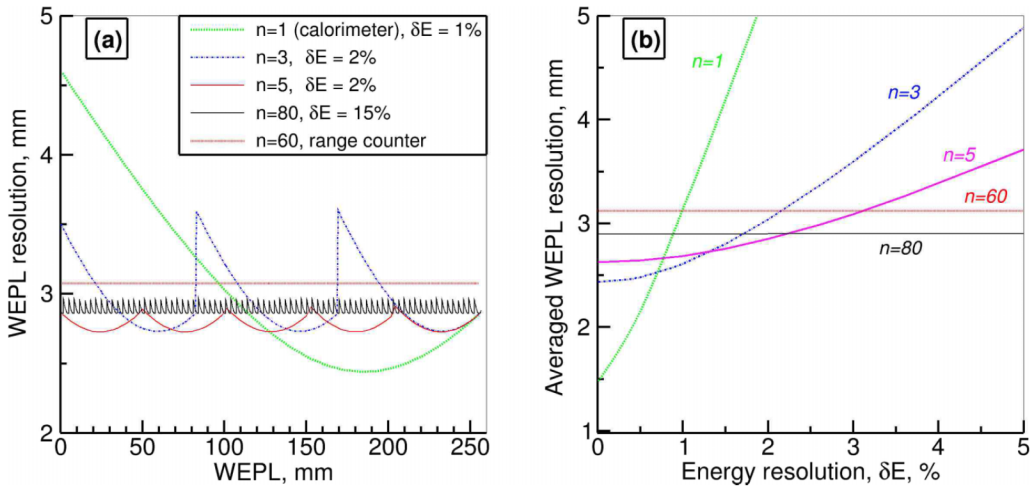


Figure 3.4: (a) Predicted WEPL resolution versus WEPL. (b) Predicted resolution averaged over 260 mm proton range as a function of intrinsic energy resolution δE for a calorimeter, a range counter and multistage scintillators with $n = 3, 5, 80$).

compares the theoretical WEPL resolutions of a calorimeter with 1% energy resolution, a range counter with 60 4 mm-thick stages and MSS's containing 3, 5 and 80 stages with 2%, 2% and 15% energy resolution respectively. The left-hand figure shows that with as few as 5 stages and 2% energy resolution, the MSS far outperforms the calorimeter, and also outperforms the range counter. The right hand plot shows the dependence of the average WEPL resolution on the energy resolution of the detector. This plot indicates that the average WEPL resolution of the 5-stage MSS is better than the range counter for energy resolutions better than 3%, which is reasonably achievable.

Design and Integration

The scintillating material used is polystyrene-based UPS-923A which provides high light output, low light attenuation and long-term stability [56]. Five stages covering a range of 260 mm requires that each stage be approximately 52 water-

equivalent millimeters. Since the relative stopping power of polystyrene is about 1.038, each stage should have a thickness of 50 mm. We chose 51 mm in order to fit a two-inch photomultiplier. The sensitive area of the detector is 10×40 cm^2 in order to overlap entirely with the sensitive area of the tracking detectors.

Each stage of the scintillator has one end beveled at an angle of 35° to form a built-in light guide that fits the two-inch square window of an R3318 Hamamatsu photomultiplier tube (PMT). All scintillator sides were optically polished and the PMTs were glued with optical epoxy to the scintillators as shown in the first panel in fig. 3.5. Each scintillator-PMT assembly was then wrapped with 65 µm-thick Vikuiti$^{\text{TM}}$ESR film with a greater than 98% reflectance (second panel). The five stages were then stacked together and the PMTs covered with a combined mu-metal/soft-steel shielding (third panel). The complete five-stage assembly is enclosed in a steel housing with a 10×37 cm^2 entrance window made of 50 μm-thick blackened aluminum foil (fourth panel).

MSS Readout

A custom printed circuit board was designed to digitize the PMT signals and package data for transmission to the event builder FPGA. Each energy detector board is capable of handling up to three channels. ADCs transform the pulses into 14-bit digital signals and can operate at up to 65 MHz. FPGAs on each energy detector board buffer the event data and, in parallel, reduce the 14-bit ADC values down to a summed pulse height per channel. Digital communication between the event builder FPGA and the digitizer boards takes place over a dual-link DVI-D cable per digitizer board.

Digitizers operate continually during the run, and samples are stored in a buffer, which is overwritten every 256 clock cycles. When a trigger is received,

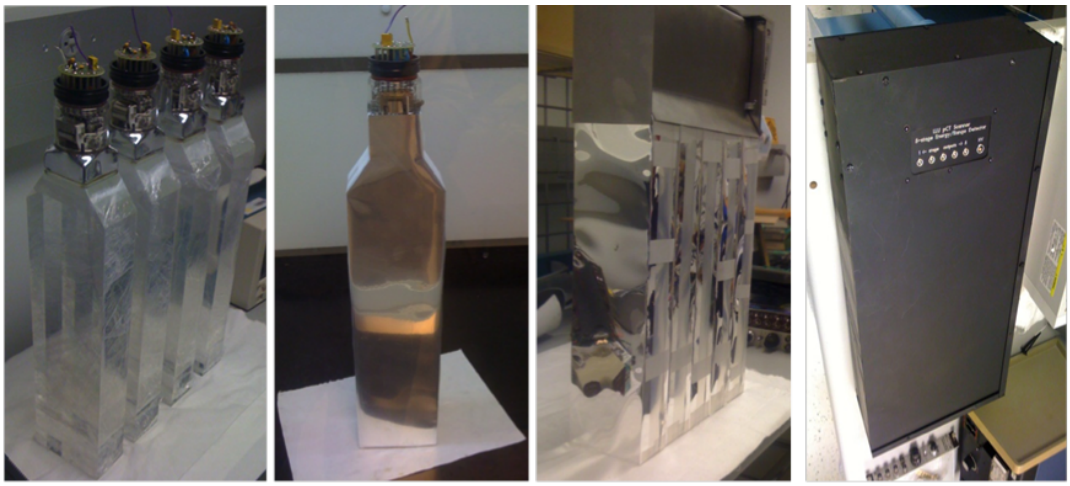


Figure 3.5: Assembly of the five stage scintillator. First panel: scintillators are beveled and outfitted with photomultiplier tubes. Second panel: The scintillators are individually wrapped in reflective film. Third panel: The scintillators are stacked together and covered with shielding. Fourth panel: The complete five stage scintillator was enclosed in steel housing with an entrance window.

the logic transfers up to 16 consecutive samples from the buffer and stores them in another buffer which is capable of storing up to eight events, until readout.

The performance of the digitizer was tested using signals from a pulse generator that were shaped to resemble PMT pulses. At 65 MHz, when the pulses arrived with a random phase relative to the ADC clock, the RMS width of the distribution of pulse sizes was 0.90%. When the pulses were phase-locked with the signal source, the RMS width was reduced to 0.15%, indicating that most of the variation was due to binning. Regardless, this variation is negligible compared with the natural variation in signals produced by protons, thus the digitization does not contribute significantly to range errors.

Data Reduction

Sending individual 15-bit samples (14 bits plus out-of-range indicator) to the event builder and the DAQ computer requires excessive data bandwith. While these samples are useful for testing purposes, they do not contribute new information to the overall data set, therefore, logic for data reduction was implemented in the digitizer-board FPGAs.

Because the data must be reduced in the FPGAs, the algorithm to estimate the pulse size in a single value must be simple and fast, therefore fitting the pulse to a functional form is impractical. Instead we use the discrete integral of the pulse as an estimator of pulse size, however, using all sixteen samples poses a danger of pulse distortion, caused by an overlap with signals from protons in adjacent RF buckets. For example, if the tail of a preceding signal has not entirely decayed before the rise of the next signal, the tail of the preceding signal will introduce a bias that increases the amplitude of the next signal. Therefore, a simple algorithm was developed in order to mitigate these effects.

The pedestals, which indicate the average signal in the electronics when no protons are present, are first subtracted off of the pulse. The pulse integral is then calculated by finding the peak sample of the pulse, and summing the peak with one sample preceding the peak, and three or four samples following the peak.

Calibration

It was found that the light collection by the PMTs varies significantly with the location of the proton path in the scintillator, relative to the PMT. Therefore, a "T-V calibration" must be implemented in order to make the detector response uniform. This was accomplished by extrapolating the tracks of protons exiting

the downstream tracker onto the scintillator being corrected, and accumulating the pulses in 5×5 mm^2 bins. The resulting distributions showed variations of up to 10% which is much greater than the <1% variation required for the desired range resolution. This required that we parameterize the positional dependence of the scintillator response. In order to do this we fitted a quadratic function in the T-direction and the V-direction and normalized the response to match the energy deposition predicted by a Monte Carlo model. Applying the correction reduced the spatial variation to an RMS of 0.4% over the sensitive area of the scintillator except for a small area near the PMT, which is not used in image reconstruction.

After the T-V correction is in place, the residual energy deposited in each of the five stages must be converted into values of WEPL for image reconstruction. The Phase-I system accomplished this using a series of runs with varying thickness degraders inserted into the device [40]; however, the number of runs required to calibrate the present system the same way would have required five times as much data (for five stages), and therefore, a custom polystyrene step phantom was designed and built to simplify the calibration procedure.

The step phantom (fig. 3.6) contains three pyramids along the T-direction with 6.35 mm steps, providing a stepwise variation of polystyrene thickness from 0 cm to 5.08 cm in the beam direction. To cover the full range of WEPL, four other 5.08 cm-thick removable bricks of polystyrene were also manufactured to be added successively to the variable part of the phantom. Therefore the maximum physical and water-equivalent thicknesses of the phantom is 254 mm and 263.7 mm respectively. Calibration is performed in five separate runs, the first with just the steps and the next four with an additional brick. The tracking information for each proton is then used in order to determine which protons passed through which steps, by projecting the tracks from the upstream tracker onto the phantom. The

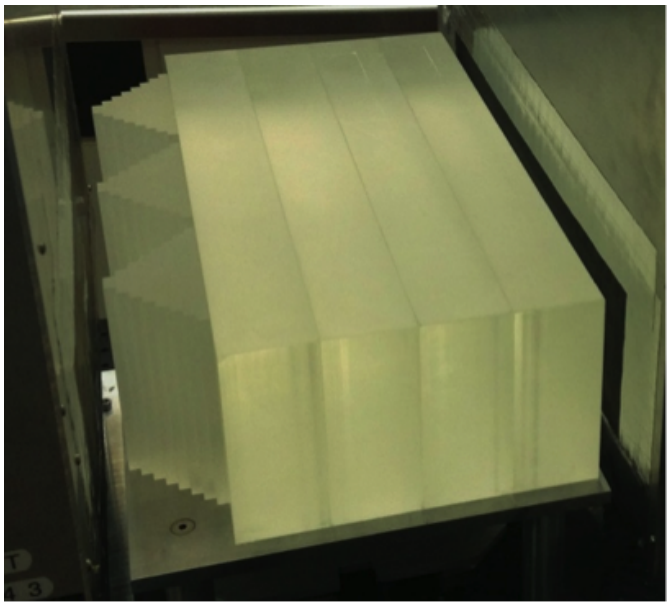


Figure 3.6: A custom step phantom composed polystyrene was designed for calibration of the five stage scintillator. The phantom contains three pyramids along the T direction with steps ranging in thickness between 0 cm and 5.08 cm in the beam direction. Four removable inserts, each of thickness 5.08 cm are used to obtain the full 26 cm range of 200 MeV protons in water.

WET traversed by each proton is calculated as the sum of air and polystyrene path lengths multiplied by their corresponding relative stopping powers. Finally, the corresponding detector response is recorded in terms of the five stage signals, corrected for track position dependence.

These data are used to generate calibration functions for each of the five stages, examples of which are shown in fig. 3.7. For each known WEPL value (given by the thickness of the step phantom), the distribution of energy responses are fitted to a Gaussian function, in order to find a stage's mean response and variance. The mean values are then fitted to a second order polynomial function

$$W = p_0 + p_1 E + p_2 E^2 \tag{3.1}$$

to produce the calibration curves. This technique was validated using Monte Carlo simulation. For a detailed description of the MSS and the calibration procedure

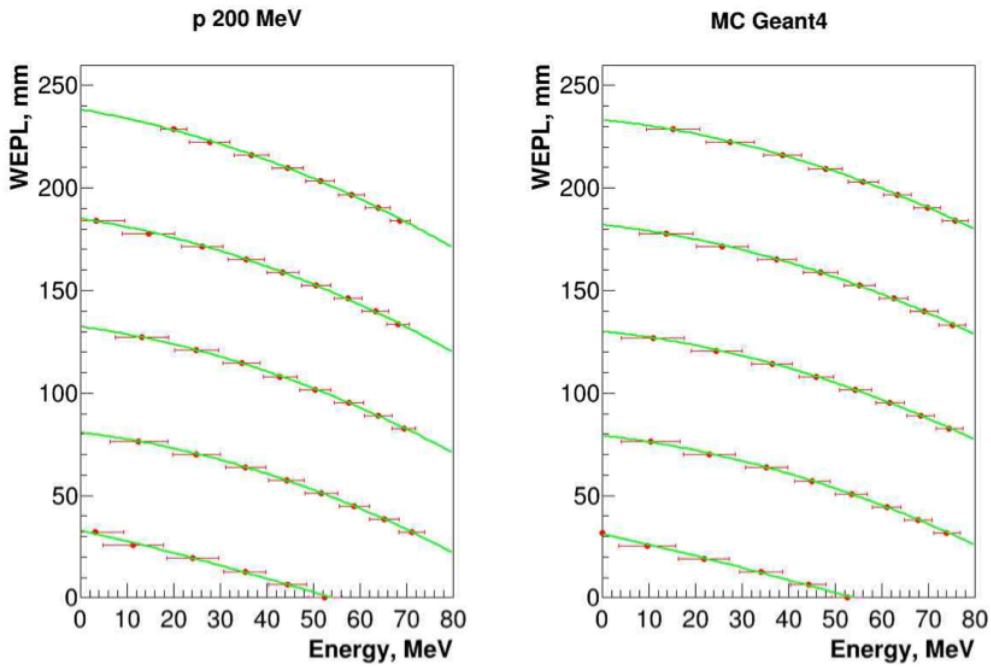


Figure 3.7: Experimental (left) and simulated (right) calibration plots for the five stages. The green curve is a second order polynomial fit used for the calibration function.

see Bashkirov et al. [57].

Detector Resolution

Using the established calibration functions, the resolution and linearity of the detector response was compared in both simulation and experiment. The results are shown in fig. 3.8.

Monte Carlo simulations were performed for a set of water slabs with thicknesses ranging from 0 to 250 mm in 1 mm increments. The curve in fig. 3.8 is shifted by about 12 mm because of proton energy loss in the 1.8 mm lead scattering foil on the LLUMC research beam line, which also contributes to a small loss in WEPL resolution. Both simulation and experiment yield an average WEPL

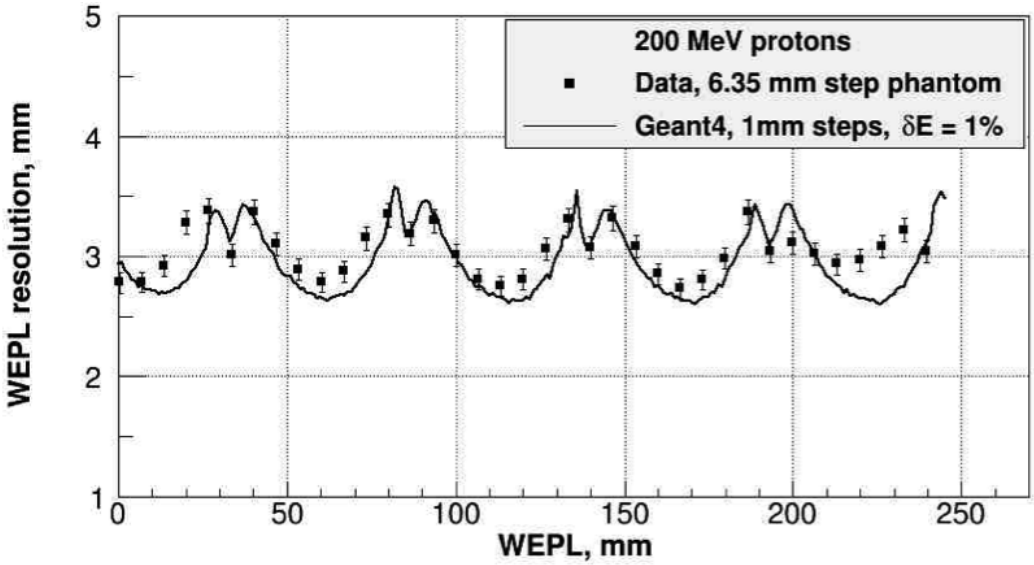


Figure 3.8: Simulated and experimental WEPL resolution for the five stage detector as a function of WEPL of the object being imaged. Points with error bars represent experimental measurements whereas the solid curve represent Monte Carlo simulation. The experimental data is shifted left by about 12 mm because of proton energy loss in the lead scattering foil on the LLUMC research beamline.

resolution of about 3 mm, regardless of degrader thickness. This is compared with a theoretical resolution of 2.8 mm obtained from the Bragg-Kleeman rule (left panel of fig. 3.4).

In addition to range straggling, uncertainty in proton range arises from leakage of energy due to nuclear interactions in the phantom and detectors leading to overestimated WEPL values, and long high-WEPL tails which pull up the mean value and broaden the distribution slightly. Nuclear interactions are mostly removed during CT reconstruction by appropriate data cuts.

3.6 Data Acquisition

The data acquisition system fig. 3.9 was designed to move raw data from at least a million events to the computer per second. The raw digitized data are delayed in

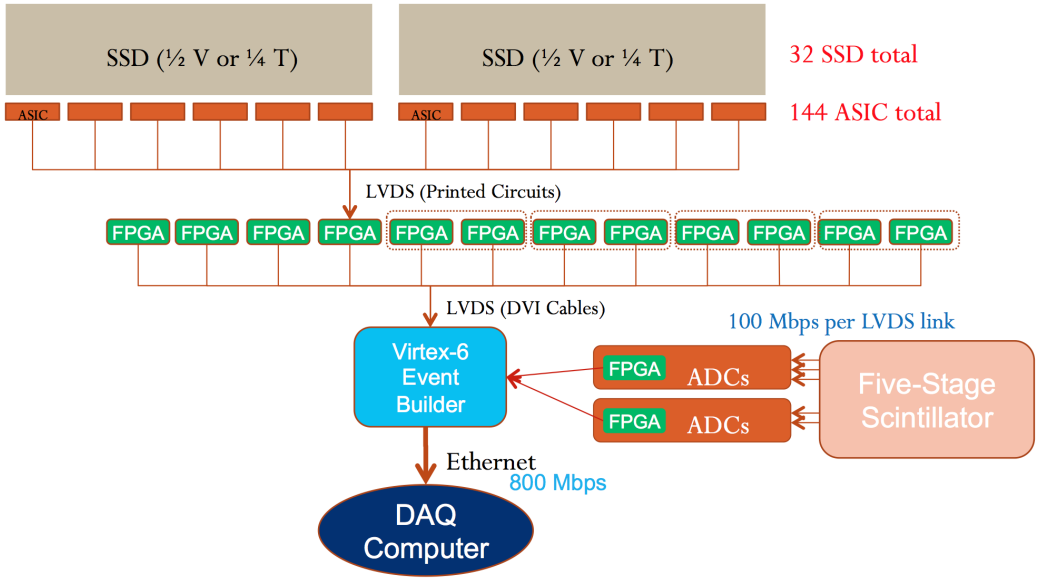


Figure 3.9: Diagram of the data acquisition flow. Each of the four tracker V boards has one FPGA and 12 ASICs. Each of the four tracker T boards has two FPGAs an 24 ASICs.

FIFOs at the front end in the ASICs for tracking boards and in FPGAs for energy detector boards until a trigger is received. Each FPGA receives lists of clusters of strips above threshold from each of the 12 ASICS with which it is associated. Each FPGA buffers the incoming data stream and builds a packet. Since in most cases there will only be one hit per layer, typically only one ASIC contributes a cluster. The FPGA then relays the compiled information to the event builder which receives information from all twelve FPGAs as well as from the two range detector FPGAs and compiles the data into a specified raw data format.

The fourteen front-end FPGAs send their data packets over DVI-D cables to the event builder, a Xilinx Virtex-6 FPGA on an ML605 evaluation board, together with a custom "mezzanine board." When all data from a given trigger have been received, they are packaged together and sent to the data acquisition computer via ethernet.

Data integrity is monitored by parity bits on ASIC-to-FPGA transmission and by an 11-bit CRC on FPGA-to-FPGA transmissions. In addition, each data packet is given a trigger tag at the front end which are matched by the system throughout event building. The data acquisition computer runs two processes during a run. One process receives data from the Ethernet data stream and writes the stream to a solid state disk. The second process controls the rotation stage during the run.

3.7 Preprocessing

After events are packaged by the event builder and sent off to the data acquisition computer, they still require a fair degree of processing before they can be used for reconstruction. Integrated pulses from the MSS must be transformed into WEPL values, and strip numbers need to be transformed into $(T\,, V)$ coordinates. This is all accomplished offline during a procedure called "preprocessing" which converts the raw bitstream into a specified binary format for reconstruction.

Track Reconstruction

A fully defined track in the upstream or downstream particle telescope is described by four hits, one hit in each of two T and V layers. These four hits together define a point and a vector in 3-D space. Tracks can be parameterized using an anchor point in the plane $u = 0$ by forward- or backward-projection, and two direction cosines. 2-D tracks are built for all combinations of hits in the upstream and downstream trackers, separately. A 2-D "supertrack" is formed by matching pairs of 2-D tracks in T and V, separately, in the upstream and downstream telescopes at the plane $u = 0$. Only supertrack pairs with a displacement of less than 10 mm

from one another at the $u = 0$ plane are considered. This cut corresponds to 12.5 standard deviations of the displacement distribution when no degrader is present, and 3.6 standard deviations in the presence of a 204 mm polystyrene degrader. Finally, a 3-D supertrack is reconstructed by matching 2-D T and V supertracks [58].

Gap Recovery

Some protons ($\approx 2\%$) pass through gaps in between sensors[2], or through dead (0.2%) or masked (noisy) strips. Most of these events (>80%) can typically be recovered. Since the tracker efficiency is so high (>99%), if a track is missing a hit, it is likely that the proton passed through a gap. If a track is missing a hit, the direction cosine is not defined because the track has an additional degree of freedom in either the T- or V-direction. However, when the location of gaps are known well, it becomes simple to infer which gap was traversed when three of the track coordinates are known. If the missing coordinate is in the front tracker, the proton path through air can be approximated as a straight line, forward-projected from the beam source, since MCS in air is very small. The gap closest to the intersection of this estimated track with the tracker layer is likely the gap that was traversed. In the rear tracker, we cannot use the beam source for track recovery, due to an unknown amount of MCS in the phantom. In this case we utilize the previously described track-matching algorithm (section 3.7), which forward- and backward-projects the tracks from the upstream and downstream telescopes, respectively, onto the plane $u = 0$. In this way, track-matching can play the role of the beam source for gap recovery in the downstream telescope.

[2]Assuming a flat distribution the probability that a proton will pass through a gap is $p = \Delta t_{\text{gaps}}/T = 2/354 \approx 0.6\%$. Then the probability of having one missing hit in 4 layers is $1 - (1 - p)^4 \approx 2\%$

Data Reduction

Six sequential data reductions are made on the data during preprocessing, in order to eliminate events that lacked information, or failed certain requirements. These data reductions sequentially remove all events that:

1. do not have at least one hit in each T-V layer pair
2. have too many hits in any given layer
3. do not have at least one 2-D track
4. fail the d<10 mm requirement ("displacement test")
5. do not have at least one T-supertrack and one V-supertrack
6. have more than one possible supertrack.

Supertracks cannot be built for events that are missing a T- and V-coordinate in the same U-plane, nor can they be made if there are more than three hits in any given layer – in this case, the ambiguities cannot be removed. Events that fail the displacement test most likely underwent nuclear scattering inside the phantom or degrader, and therefore do not provide useful information for reconstruction. Finally, ambiguities cannot be resolved for events that have more than one supertrack. These are most likely multiple-proton events (2 or more) which cannot be resolved due to a lack of segmentation in the range detector. For this reason, all multi-proton events must be eliminated. These events can often be identified by the amount of energy deposited in the range detector. Finally, information from the range detector is used in order to eliminate any events that leave an uncharacteristic energy trace. Roughly 30% of raw events are eliminated by these data reductions.

3.8 Chapter Summary

A prototype proton CT scanner has been built and successfully tested in two beam lines. Hardware and data acquisition have operated reliably at the design data rate in both synchrotron and isochronous-cyclotron facilities. The novel multi-stage scintillator combines elements of calorimetric measurement with proton range measurements and we have shown that the range resolution achieved in both simulation and experiment is close to the theoretical limit. The system is now ready to support testing of reconstruction algorithms as well as a thorough evaluation of proton CT in terms of image resolution, RSP measurement, spatial resolution, and dose deposition in a variety of phantoms.

Chapter 4

Geant4 Simulation Platform

4.1 Introduction

Monte Carlo (MC) simulations are a useful tool to study the performance of detectors in many applications in particle physics and medical physics. The Geant4 Monte Carlo toolkit is a platform for the simulation of the passage of particles through matter and is a successor of the GEANT series of software toolkits developed by CERN. In this section we will introduce a modular Geant4-based simulation platform developed by Giacometti et al. [59] that is a virtual representation of the prototype proton CT scanner mounted on existing beamlines at LLUMC and NMCPC.

4.2 Platform Description

Geometry

The MC simulation of the prototype proton CT scanner described in chapter 3 was implemented in Geant4 version 10.1. Figure 4.1 shows the schematic geometry

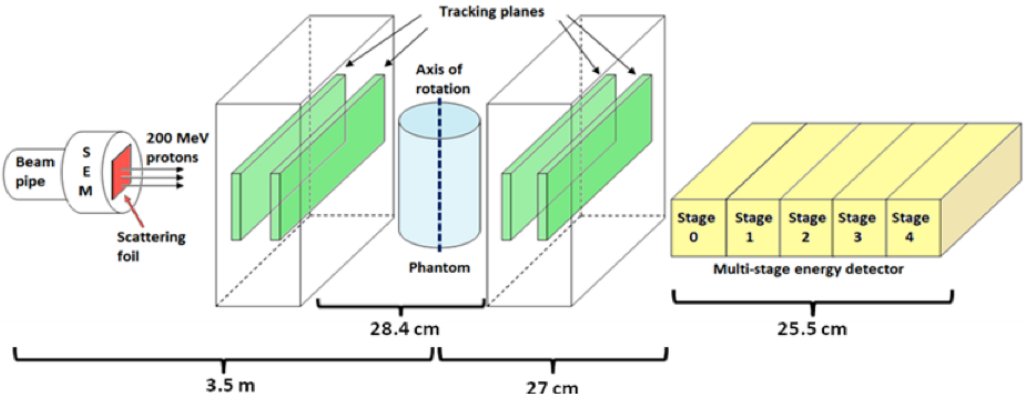


Figure 4.1: Schematic representation of the Geant4 model of the prototype proton CT scanner installed on the research beamline at LLUMC.

of the head scanner simulated in Geant4. The simulation described in this thesis is modeled on the research beamline at LLUMC. At the time of this thesis, the majority of our modeling has been done using the LLUMC beam line for simplicity and historical reasons, however it is important to mention that this the beamline module of the simulation can and will be replaced with alternate beam lines of known geometry or phase space.

In the case of the LLUMC beamline, a 200 ± 0.005 MeV proton pencil beam with diameter 0.2 mm and without angular divergence is modeled inside a vacuum enclosed stainless steel pipe with a length of 5 cm, an inner diameter of 3.52 cm and a wall thickness of 2.9 mm. The proton beam passes through five 12.7 µm-thick aluminum foils which represent the secondary-electron emission monitor (SEM) detectors that are integrated at the distal end of the vacuum pipe. The beam then exits through a 25 µm-thick titanium foil. At the exit from the beam pipe, the beam passes through a 1.9 mm-thick lead foil which scatters the beam into a Gaussian cone beam of approximately 16 cm (FWHM) at the entrance of the front tracker.

The silicon trackers are modeled realistically as described in section 3.3. The

silicon strips are modeled as sensitive Si volumes of identical size and spacing, and are grouped according to chip, SSD, and FPGA number used by the DAQ system (section 3.6). The phantom is placed at the scanner isocenter between the trackers. After passing through the tracker, protons typically stop in the multi-stage scintillator (MSS) located 27 cm downstream from the isocenter. The MSS has been realistically implemented as described in section 3.5.

Physics Packages

The Low Energy Package, based on the Livermore data libraries [60], was selected to model electromagnetic interactions of protons in the proton CT scanner. The threshold of the production of secondary particles was optimized to speed up the simulation without compromising the accuracy of its results. The G4HadronPhysicsQGSP_BIC_HP and the G4HadronElasticPhysicsHP were chosen to describe neutron interactions up to 20 MeV. Ion hadronic interactions were described by G4IonBinaryCascadePhysics.

Output

The output of the simulation platform consists of the position of the proton intersection with the tracking planes and the energy deposited in every stage of the MSS. The user can select from two output formats for the hit positions:

- Hit coordinates (in millimeters) with respect to the origin of the isocenter. This simulates the data format required for reconstruction.
- Strip, chip, SSD, and FPGA numbers of the strips hit by the protons in each tracking plane. This simulates the prototype data bit stream.

The energies of the protons are converted to WEPL using the calibration procedure described below. The user may also specify the number of projections and the projection interval, or simulate a continuous scan.

4.3 Calibration

The goal of the WEPL calibration is to correlate the signal generated by protons in the stages of the MSS to the known WET that the proton traversed in a calibration phantom of known composition. The calibration phantom described in section 3.5 was implemented in Geant4. The calibration procedure was simulated in the proton CT scanner platform and the calibration curves produced were compared with the curves obtained with the experimental calibration.

To reproduce the experimental calibration procedure, five calibration runs, each containing approximately one million protons were simulated, one with the step phantom alone, and four more with additional degraders after the step phantom. Only protons entering a step near the middle were used in the analysis – any proton entering less than 0.35 mm from the edge of the step was excluded. In addition, protons that scattered out of one step and into another were excluded. The setup for the calibration of the simulation is shown in fig. 4.2.

As in experiment, a Gaussian fit to the central part of the histograms of the scintillator responses was applied to define the mean energy deposited in every stage for each step. Polynomial curves of energy versus WET were fitted to those mean energy values. Ambiguities in calibration for protons that stopped within 2 mm of an interface between two stages were resolved by taking the transmission curves into account in addition to the stopping curves.

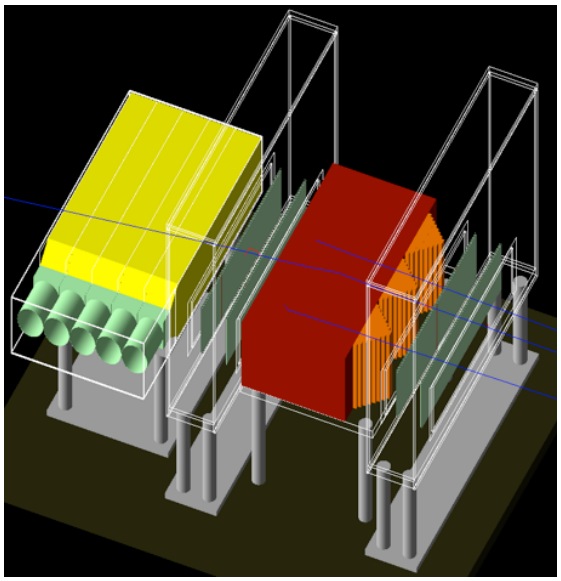


Figure 4.2: The simulated calibration setup with the calibration phantom (orange) and four additional polystyrene degraders (red) installed. The MSS (yellow) is shown with photomultiplier tubes (green).

4.4 Validation

The validation of the Geant4-based simulation platform required comparing each of the following in simulation and experiment: the tracker and MSS responses, the calibration curves, and the reconstructed RSP values from scans of a variety of phantoms.

The tracker responses were validated by comparing the strip responses in each of the eight tracker layers in experiment to those in the simulation. This included a comparison of the measured and simulated horizontal and vertical beam profiles in the tracking planes, respectively, generated by the proton beamline and scattering foil.

The MSS was validated by comparing the WEPL distributions of protons passing through polystyrene degraders of thicknesses 50.8 mm, 101.6 mm and 203.2 mm with the experimental WEPL distributions. The central parts of the distributions

Table 4.1: Comparison of experimental and simulated WEPL using degraders of varying thicknesses.

Physical Thickness (mm)	WET (mm)	Experimental WEPL (mm)	Simulated WEPL (mm)	% Diff.
50.80	52.73	52.36 ± 3.38	52.8 ± 3.48	0.84%
101.60	105.46	104.80 ± 3.31	105.5 ± 3.42	0.66%
203.20	203.92	211.12 ± 3.17	211.1 ± 3.39	0.009%

were fitted with Gaussian curves. The mean WEPL values calculated for experimental and simulated data were compared with the expected WET of the degraders. The agreement between measured and expected WEPL is less than 1% in both experiment and simulation (table 4.1).

The validation of the performance of the simulated proton CT scanner was accomplished using the sensitometry and line pair modules of the Catphan 600 series (The Phantom Laboratory, Salem, New York, USA) and a pediatric anthropomorphic head phantom (model HN715, CIRS, Norfolk, Virginia, USA). These phantoms will be described in greater detail in chapter 6. The reconstructed RSPs of the materials contained in sensitometry module and the head phantom were compared in images reconstructed both from simulation and experiment and have been shown to agree to within 1%. Good agreement in spatial resolution between experiment and simulation has been found by comparing the discrete modulation transfer functions obtained using the line pair module. A detailed account of the validation process is described in [59].

4.5 Chapter Summary

We have introduced a modular Geant4-based simulation platform that is a virtual representation of the prototype proton CT scanner which has been validated by

comparing its output to that of experiment. MC simulations provide a useful and essential means to study the performance of our system in a controlled setting. The simulation described in this chapter, as well as modified versions of it, have been used throughout the course of this thesis to develop theories and test the hypotheses of proton CT.

Chapter 5

First Studies in Proton Radiography

5.1 Introduction

With increasing use of proton radiation therapy for cancer patients, research into new imaging methods that can improve the accuracy of proton range estimates in radiation therapy planning have become a high priority. In many cases, protons are particularly desirable for treating cancerous tissue in close proximity to radiosensitive normal tissues, such as at the base of the skull and near the spinal cord. Protons are preferable to photons because their energies are easily tuned and the high dose of the Bragg peak can be localized reducing the threat of damaging otherwise healthy tissue. Accurate treatment of tumors at the base of the skull and near the spinal cord requires accurate knowledge of proton stopping and scattering power.

In order to obtain relative stopping power (RSP), Hounsfield units (HU, i.e. units of x-ray attenuation used in x-ray CT) are transformed using a calibration

curve. However, there is no unique relationship between HU and RSP, especially in the regime of RSP=1 (i.e. water, human tissue). This means that errors in proton range are consistently at least 2% of the nominal proton range or even higher in regions containing tissue, bone and air interfaces. A recent survey by the American Association of Physicists in Medicine (AAPM) showed that 33% of attendees polled said that range uncertainties are the main obstacle to making proton therapy mainstream [3]. Simulations and first experimental results have shown that by using a proton CT imaging system one may be able to reduce this range uncertainty to about 1% or less without increasing the dose to the patient [4].

Proton radiography has been recognized as being valuable for patient-specific optimization of the calibration from CT HU to RSP, and researchers at the Paul Scherrer Institute (PSI) have pioneered its technical development and use [24, 61]. Proton radiography differs in several key aspects from x-ray radiography. While unscattered x-rays travel in straight line paths, protons undergo multiple Coulomb scattering (MCS) events. This limits spatial resolution since the proton path deviates from the assumed straight lines by up to several millimeters in anatomical objects encountered in medical proton imaging. The accuracy of those path estimates is critical for achieving a high spatial resolution in proton radiography. An advantage of proton radiography is that it may allow us to distinguish small differences in RSP of various tissues better, compared with x-ray radiography at the same dose. Proton radiography also provides us, directly, with the water equivalent thickness (WET) of the object being imaged, which is useful for quality assurance in pre-treatment verification in proton therapy.

In this chapter, we report an experiment which used 200 MeV protons to generate proton energy-loss and scattering radiographs of a hand phantom. The

experiment used the phase-I proton CT scanner prototype[1], which was installed on the research beam line of the clinical proton synchrotron at Loma Linda University Medical Center (LLUMC). It was found that while both radiographs displayed anatomical details of the hand phantom, the energy-loss radiograph had a noticeably higher resolution. Nonetheless, scattering radiography may yield more contrast between soft and bone tissue than energy-loss radiography; however, this requires further study. This study contributes to the optimization of the performance of the next-generation of clinical proton CT scanners. Furthermore, it demonstrates the potential of proton imaging (proton radiography and CT), which is now within reach of becoming available as a new, potentially low-dose medical imaging modality.

5.2 Materials and Methods

Phase I prototype proton CT scanner

The first generation prototype proton CT scanner, shown in fig. 5.1, is based on the design principle described in [62]. It consists of two principal components: The silicon tracker, which tracks individual protons, and the segmented cesium iodide calorimeter used for measuring residual energy of protons.

Eight silicon tracker tiles are arranged into four planes, each 400 µm thick, with a sensitive area of 8.95 cm × 17.4 cm. The silicon strip detectors (SSDs) have a strip pitch of 228 µm. The tiles are arranged into two "telescopes," one upstream and one downstream from the phantom, so that the entry and exit vectors can be determined. The tracking planes interface through a high speed data acquisition

[1]This study was performed before the completion of the current, Phase-II prototype which is the primary focus of this thesis

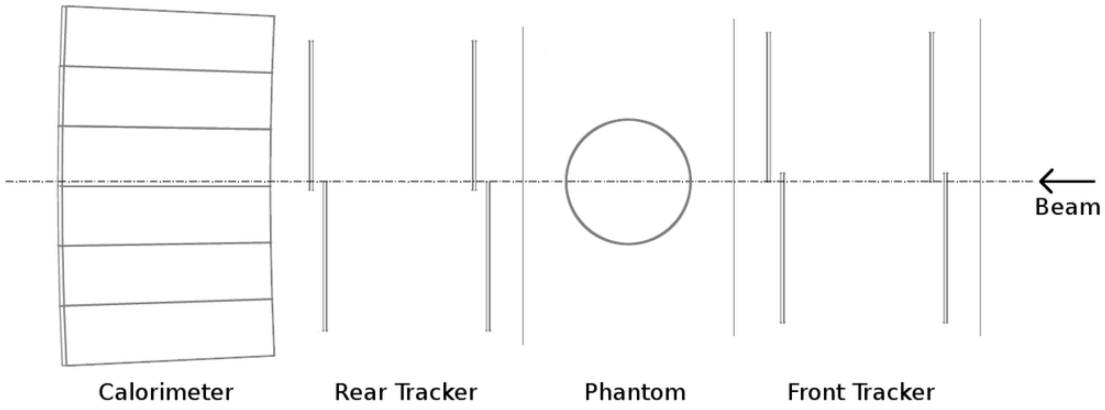


Figure 5.1: Schematic of the phase-I prototype proton CT scanner design. The distance between the two planes of the front and rear telescopes is 9.8 cm. The distance between the front and rear tracker is variable.

system based on field programmable gate arrays (FPGAs).

Each tracking detector has a fast trigger output which is connected to the trigger logic within the FPGA. A trigger decision is then made each clock cycle (50 ns) and the readout sequence begins for the entire system when a coincident event is detected between any two detectors forming an x-y pair. Offline, further data rejection is performed to select events with higher order coincidences only (8-fold, in this case).

The calorimeter is composed of an array of 18 thallium-doped CsI crystals arranged in a 3×6 matrix. Each crystal is 12.5 cm long, which is sufficient to stop 200 MeV protons. The residual energy of the protons is converted to light by scintillation, which is read out by photodiodes and converted to a digital value by an analog-digital converter (ADC) at a rate of up to 100 kHz.

For the radiography experiment, the proton flux in the accelerator at LLUMC was tuned to very low values (less than 50k protons per 0.5-second spill) to match the rate capability of the first-generation data acquisition system.

A realistic hand phantom, consisting of a human hand skeleton embedded in tissue-equivalent plastic, was mounted vertically between the upstream and downstream SSDs and imaged with a low-intensity 200 MeV proton cone beam, emanating from the vacuum exit window of one of the proton research beam lines at LLUMC. The entry and exit vectors, as well as the calorimeter response of each individual proton, were measured. This information allowed us to reconstruct the estimated trajectories and water equivalent path length (WEPL) of each proton through the phantom. The WEPL values were derived from the calorimeter response using a calibration process for the first-generation proton CT scanner, which is described in detail in Hurley et al. [40].

Radiography algorithm

For the production of the radiographic images, software developed for a full proton CT reconstruction described in chapter 6 was executed on a workstation equipped with two dual-core central processing units and 8 GB of RAM. The data input into the radiography software contained the proton tracker coordinates and the calorimeter response for each proton. A cylindrical reconstruction volume, containing the entire hand phantom, was created. The main steps of the algorithm used for the proton radiograph reconstruction are as follows:

1. Read the input data (proton histories) from file.

2. Determine which histories traversed the reconstruction volume and calculate the locations where the proton entered and exited the volume.

3. Assign angular and spatial bins to each history.

4. For histories that traverse the reconstruction volume and belong to the zero-degree angular bin, write data to file.

5. Write scattering angle values in the t and v directions (ϕ_t, ϕ_v) to files corresponding to each pixel.

6. Calculate the total scattering angles, defined with respect to the beam axis, for each proton in each pixel.

7. Plot the median total scattering angle in each pixel to obtain the scattering radiograph.

8. Convert residual energy values to WEPL using the calibration curve.

9. Perform data "cuts": For each pixel, define a "mode window" of WEPL that accepts protons within $\pm$ 30% of the mode, or $\pm$1 cm if 30% is less than 1 cm. Define the "peak" WEPL as the mean of the distribution contained within the mode window.

10. Plot peak WEPL value to obtain the energy-loss radiograph.

The radiographic reconstruction made use of the phantom-based coordinate system $(x\,, y\,, z)$ and the detector-based coordinate system $(t\,, u\,, v)$ of the proton CT scanner (fig. 2.1). The entry and exit vectors were projected onto the reconstruction volume so that the entry and exit coordinates could be determined. These points were connected by a straight line, for the purposes of radiography, which estimated the path of the proton. The midpoint of the estimated proton path was calculated as the average of the entry and exit coordinates. The proton was then binned into the nearest integer spatial bin (pixel) in t and v.

Cuts in relative angle, defined as the difference between entry and exit angle, were made at 3σ from each pixel's mean relative scattering angle in both the t

and v directions. These cuts were made to exclude events that had very large scattering angles, caused by inelastic nuclear interactions or elastic large-angle scattering events inside the phantom. The software also made cuts in WEPL at 3σ from the mean pixel value, and were necessary to exclude histories that were affected by inelastic nuclear interactions or resulted from unresolved coincidence of two or more particles in the calorimeter.

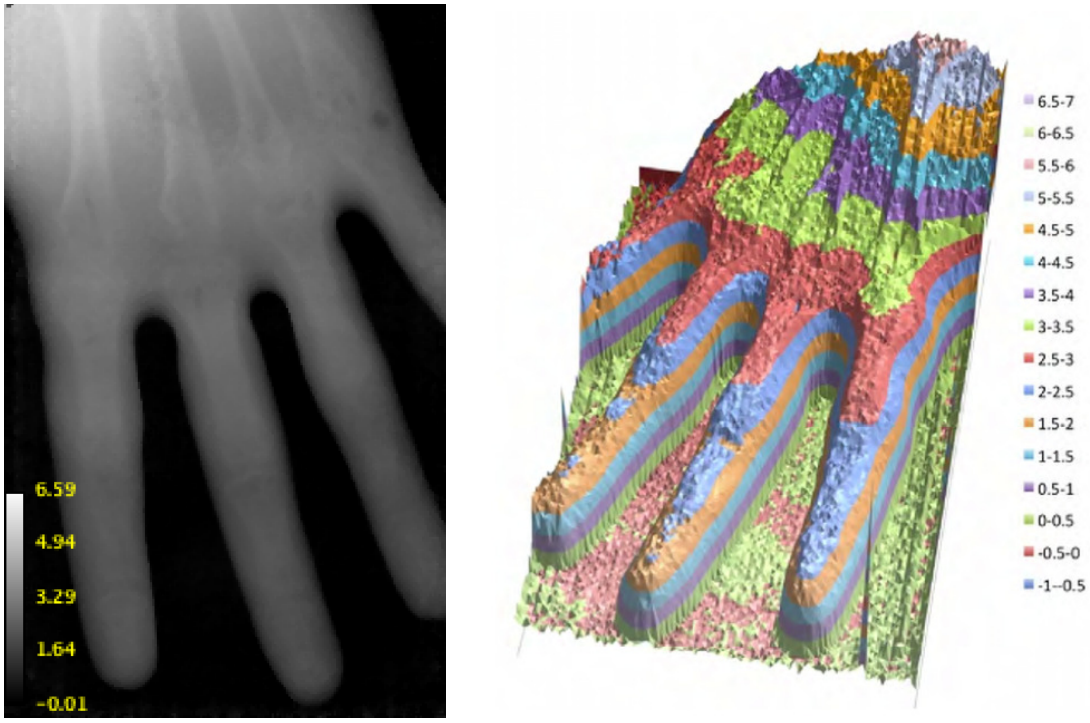
The reconstruction of the radiograph was performed in the t-v plane located at $u = 0$, at the center of the reconstruction volume.

5.3 Results and Discussion

Proton energy-loss radiography

Figures 5.2a and 5.2b show two expressions of radiographic images of the hand phantom in terms of WET. A total of about 3.5 million proton events entered the reconstruction volume and were detected by the SSDs with near 100% efficiency. About 2.6 million of these events were selected by excluding pileup events, which were recognized by their larger-than-usual energies deposited in the calorimeter, and by removing events that did not pass through all four tracker planes. Of the 2.6 million events, 67.1% passed cuts in angle and WEPL and were used to construct figs. 5.2a and 5.2b. A pixel size of 0.5×0.5 mm^2 was used to create these images; therefore, each pixel contained the data of approximately 40 protons.

The contrast in 5.2a is low compared with an x-ray radiograph. This is due to the fact that the RSP of bone is only 50%-80% greater than that of water while the x-ray absorption power of bone can be several times greater, depending on the photon energy. The relief-map display of the phantom in Fig. 5.2b more clearly depicts regions of varying WET of the hand, and shows clear structural details.

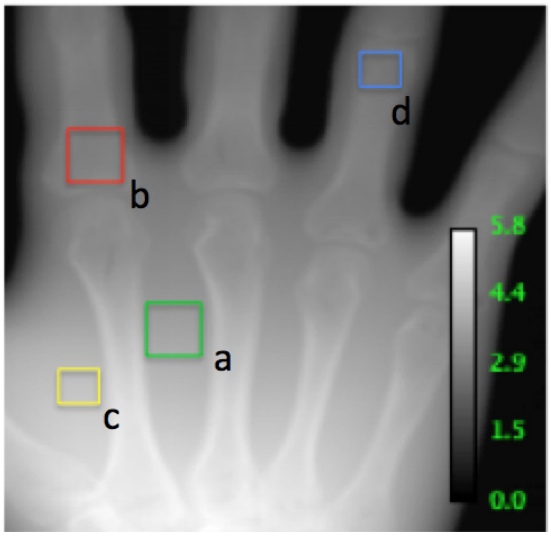


(a) WET Radiograph (b) WET relief map

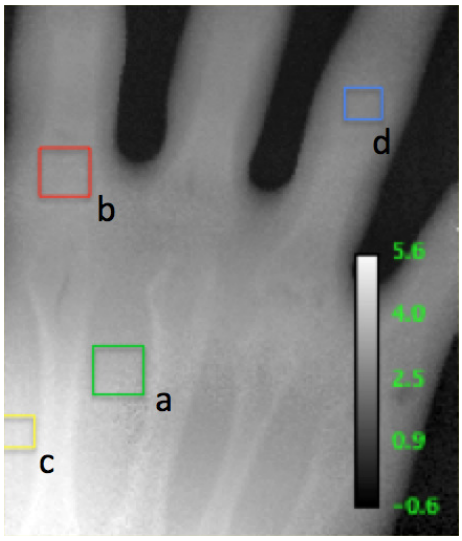
Figure 5.2: The first radiograph of a hand phantom with 0.5 mm pixels (left) and a relief map of WET calculated from the summed-up stopping power of the phantom (right). The scales are in centimeters of WET.

In order to generate a comparable WET image from x-ray data, an x-ray CT scan of the hand phantom was performed with the 64-slice GE LightSpeed VCT scanner (GE Healthcare Technologies, Waukesha, WI), normally used for proton treatment planning in the Department of Radiation Medicine at LLUMC. A Python script was used to convert voxel values to RSP with the Hounsfield-Unit-to-RSP calibration curve provided by the medical physics team. An x-z projection of RSP values was created by summing over pixels with constant x-value in the y-direction for each slice of constant z, and multiplying the sum by the pixel size.

Figure 5.3 contains a side-by-side comparison of the transformed x-ray CT scan and the proton radiograph with the same window and level settings. The



(a) X-ray radiograph



(b) Proton radiograph

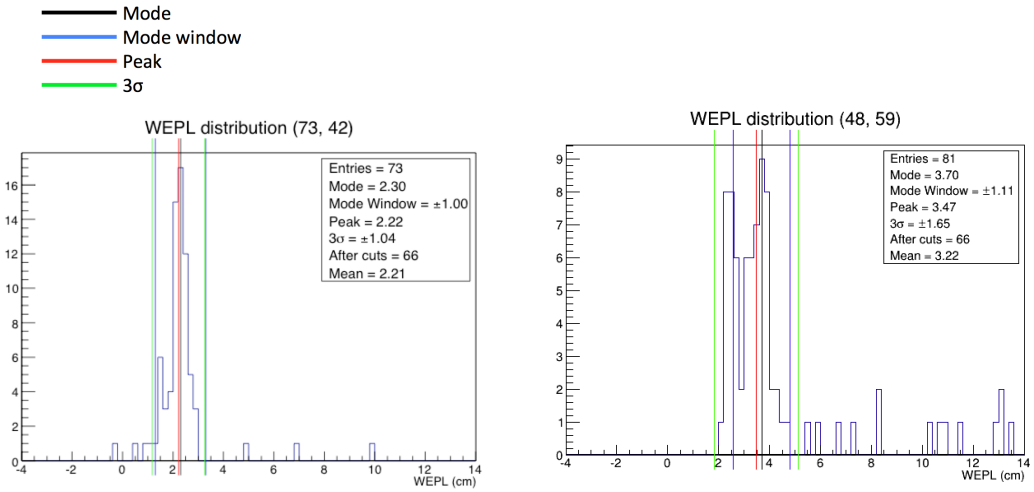
Figure 5.3: (a) The WET projection image of the hand phantom based on a high-resolution x-ray CT scan. (b) The corresponding WET proton radiograph. The scales are in centimeters of WET.

Table 5.1: Comparison of $\mathrm{WET}_{\mathrm{xray}}$ and $\mathrm{WET}_{\mathrm{proton}}$ for Selected ROIs

ROI	$\mathrm{WET}_{\mathrm{xray}}$ (cm)	$\mathrm{WET}_{\mathrm{proton}}$ (cm)	% Diff.	Relative Diff. (σ)
(a.)	3.6 ± 0.1	3.5 ± 0.1	2.8%	0.7
(b.)	2.9 ± 0.1	3.0 ± 0.1	3.3%	0.7
(c.)	4.2 ± 0.1	4.2 ± 0.2	0.%	0.0
(d.)	2.5 ± 0.1	2.5 ± 0.0	0.%	0.0

calibration scales are in centimeters of WET. Table 5.1 shows that the WET values averaged over selected regions agree to within about 3%. This is within the realm of what is expected.

The WEPL distribution of protons before data reduction in each pixel was roughly Gaussian, as seen in fig. 5.4a. The distribution is usually right-skewed (high WEPL) which corresponds to the left-skewed (low-energy) distributions in energy. The protons in the tails are protons that underwent nuclear scattering



(a) (b)

Figure 5.4: Distribution in WEPL for pixels described by the coordinates (a) $(t = 73, v = 42)$ and (b) $(t = 48, v = 59)$ before cuts are made. The black line defines the mode of the distribution and the red line defines the mean or "peak" of the distribution. The blue lines indicate the mode window which contains the particles within ±30% of the mode, and provides the distribution on which the 3σ cuts are based. The green lines indicate the cuts made on this specific pixel. Notice the straggling in the large WEPL range. These values correspond to particles that underwent nuclear interactions. (a) illustrates an example of a roughly Gaussian WEPL distribution. (b) is that for a boundary pixel with a bimodal WEPL distribution.

events. These are the events that are eliminated using the appropriate cuts.

We did find, however, that a significant percentage of pixels contained non-Gaussian WEPL distributions. These distributions, as in fig. 5.4b, are bimodal and correspond to pixels that lie on the boundary between two materials of different RSP. In this case, the reconstruction algorithm selected the leftmost mode, and the appropriate cuts were determined based on that value. This, however, reduced spatial resolution. Methods such as averaging the two modes, or "splitting" pixels have been proposed and have yet to be explored.

The image quality of proton radiography in comparison to x-ray radiography would best be studied as described in the work by Depauw and Seco [63]. These

investigators used the contrast-noise-ratio (CNR) expressed in units of decibel (dB) in a Monte Carlo study of x-ray and proton radiography with a high-contrast resolution phantom. However, for a similar analysis, the hand phantom used in this study was too non-uniform to provide meaningful data. A phantom better suited for a CNR analysis will be constructed and used in future experiments.

Proton scattering radiography

The amount that a proton is scattered by MCS between its entry and exit from a phantom is inversely proportional to the product of its velocity and momentum. In the Gaussian approximation of MCS, given by Lynch and Dahl [27], the width of the scattering distribution is described by:

$$\theta = \frac{13.6\text{MeV}}{\beta cp} z \sqrt{\frac{x}{X_o}} [1 + 0.038 \log \frac{x}{X_o}] \tag{5.1}$$

where θ is the width of the Gaussian approximation for angular deflection in a plane, β, p are the velocity and momentum of the proton, respectively, z is the charge of the proton and x/X_o is the thickness of the material traversed in units of radiation length. We calculate X_o of the material using:

$$\frac{1}{X_o} = \sum_i \frac{w_j}{X_j} \tag{5.2}$$

where the w_j's are the fractions by weight of each element in a given material and the X_j's are the radiation lengths of each material. The second term in eq. (5.1) tends to be small and can thus be ignored for purposes of estimation. Note that this approximation is good only for relatively thin objects (i.e. $10^{-3} < x/X_o < 100$) where the velocity and momentum are assumed to be approximately constant. For a thicker phantom, we would need to account for energy-loss by introducing an integral over x (see [31] for details).

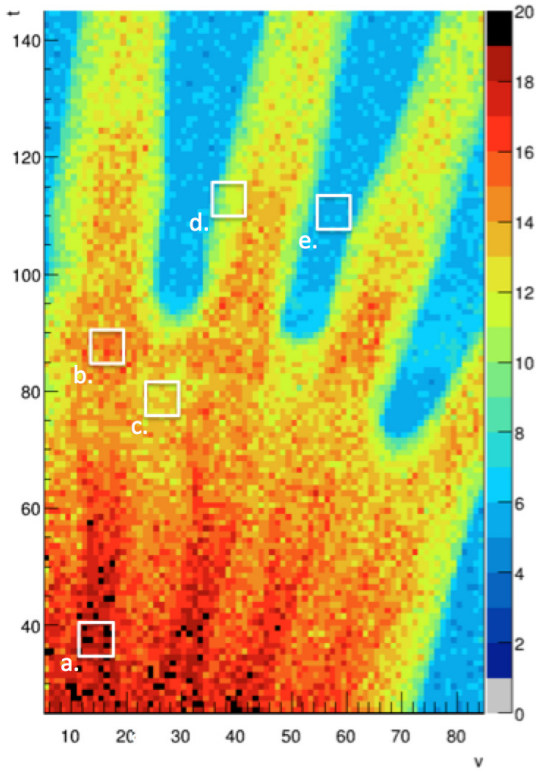


Figure 5.5: Scattering radiograph of the hand phantom showing the median 3-dimensional scattering angle (in milliradians) in each voxel. The blue color code corresponds to the scattering of about 5 mrad of 200 MeV protons in the SSDs.

Table 5.2: Expected and Observed Scattering Angles for 200 MeV Protons ($\beta = .566$, $p = 644$ MeV/c) for Selected ROIs (units in milliradians)

ROI	Expected	Observed (fig. 5.5)	Relative Diff. (σ)
(a.)	17.	$17.8. \pm 0.9$	0.9
(b.)	15.	$14.1. \pm 0.9$	1.0
(c.)	11.	12.3 ± 0.7	1.0
(d.)	10.	$10.3. \pm 1.3$	0.2
(e.)	5.	5.2 ± 0.4	0.5

Table 5.3: Densities and radiation lengths of materials commonly encountered in proton CT. Data for bone: [26]; for tissue, water and silicon: [27]

Material	Density (g/cm^3)	Radiation Length (g/cm^2)
bone	1.45	16.6
tissue	1.00	38.2
water	1.00	36.1
silicon	2.33	21.8

The scattering radiograph of the hand phantom in fig. 5.5 was obtained as follows: for each pixel, the scattering angle of each proton history in the t-u and v-u planes was obtained from measurement of the entry and exit directions. The total scattering angle with respect to the u-axis was then calculated as

$$\Phi = \sqrt{\phi_t^2 + \phi_v^2} \tag{5.3}$$

for every t, v pair. The median Φ in each pixel was plotted as a color map on a two-dimensional graph.

Table 5.2 displays the expected and observed median scattering angles for select regions of interest (ROIs) corresponding to fig. 5.5. The expected values were calculated using eq. (5.1), the measurements of bone and tissue thicknesses obtained from the x-ray CT scan of the phantom, and the radiation lengths of standard bone, soft tissue, water and silicon (table 5.3).

Table 5.2 shows good agreement between expected and observed scattering angles for the hand phantom. All of the selected scattering regions fall within 1σ of the expected value. With the exception of region (b), the measured scattering values were consistently slightly higher than the expected values, which may be explained by the right skew of the distribution in Φ.

Despite the higher noise in the scattering radiograph compared to the energy-loss WET radiograph (fig. 5.2a), it shows a great amount of structural detail.

Skeletal structures such as metacarpal bones and soft tissue can be clearly distinguished. It should also be noted that the overall scattering value of regions (a) and (b) are similar, despite the difference in the structure of the phantom, because region (a) contains a large amount of tissue and a small amount of bone, while region (b) contains a large amount of bone and a very small amount of tissue.

Figure 5.6 compares two image profiles for the energy-loss radiograph (dashed curve) and the scattering radiograph, (solid curve) for a pixel size of 1 mm. When the profile of the scattering curve was normalized to the energy-loss profile, we found that the general shapes of the two curves of each plot were very similar, which shows that in this case, regions of greater stopping power were also regions of higher scattering power, as one would expect. Furthermore, qualitative observation of many of these types of line profiles suggest that scattering radiography may yield a higher CNR; however, a phantom better suited for this measurement would be required to confirm this. In this case, proton scattering radiography may prove beneficial for patient alignment based on skeletal features.

While the energy-loss radiograph provides us with WET for energy loss which is most important for quality control in proton range estimation, the scattering radiograph may provide us with the WET for proton scattering power. The usefulness of this information for treatment planning requires further investigation.

Dose to Phantom

Three-and-a-half million proton events entered the reconstruction volume in this experiment. The fluence near the central axis of the proton cone beam was approximately 24,800 cm^{-2}. Assuming a proton weighting factor of unity and a total stopping power of 4.5 MeV/cm for 200 MeV protons, the estimated dose delivered to the center of the phantom was approximately 20 μSv which is within the same

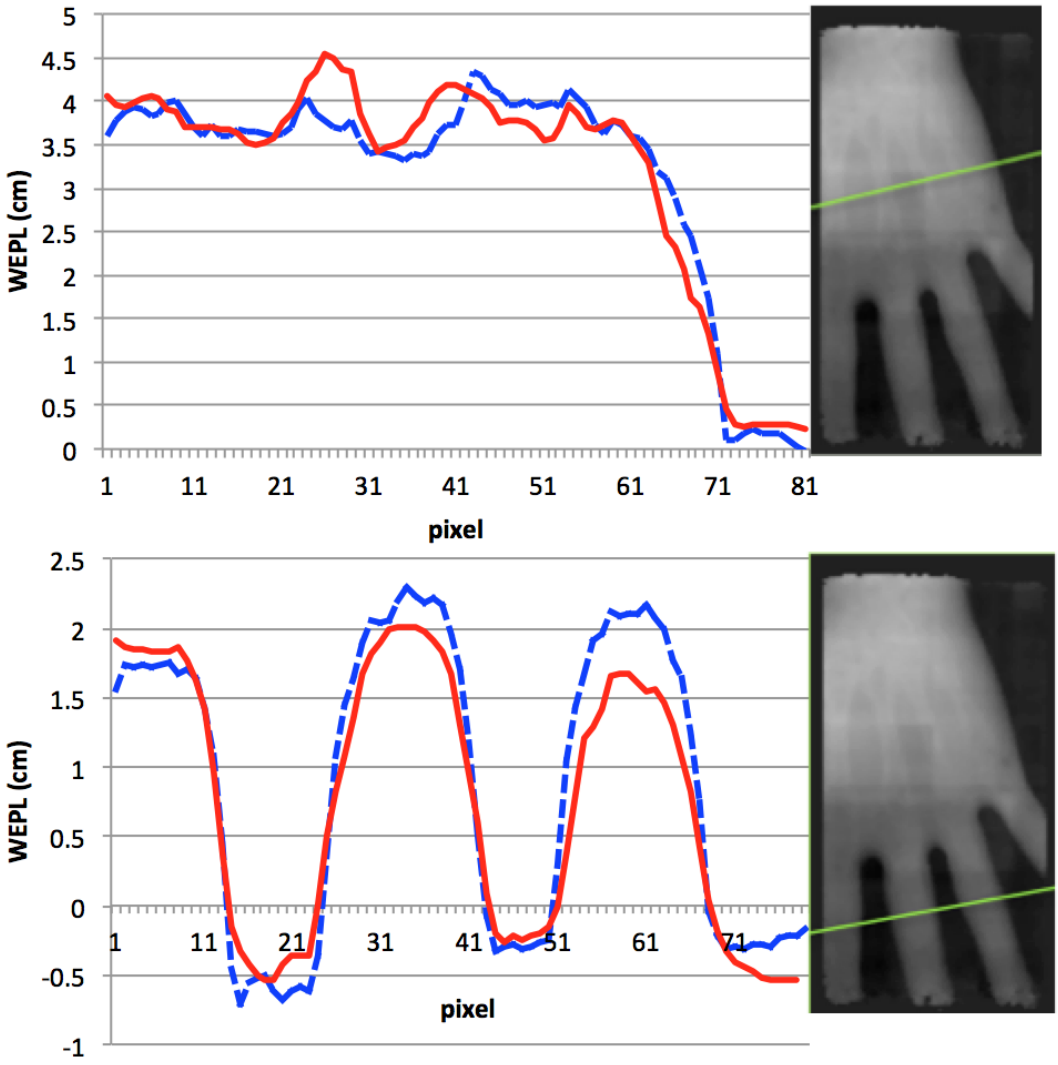


Figure 5.6: Normalizing the scattering radiograph (solid curve) to the energy-loss radiograph (dashed curve), we see roughly the same shape and even some subtle features for radiographs with 1-mm pixels. While the quality of the scattering curve is not as good as that of the energy-loss curve, it may still be useful for patient positioning.

order of magnitude as the dose delivered in state-of-the art x-ray radiography for a hand or foot [64]. There was no intention to formally study the relationship between dose and image quality in this work. Based on theoretical calculations [31] and early experimental proton CT experience, [20] the dose needed for the same level of image noise could be much lower with proton radiography than was used in this study.

5.4 Chapter Summary

Our proton radiographs demonstrate the promise of proton imaging (proton radiography and CT) now within reach of becoming a new, potentially low-dose clinical imaging modality. For future investigations, new phantoms will be developed and built in order to systematically study these properties. Region of interest studies of the scattering radiograph indicate strong agreement between our measurements and the predictions made using Lynch and Dahl's equation. Although the image is lower quality than the energy-loss radiograph, scattering radiography provides information about the proton scattering power of materials, which is inversely proportional to the effective atomic number distribution in the tissue. Energy-loss radiography cannot provide this information since stopping power depends only on Z/A, which is practically identical for most soft tissues and water, leading to very low contrast. Therefore, scattering radiography and tomography will likely have useful applications in proton treatment planning.

Presently, our radiography procedure assumes straight-line trajectories of protons through the phantom. It is likely that we can improve our images by incorporating the most likely path of the proton into our binning algorithm. Additionally, our reconstruction algorithm presently makes angular cuts in the u-v

and u-t planes, separately. We believe that by adding these angles in quadrature and cutting based on this total angle (as in fig. 5.5) will improve the quality of our event selection. Better yet, instead of removing large-angle scattered protons from our analysis by cuts based on Gaussian distributions, we would rather like to include these data by using more appropriate scattering models. In this case, it may be possible to obtain additional information about the elemental composition of tissues using proton radiography and tomography.

Chapter 6

Proton CT Reconstruction

6.1 Introduction

Proton beams have an advantage compared with other types of radiation treatment since they have the potential to deliver radiation energy to the tumor while leaving the surrounding tissue mostly undamaged. This is possible because of the properties of the depth-dose curve, which is characterized by a low-dose leading edge and the Bragg peak near the distal edge. In order to realize the full potential of this property, however, the RSP of the proton must be accurately known in the patient. Without very precise knowledge of RSP, significant margins around the primary target must be included in order to insure irradiation of the entire target.

The goal of proton CT is to obtain a three-dimensional map of proton stopping power within a patient. Using a proton beam for CT imaging would make the proton radiation treatment more precise by defining the Bragg peak position in the patient more accurately, and by enabling the verification of the actual patient and tumor position with respect to the proton beam.

In this chapter I present several proton CT images of phantoms reconstructed

from data obtained at the Northwestern Medicine Chicago Proton Center (NMCPC) in May 2015, and I compare the reconstructed RSP values to their expected values. We find strong agreement between the reconstructions and their truth values.

6.2 Materials and Methods

Phantoms

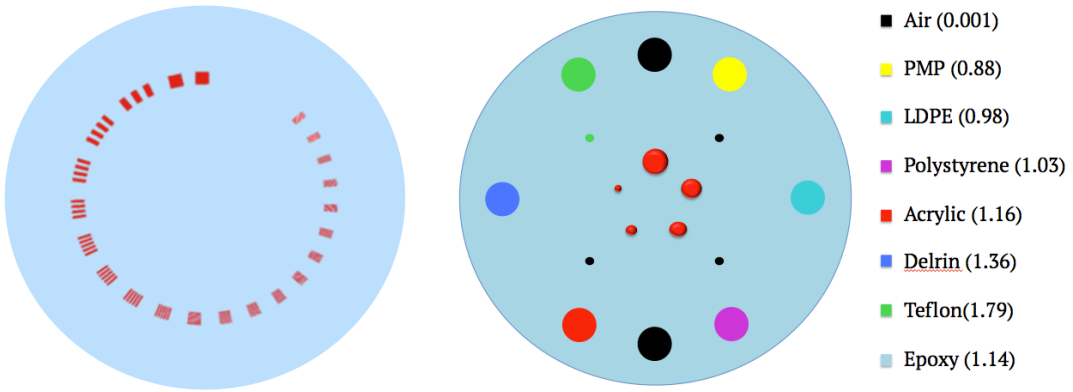
Four phantoms were scanned for this experiment. These phantoms include:

1. Water Phantom: A PMMA cylinder with 5 mm-thick walls and a 75 mm outer radius filled with distilled water, which was degassed in a vacuum chamber.

2. CTP528 "Line Pair" Phantom: A PMMA cylinder with a diameter of 150 mm and a height of 40 mm containing aluminum inserts that are 2 mm tall, 5 mm wide and of varying width and spacing. These inserts are arranged into line pair patterns ranging in spatial frequency from 1 to 21 line pairs per centimeter (lp/cm) at a radius of 48 mm (fig. 6.1a).

3. CTP404 "Sensitom": An epoxy cylinder with a diameter of 150 mm and a height of 20 mm containing eight 12.2 mm-diameter inserts composed of LDPE, Polystyrene, PMMA, Delrin, Teflon and PMP (fig. 6.1b).

4. Pediatric Head Phantom (model HN715 CIRS): a realistic phantom reproducing anatomical details of the head and cervical spine of a 5-year-old male, including skeletal and soft tissue features including the tongue, intra-cranial and paranasal sinuses filled with a lung-tissue equivalent material, ear canals,



(a) CTP528 (b) CTP404

Figure 6.1: (a) Schematic of the CTP528 Line Pair module. (b) Schematic of the CTP404 sensitometry module with color-coded inserts. The materials contained in the phantom and their RSPs are shown.

and realistic facial features. It is composed of eight tissue-equivalent materials including brain-, pediatric cortical bone-, trabecular bone-, enamel-, dentin-, spinal disc-, and soft tissue-equivalent polymers. A photograph of the phantom and a 1.25 mm slice of a digitized model of the phantom are given in fig. 6.2.

Each phantom was mounted on the rotation stage of the prototype proton CT scanner on a 200 MeV beam at NMCPC as shown in fig. 6.3 for the pediatric head phantom. Projections were measured in 4 degree increments from 0 to 360 degrees.

Data Preparation

The raw data were preprocessed as described in section 3.7. Roughly 65% of the recorded events passed the preprocessing stage. After preprocessing the data for the water phantom contained about 200 million histories, data for the sensitom

(a) Pediatric head phantom (b) Slice of digitized pediatric head phantom

Figure 6.2: Photograph of the anatomically accurate head phantom of a 5-year-old male child (left) and a 1.25 mm slice of the digitized model of the head phantom (right). The tissue-equivalent materials contained in the phantom and their RSPs are color coded in the digital model.

Figure 6.3: The anatomical pediatric head phantom installed on the beamline at NMCPC.

contained roughly 340 million histories, and the data for the Line Pair phantom contained roughly 700 million histories. At this point, the data set for the Line Pair phantom was further reduced by randomly selecting events from each projection such that the resulting data set consisted of 340 million events. This was done in order to keep the time required for reconstruction at a reasonable length, and because 340 million events is more than sufficient for a phantom of this size. For the head phantom, two scans are required, one for each the upper and lower halves of the phantom. For each of these scans roughly 340 million events remained after preprocessing.

The data for all phantoms except for the Line Pair phantom were reconstructed using 0.625 mm pixels and a slice thickness of 2.5 mm. These specifications were selected in order to insure that each bin contained high enough statistics so that the images would not be corrupted by noise. The slice thickness was sacrificed in favor of a smaller pixel size for the sake of spatial resolution in the resulting images.

Reconstruction Software

The proton CT reconstruction software used to produce the images in this thesis was executed on a workstation equipped with two dual core central processing units, 8 GB of RAM and an EVGA GeForce GTX680 GPU. The data input to the reconstruction software contained proton tracker coordinates, the WEPL for each proton and a run header that specified the projection angle, along with other descriptive information.

Image reconstruction requires the definition of a reconstruction volume (a cylindrical volume of user-specified diameter) that contains the entirety of the object being imaged. For each proton, the coordinates of entry into and exit from

the reconstruction volume are determined by forward and backward projection of the entry and exit vectors measured by the upstream and downstream silicon trackers, respectively, of each proton onto the reconstruction volume. Protons that do not enter the reconstruction volume, as well as those that exit from the top or bottom surfaces of the cylinder, are ignored during image reconstruction.

After the intersections with the reconstruction volume are calculated, a straight line between the point of entry to and exit from the reconstruction volume is used to approximate the proton path through the reconstruction volume. Protons are then binned at the plane u=0, based on where the midpoint of the proton's path through the reconstruction volume is located. Proton histories with WEPL values greater than 3 standard deviations from the "peak" WEPL value are eliminated[1]. Three-sigma data cuts are also made on relative angle ($\theta_{\text{rel}} = \theta_{\text{exit}} - \theta_{\text{entry}}$) in order to eliminate particles with very large scattering angles[2]. These cuts are effective at reducing errors from protons that undergo hadronic interactions in the scanner or phantom.

After cuts have been made, the binned data is arranged into a sinogram, which is then passed through a Shepp-Logan filter and is used as input to the filtered back projection (FBP). The Feldkamp-Davis-Kress (FDK) algorithm, the cone-beam version of the FBP, is used both for boundary detection and as a starting point for iterative reconstruction. Boundary detection is achieved by thresholding the FBP image. Any voxel assigned an RSP of 0.6 or larger is included within the object. This information is then passed to the MLP algorithm to define the volume in which multiple Coulomb scattering occurs.

[1]The "peak" is defined by first finding the left-most mode of the WEPL distribution in the bin. From that mode, a mode window is defined to contain +/- 30% of the data from the mode. The mean of this reduced distribution is defined as the peak, and the standard deviation of the reduced distribution is the one used in the 3σ cuts.

[2]Angle cuts may not be necessary. WEPL cuts appear to eliminate all protons scattered at significantly large angles.

For iterative reconstruction, a method based on diagonally relaxed orthogonal projections (DROP) onto convex sets is used. This method requires knowledge of WEPL and the MLP of each proton, forming a large linear system of equations, which is then solved for the unknown RSP image vector. The DROP method has been further enhanced by interleaved superiorization of the total variation in the reconstructed RSP map [65].

DROP-TVS

Many problems in mathematics and real-world applications can be modeled as convex feasibility problems. These are problems in which one seeks a point in a convex set that is contained in the region where all hyperplanes intersect, where a convex set is defined as one in which all the points on a straight line between any two points of the set are also contained in the set[3]. A feasibility-seeking method is fundamentally different from an optimality-seeking method. In the feasibility-seeking method, consistency with all prior information defines a *set* of equally acceptable solutions, while an optimality-seeking method seeks a single "best" solution.

In proton CT, we prefer the feasibility-seeking heuristic for several reasons. Firstly, the solution vector is described by the intersection of hyperplanes in an N-dimensional space, where N is very large (order $\sim 10^8$) and the hyperplanes contain inherently noisy data. Given the presence of this noise, a unique solution typically does not exist, rather the hyperplanes intersect in a region of space, all of which must be explored to find a preferred solution. Secondly, because the image will ultimately be judged by the human eye which is insensitive to mathematical optimality, the definition of a mathematically best solution is largely

[3]All linear (equality or inequality) or affine constraints lead to convex sets.

inconsequential.

Projection algorithms are iterative algorithms that use projections onto (usually closed and convex) sets. Projection algorithms are a popular choice for solving linear systems of equations because of their conceptual simplicity. Orthogonal projections are a particular type of projection algorithm that uses the nearest Euclidean distance between two hyperplanes.

Diagonally-Relaxed Orthogonal Projections (DROP) is derived from the simplest orthogonal projections algorithm, Algebraic Reconstruction Technique (ART) which is a fully sequential algorithm (eq. (2.47)) and therefore very slow when N is very large. DROP accounts for the sparsity of the $\boldsymbol{A}$ matrix (section 2.6) which speeds up its convergence compared with sequential ART. In proton CT we use the block-iterative version of DROP, which further increases the speed of convergence [66].

The DROP algorithm can be expressed:

$$x^{k+1} = x^k + \lambda_k \boldsymbol{U}_{t(k)} \sum_{i \in I_{t(k)}} \frac{b_i - \langle a_i, x_i^k \rangle}{||a_i||^2} a_i \tag{6.1}$$

where x^k is the current iterate of the solution, λ_k are user-defined relaxation parameters that ensure the solution is not overstepped, and $\boldsymbol{U}_{t(k)} = \text{diag}(\min(1, 1/h_j^t))$, where the h_j^ts are the number of proton histories in the tth block that intersect the jth voxel. In our reconstruction process we use a constant value of λ_k typically between 0.1 and 0.2 and block sizes on the order of 10^5 to 10^6.

To improve the feasibility-seeking aspects of our algorithm, we also include a step of total-variation superiorization (TVS) once per block per iteration. Total variation (TV) is defined as:

$$\text{TV}(z) = \sum_{i=1}^{M-1} \sum_{j=1}^{M-1} \sqrt{(z_{(i+1)j} - z_{ij})^2 + (z_{i(j+1)} - z_{ij})^2} \tag{6.2}$$

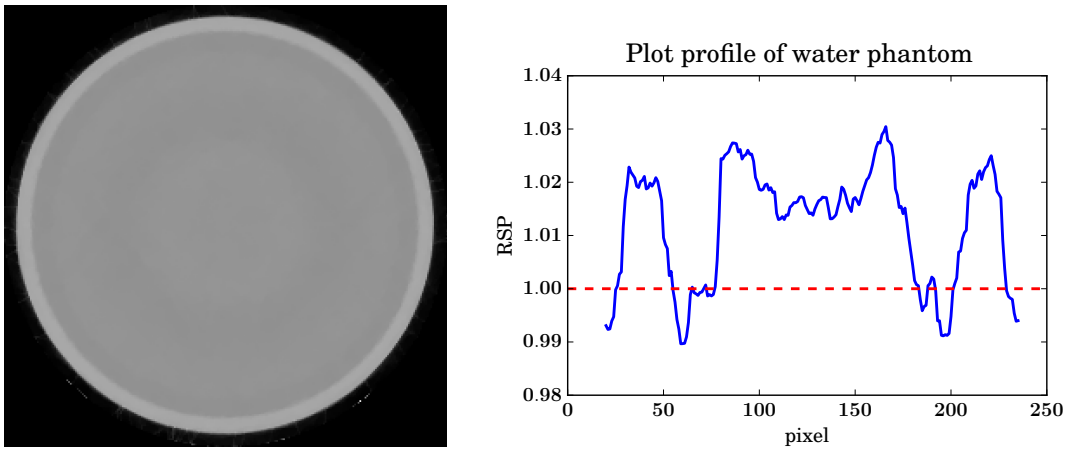
where z is the 2-dimensional ($M \times M$) representation of the image vector. Superiorization refers to the process of finding a "superior" solution with respect to the cost function (in this case, TV) that is also a feasible solution to the linear system $\boldsymbol{Ax} = \boldsymbol{b}$. A superior solution is a feasible solution for which the value of the merit function is smaller than the value of the function at the feasible point that would have been reached if the superiorization process had not been applied, but not necessarily minimal. Superiorization differs inherently from minimization in that rather than seeking a constrained minimum point, it balances the search for feasibility and minimality by seeking a feasible point that is also better than other feasible points. An essential benefit of using TVS for proton imaging is that it smooths the image while preserving edges, which makes it a good choice for low-contrast imaging.

TVS in proton CT image reconstruction is achieved first by perturbing the image vector by a subgradient of the TV multiplied by a relaxation parameter, β. The perturbed image vector is then used for the next iteration of DROP. This is repeated for each block for each iteration. For every iteration, the relaxation parameter, β is decreased by a factor of 2.

6.3 Results and Discussion

Water Phantom

A slice of the reconstructed water phantom is shown in fig. 6.4a. Roughly 80 million proton histories were reconstructed to produce this image. The average reconstructed RSP of the water phantom is 1.01 ± 0.01 which agrees with the RSP of water (RSP=1.00) to within one standard deviation. The reconstructed RSP of the outer cylinder is 1.17 ± 0.01 which agrees with the theoretical RSP of PMMA



(a) Reconstructed water phantom

(b) Plot profile of the water phantom

Figure 6.4: (a) Water phantom reconstructed with pixel size 0.625 mm pixels and 2.5 mm slices. Approximately 80 million histories were used in this reconstruction. (b) Plot profile of water phantom. The profile indicates regions of enhanced and reduced RSP.

(RSP=1.16) to within one standard deviation.

A notable feature of this reconstruction is the presence of low contrast ring artifacts. Figure 6.4b shows a line profile across the homogeneous part of the phantom and indicates that there are regions both of enhanced and reduced RSP. These types of azimuthally symmetric ("ring") artifacts are due to miscalibrations in the MSS. Ring artifacts are a fairly well-understood and pervasive phenomenon in CT. We have studied ring artifacts at length in proton CT and the mechanism by which these form are discussed in Geoghegan's 2015 bachelor's thesis [67].

When protons stop in the interfaces between two stages, energy which should have been deposited in the downstream stage may instead be absorbed by the reflective wrapping, resulting in an assigned WEPL value that is greater than the true WEPL of the proton leading to an RSP enhancement in a ring, which corresponding to that particular WEPL. Similarly, if energy leaks from an upstream to a downstream stage, a WEPL that is too low may be assigned, leading to a ring of

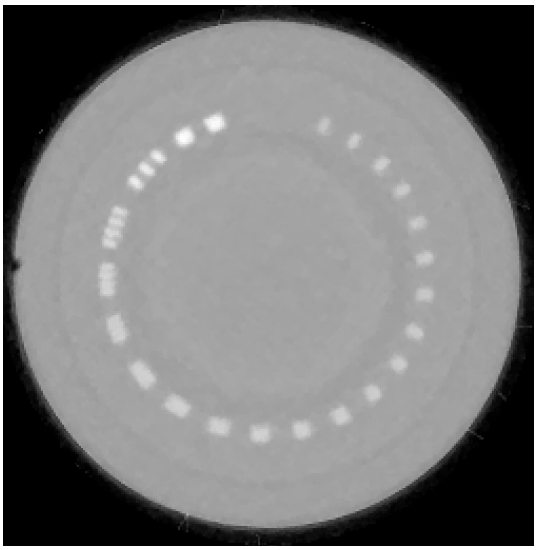


Figure 6.5: Central slice of the Catphan 528 "Line Pair" Phantom. The pixel size is 0.625 mm and the slice thickness is 1 mm. Approximately 140 million histories passed the 3σ cuts.

reduced RSP. These rings have finite thickness due to MCS and range straggling.

Several improvements have been made to the calibration of the MSS with the aim of reducing the presence of ring artifacts in our reconstructions. While these have had some effect on minimizing the prevalence of ring artifacts, more work must be done to correctly handle Birks' effect, which relates light yield and path length in a scintillator, specifically for our energy/range detector.

Line Pair Phantom

The Line Pair phantom was reconstructed both with a slice thickness of 2.5 mm and 1 mm. Approximately 140 million protons passed the 3σ cuts and were used in this reconstruction, which is enough that even with a 1 mm slice thickness, there are still an average of 80 protons per bin. This is enough that statistical noise in the data is sufficiently small. The central 1 mm slice of the Line Pair phantom, which contains the aluminum inserts, is shown in fig. 6.5. The function

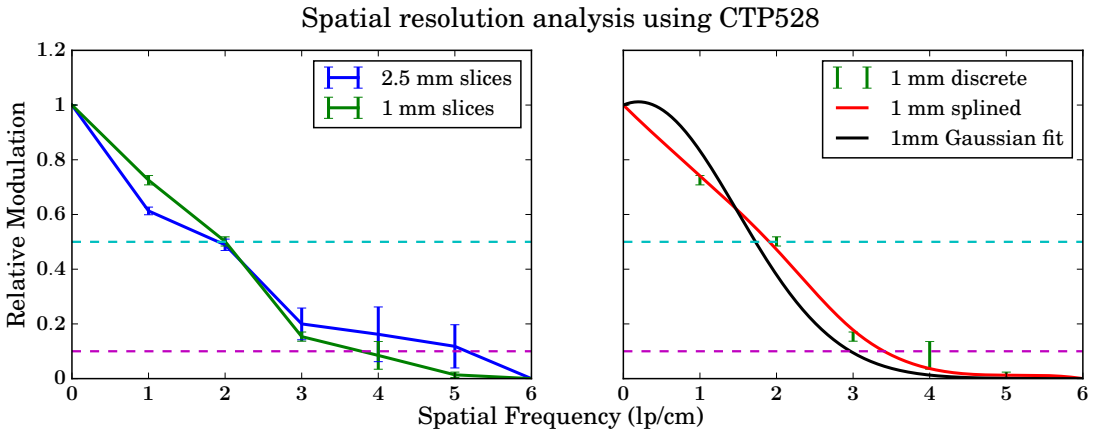
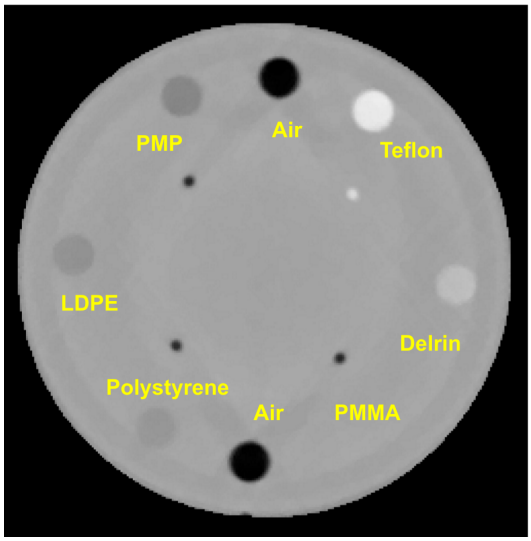


Figure 6.6: The left hand figure compares the spatial frequency measurements of the $z = 2.5\,\text{mm}$ and the $z = 1\,\text{mm}$ reconstructions. The error bars for the $z = 2.5\,\text{mm}$ reconstruction are significantly larger. The right hand plot shows the 1 mm data with both a (red) splined function and the (black) Gaussian fit to the data.

of the Line Pair phantom is to infer the spatial resolution of the imaging system by estimating the relative modulation transfered to the image using the line pair patterns. This is not as robust a technique as the edge spread function method which will be described in chapter 7.

The spatial frequency measurements for both the $z = 2.5\,\text{mm}$ and $z = 1\,\text{mm}$ reconstructions are shown in the left hand panel of fig. 6.6. The error bars for the $z = 2.5\,\text{mm}$ reconstruction are much larger than for the $z = 1\,\text{mm}$ reconstruction because in the $z = 2.5\,\text{mm}$ reconstruction the slice thickness is larger than the height of the inserts (2 mm) and therefore the relative contrast is compromised due to averaging over aluminum and PMMA. The 1 mm reconstruction avoids this and the data is therefore more consistent with itself. In any case, the two curves agree very well indicating that that the spatial frequency measurements in the azimuthal direction are independent of z-slice thickness as expected.

The right hand panel of fig. 6.6 shows the z =1 mm spatial frequency measurements as well as a splined function (red curve) that describes the data, and



Material	RSP
Epoxy	1.14
LDPE	0.98
Polystyrene	1.03
PMMA	1.16
Delrin	1.36
Teflon	1.79
PMP	0.88

Figure 6.7: 2.5 mm slice of the sensitom phantom with 0.625 mm pixels. The RSP of the inserts are given in the table on the right. Approximately 140 million histories passed the 3σ cuts and were used in this reconstruction.

a Gaussian fit (black curve) to the data. The data shown in this plot give a rough estimate of the modulation transfer function (MTF) of the data which will be described in detail in chapter 7. While the MTF of the data is, in general, expected to be approximately Gaussian, fig. 6.6 indicates a Gaussian model does not conform well to the data. The splined curve suggests that maximum spatial resolution in the azimuthal direction at a radius of 48 mm in this phantom is in the vicinity of 3.4 lp/cm.

Sensitom Phantom

A reconstructed slice of the Sensitom phantom is shown in fig. 6.7. Approximately 140 million histories passed the 3σ cuts and were reconstructed to produce this image. The RSP truth values are given in the table to the right of the figure. These were obtained using the PeakFinder instrument with a carbon ion pencil beam at the Heidelberg Ion-Beam Therapy Center (HIT). The truth values are

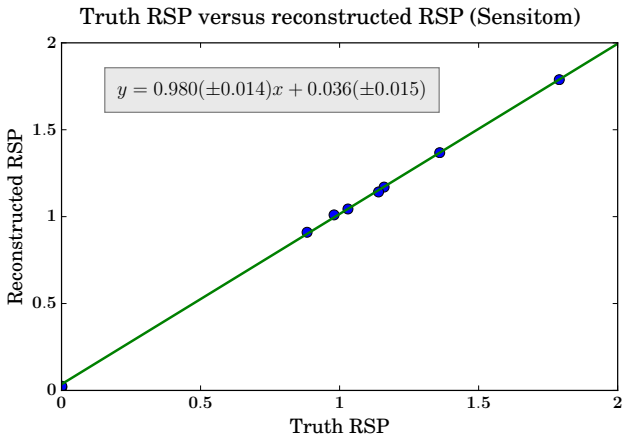
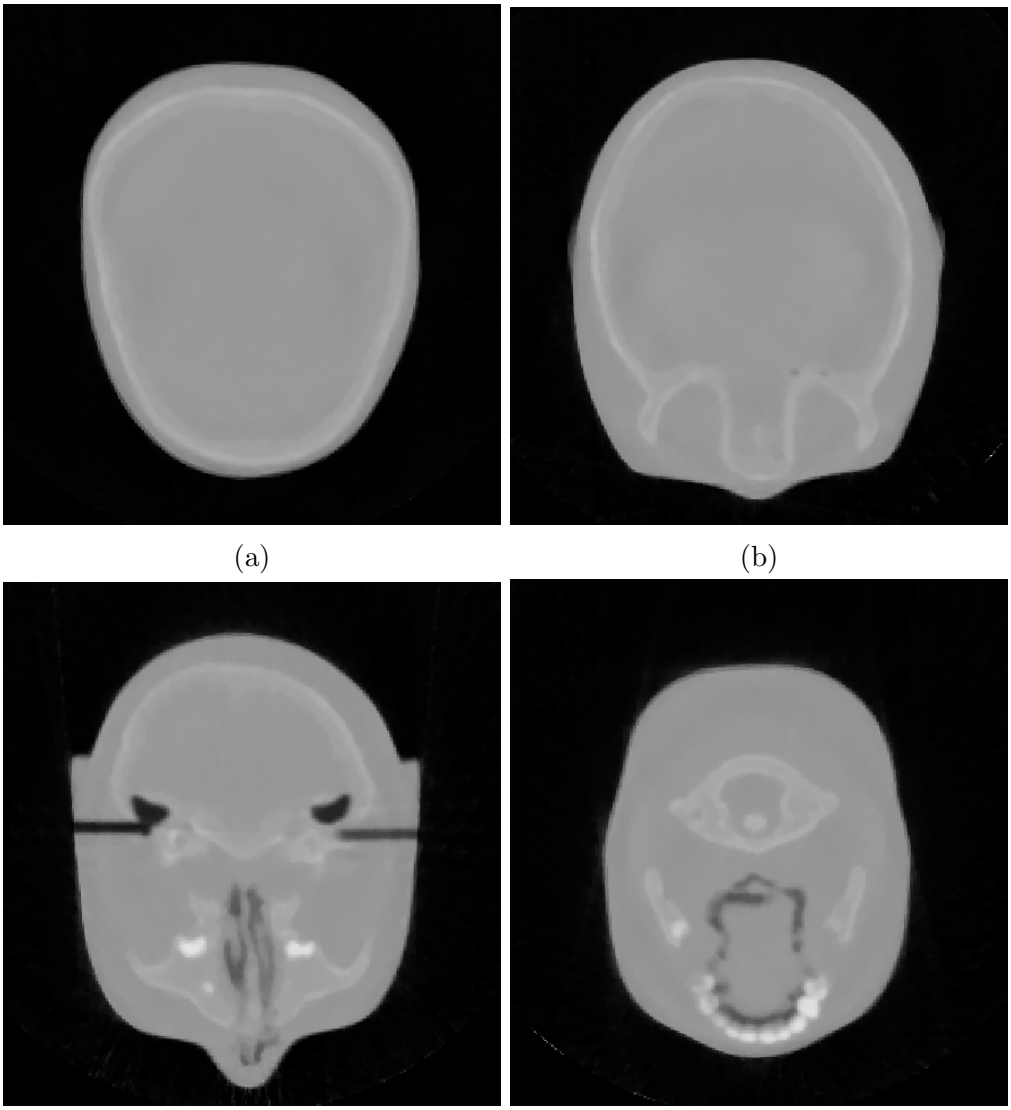


Figure 6.8: Plot of ideal versus reconstructed RSP in the Sensitom. The line of best fit with the uncertainties on the fit parameters is also given.

plotted versus the reconstructed values in fig. 6.8. As indicated in the plot, the line of best fit is described by the equation $y = 0.98x+0.036$. Exact reconstruction of the RSPs would result in a line with the equation $y = x$. The error on the slope was found to be $\sigma_m = 0.014$ and the error on the intercept was found to be $\sigma_b = 0.015$ indicating that the calculated slope agrees with $m = 1$ within 1.4 standard deviations and the intercept agrees with $b = 0$ to with in 2.4 standard deviations.

The individual data points agree with the truth values to within 3% or less for all materials excluding air. Because of the low RSP of air, the percent error is not a good measure of agreement. The reconstructed values of air agree to within 1.7 and 2.3 standard deviations for the top and bottom inserts, respectively.

The triangular artifacts radiating outward from the air inserts in fig. 6.7 are the result of the same phenomenon that causes the ring artifacts in the water phantom. In the Sensitom, though, the RSP is non-uniform where there are inserts, and therefore the rings are distorted according to the WET traversed by the proton.



(a) (b)

(c) (d)

Figure 6.9: Four reconstructed z-slices of the pediatric head phantom. The slice thickness is 2.5 mm and the pixel size is 0.625 mm

Pediatric Head Phantom

Four reconstructed slices of the anatomical pediatric head phantom are shown in fig. 6.9. Figure 6.9a shows a slice through the top of the skull and contains reconstruction of the brain, skull and soft tissue surrounding the skull. Figure 6.9b shows a slice through the orbital sockets near the upper bridge of the nose, and

Table 6.1: CIRS Pediatric Head Phantom Theoretical RSP

Material	**Predicted RSP**
Sinus Cavity	0.22
Soft Tissue	1.03
Brain	1.04
Spinal Disc	1.06
Trabecular Bone	1.10
Dentin	1.56
Pediatric Cortical Bone	1.63
Enamel	1.77

fig. 6.9c shows a slice through the lower part of the nose and sinuses. The ear channels are also visible in this slice, as well as non-descended "adult" teeth. The final slice, fig. 6.9d shows a cross section of the spinal column, the tongue, jaw bone and the lower row of teeth.

The theoretical estimates for materials contained in the head phantom were obtained using eq. (2.6) and are given in table 6.1. The theoretical values are plotted against the reconstructed values in fig. 6.10. The left hand plot in fig. 6.10 shows all eight materials contained in the head phantom. The line of best fit is described by the equation $y = 0.843x + 0.163$ where the uncertainty on the slope is $\sigma_m = \pm 0.124$ and the uncertainty on the y-intercept is $\sigma_b = \pm 0.129$. These parameters agree with $m = 1$ and $b = 0$ to within 1.3 standard deviations. Because the point corresponding to pediatric cortical bone is significantly different from the line of fit, this point was removed and the resulting fit is shown in the right panel of fig. 6.10. The line of fit in this panel is described by the equation $y = 0.972x + 0.029$ where the uncertainty on the slope is $\sigma_m = \pm 0.085$ and the uncertainty on the intercept is $\sigma_b = \pm 0.089$ which agree with $m = 1$ and $b = 0$ to within 0.3 standard deviations.

The reconstructed value of pediatric cortical bone (RSP=1.33) deviates from

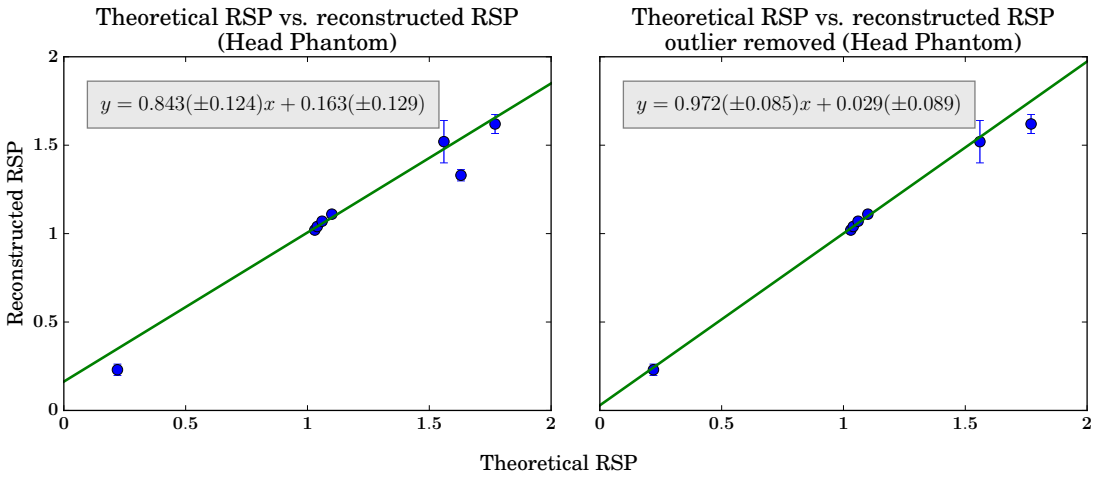


Figure 6.10: Theoretical vs. reconstructed RSP in the pediatric head phantom. The left hand plot shows the comparison for all materials in the head phantom. The right hand plot shows the relationship with the outlier (pediatric cortical bone) removed.

the theoretical value (RSP=1.63) by nearly 20% – by far the largest disagreement out of all the materials reconstructed. This is because the cortical bone in the phantom is very thin (1-2 mm) so it is impossible to measure a homogeneous region of cortical bone when the slice thickness is 2.5 mm, and therefore the RSP of cortical bone gets averaged with the surrounding tissue. These tissues are typically trabecular bone and soft tissue, which have RSPs closer to 1 and so pull down the average RSP value of the voxel.

The largest uncertainty in RSP was in dentin, which is likely because only very small regions of the phantom contain dentin, and therefore several different regions needed to be combined in order to get a sufficient sample. In addition, these regions may also be influenced by surrounding tissues which are typically enamel and trabecular bone. Therefore, in some regions the average measured value may be pulled up by enamel, whereas in other regions the average may be pulled down by trabecular bone, resulting in a large spread in measured values.

Enamel also has large error bars compared to most of the other data points, and

also had the second largest deviation from theory, with a percent error of -8.5%. This is due to a combination of effects including the two previously described. In addition, we have generally found in both the Sensitom and in the pediatric head phantom that for high contrast materials, the RSP of high density inserts tends to be underestimated, while the RSP of low density inserts tends to be overestimated. This is because most of the voxels in our phantoms are near RSP=1, so the average values tend to be pulled in this direction. Since the WEPL represents the line integral of RSPs through the phantom, protons that pass through the nonhomogeneous inserts differ only slightly from protons that pass mostly through voxels with RSP=1.

6.4 Chapter Summary

We have used the prototype proton CT scanner to acquire proton CT data for several phantoms. Three simple cylindrical phantoms and one anatomically accurate pediatric head phantom have all been successfully reconstructed to a high degree of accuracy using the algorithm described in this chapter.

We are currently in the process of improving data preprocessing in order to reduce the presence of ring artifacts, which are evident in the reconstructions of the water phantom and the Sensitom – this, however, may not be possible with the current energy/range detector, and as such may be a good reason to investigate other options (see chapter 8). Additionally, the reduced accuracy in the RSP reconstruction of high-contrast materials suggests a weakness in the current image reconstruction algorithm. Our collaborators at Baylor University are further developing the DROP-TVS algorithm to incorporate robustness in order to improve the fidelity of the RSP reconstruction [68]. We hope future

versions of the reconstruction algorithm will perform even more accurately ($< 1\%$) than the present algorithm.

Chapter 7

Spatial Resolution in Proton CT

7.1 Introduction

The initial discussion of proton imaging versus x-ray imaging in the late 1970s pointed to the lack of spatial resolution in proton imaging compared with the already-successful x-ray CT technology. This led to the abandonment of proton CT as a diagnostic imaging modality. The recent increase in the number of proton therapy facilities, and the lack of imaging support for proton therapy in the treatment room, has led to a renewed interest in proton radiography and CT for improved range definition and treatment verification. The present procedure for proton therapy planning involves converting the Hounsfield value of each voxel in x-ray CT planning scans of the patient into proton stopping power via a stoichiometrically-acquired calibration curve. However, since there is no unique relationship between Hounsfield values and proton stopping power, this procedure has inherent uncertainties of a few percent in the proton range, requiring additional distal uncertainty margins in proton treatment plans. Cone beam x-ray CT is now becoming available for image guidance and treatment plan verification,

however it has distinct disadvantages due to reconstruction artifacts and even larger range uncertainties.

In contrast to x-ray CT, proton CT measures the relative stopping power (RSP) with respect to water of the object directly, eliminating the need for Housfield-value-to-RSP conversion. In the prototype proton CT scanner that we have developed in recent years [54, 55, 57], a low-intensity energetic beam of protons traverses the phantom entirely and stops in a downstream energy/range detector. The entry and exit vectors of each proton are measured in order to determine a most-likely path (MLP) of the proton through the object, and the response of the energy/range detector is converted to the water equivalent path length (WEPL) of each proton in the object. These measurements are made at many angles between 0° and 360° in order to reconstruct a three-dimensional map of proton RSP using an appropriate image reconstruction algorithm [65, 66, 69].

The spatial resolution of proton CT images is fundamentally limited by multiple Coulomb scattering (MCS), which determines the uncertainty of the MLP prediction. Due to MCS, straight-line projections are not accurate enough for clinically acceptable spatial resolution. For this reason, image reconstruction is better accomplished by using a MLP formalism [28]. In its current form, the MLP formalism applies a Gaussian approximation of MCS in water in order to estimate the MLP of the proton through the object assuming it is uniform and consists of water, which is an approximation. In addition to the MLP estimate, the formalism provides an uncertainty envelope for given entry and exit vectors of each proton. This uncertainty envelope provides a theoretical approximation of the limit of spatial resolution imposed by MCS. It is important to note that proton CT is not intended to produce images of high visual contrast, but rather to create an accurate map of RSP for proton radiation therapy planning. Nonethe-

less, spatial resolution is intimately related to the accurate prediction of proton range when protons pass along high-contrast tissue interfaces, i.e., the prediction of range dilution effects [70].

The modulation transfer function (MTF) is widely used to characterize the spatial resolution of imaging systems. The MTF of proton CT has several individual components: the frequency response of the tracking detector measurement, the MCS in the object, and the reconstruction processes. The total MTF of the system is the product of these components. Thus, the MTF of the proton CT scanner can be modeled in the following way:

$$\mathrm{MTF}_{\mathrm{Total}} = \mathrm{MTF}_{\mathrm{MCS}} \times \mathrm{MTF}_{\mathrm{detector}} \times \mathrm{MTF}_{\mathrm{recon}} \tag{7.1}$$

The most common way to measure the MTF is to use a phantom with a thin, high-density metallic wire embedded orthogonally to the scanning plane. The reconstructed image of the wire yields the point spread function (PSF), while the Fourier transform of the PSF yields the MTF. However, in noisy or low-contrast images the use of an edge phantom is the preferred alternative. In this method, blocks of material with high contrast compared with the background, and sharp, straight edges are used as test objects. The edge spread function (ESF) is obtained by overlaying many edge profiles, and is then differentiated to obtain a line spread function (LSF). The LSF can then be Fourier-transformed to obtain the MTF.

The acceptability of any proton imaging system depends on the relative importance assigned to visual quality of an image, (spatial resolution and noise), and RSP accuracy [25]. A theoretical study of spatial resolution in proton and carbon radiography using Monte Carlo simulations was performed by Seco et al. [71], and theoretical estimates of spatial resolution in proton CT have been published by Schneider et al. [72], but, to the authors' knowledge, no comprehensive study

of spatial resolution of an experimental proton CT scanner has been published. In this paper, we present the radial and azimuthal MTFs for a prototype proton CT scanner to characterize the spatial resolution that can be achieved with such a system using 200 MeV protons.

7.2 Defining Modulation Transfer Function

The MTF measures how faithfully an imaging system transfers contrast from the object to the image. For example, in fig. 7.1, row A represents the original pattern being imaged and row B is the pattern after it has been imaged. Row C is the relative modulation of A (essentially a line profile), whereas row D is the relative modulation of B. Notice how B is blurred compared to the original pattern, especially toward the right where the line spacing gets closer together. Row D reflects this blurring, since it appears more sinusoidal than C. The relative modulation is described by the distance between two adjacent peaks and troughs and is indicated on the horizontal axis. The MTF is normalized to 1, so an ideal system will have an MTF of 1, (i.e. profile D would look exactly like profile C) while a completely unresponsive system would have an MTF of 0.

Consider an image given by

$$f(x\,,y) = s(x\,,y) \otimes r(x\,,y) \tag{7.2}$$

where $s(x\,,y)$ is the ideal image, $r(x\,,y)$ is the system response and $\otimes$ is the convolution operator. This says that the reconstructed image, $f(x\,,y)$ is simply a convolution of the ideal image with the system's response to an impulse, which is the PSF (in 2D) or LSF (in 1D). The LSF can be derived from the ESF, which in this case is just the response of a system at the interface of two juxtaposed

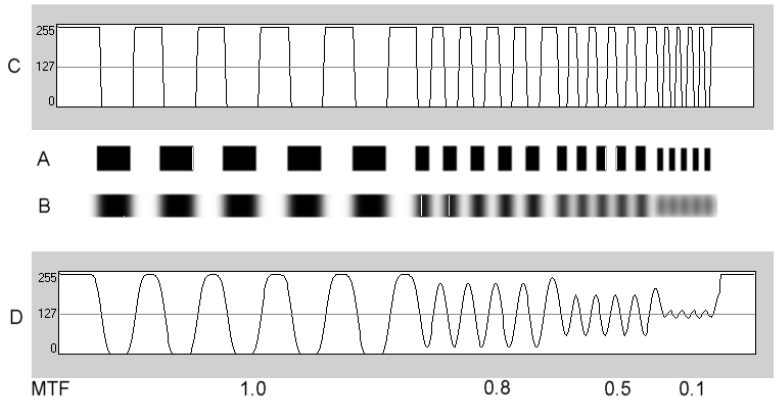


Figure 7.1: A is the original test pattern. B is the image of the test pattern. C is the line profile of the original test pattern. D is the line profile of the image of the test pattern [73]

materials of high contrast[1].

The ESF takes on a characteristic sigmoidal shape which can be described by

$$\mathrm{ESF}(x) = a + \frac{b}{1 + \mathrm{e}^{\frac{x-\mu}{\sigma}}} \tag{7.3}$$

where a is the vertical shift of the function, b is the scaling factor, μ is the mean value of the function, σ is the width. The LSF is found from the ESF by taking the derivative

$$\mathrm{LSF}(x) = \frac{\mathrm{d}}{\mathrm{d}x}\mathrm{ESF}(x). \tag{7.4}$$

Since the MTF is defined to be the normalized frequency response of the system, this requires us to take the Fourier transform of the LSF in order to obtain the MTF. We define

$$H(\omega) = \mathscr{F}\{\mathrm{LSF}(x)\} \tag{7.5}$$

[1]A high contrast-to-noise ratio (CNR) is required in order to evaluate the MTF to avoid contamination by image noise.

and therefore

$$\mathrm{MTF}(\omega) = \frac{|H(\omega)|}{|H(0)|}. \tag{7.6}$$

The value of ω corresponding to the point along the curve where MTF = 0.1 is known as the $\mathrm{MTF}_{10\%}$ which typically characterizes the useful maximum resolution of the system.

There are many components of the system that affect the MTF. Since a convolution in image space is just multiplication in Fourier space, the total MTF is the product of the MTFs resulting from each of the contributing factors. Contributing factors include

1. multiple Coulomb scattering in the object and tracking detectors
2. pixel size
3. the pitch of the SSDs
4. the reconstruction algorithm.

7.3 Materials and Methods

Prototype proton CT scanner

The prototype proton CT scanner consists of a particle tracker composed of silicon strip detectors (SSDs) and a multi-stage scintillator (MSS) for WEPL measurement. The particle tracker is composed of two particle telescopes, one upstream and one downstream from the phantom, each arranged into four layers of four 400 µm-thick silicon wafers with a strip pitch of 228 µm. Each layer of the silicon tracker measures the position of the protons with greater than 99% efficiency. The

front telescope measures the coordinates and the angle of the proton before it enters the phantom while the rear telescope measures the coordinates and the angle of the proton after it exits the phantom. The telescopes have a total sensitive area of $8.6 \times 34.9\,\mathrm{cm}^2$. The tracking planes interface through custom readout ICs and a high-speed data acquisition system based on field-programmable gate arrays (FPGAs). A more complete description of the proton CT scanner hardware can be found in [54].

The MSS measures the residual energy and range of each proton. It is composed of 5 stages of scintillating plastic (UPS-923A, polystyrene), each $36.0 \times 10.0 \times 5.1$ cm^3. Stages through which protons pass entirely contribute directly to total range, while the stage in which the proton stops measures residual energy, which is converted via calibration into residual range. Integrated light guides interface through photomultiplier tubes (PMTs). PMT signals are digitized on a custom board by fast pipelined analog-digital converters and interfaced to the data acquisition by FPGAs. A detailed description of the MSS can be found in [57].

Custom edge phantom

I designed a custom edge phantom (fig. 7.2), which was fabricated by Computerized Imaging Reference Systems, Inc. (CIRS), Norfolk, VA. It was designed for measuring the MTF of the proton CT scanner. The phantom has a diameter of 200 mm, a height of 60 mm, and is composed primarily of water-equivalent plastic (CIRS Plastic Water-LR, RSP=1.007). It contains four groups of rectangular inserts composed of three different tissue-equivalent polymers representing enamel (RSP=1.770), adult cortical bone (RSP=1.685) and lung (RSP=0.217), as well as air (RSP=0.007). The inserts are completely contained within the body of the phantom and have dimensions of $15 \times 15 \times 45\,\mathrm{mm}^3$. The centers of the inserts are

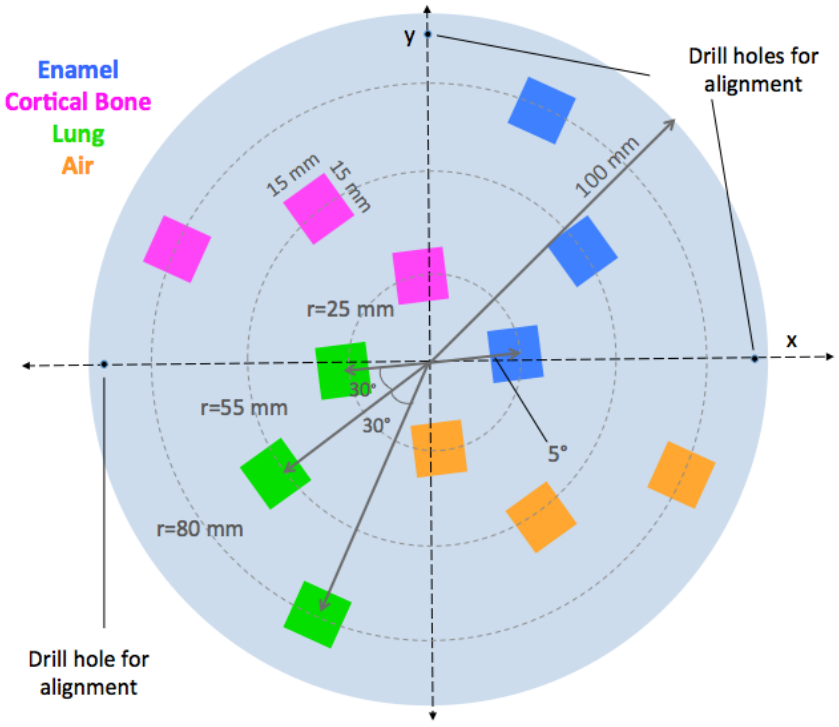


Figure 7.2: Cylindrical edge phantom composed of water-equivalent polymer containing rectangular inserts composed of enamel (blue), adult cortical bone (magenta), lung (green) and air (orange). The phantom is 200 mm in diameter and 60 mm tall.

positioned at radii of 25 mm, 55 mm and 80 mm, respectively, such that their innermost and outermost edges are orthogonal to the radius of the cylinder. Three drill holes of 1 mm radius are located 95 mm from the center of the phantom defining the x-y coordinate axes. The first enamel insert is rotated 5° with respect to the x-axis. The angle between adjacent inserts is 30°.

Image reconstruction

The software for the reconstruction of images was executed on a workstation equipped with two dual-core central processing units, 8 GB of RAM, and an EVGA GeForce GTX680 GPU. The data input to the reconstruction software contained proton tracker coordinates and the water-equivalent path length (WEPL) for each

proton.

Image reconstruction was accomplished first by projecting the entry and exit vectors measured by the silicon trackers onto the reconstruction volume, and then binning the WEPLs of these protons into a sinogram according to the midpoint of the straight-line path between the points at which the protons enter and exit the reconstruction volume. The distribution in each bin was then analyzed and data cuts were performed in order to eliminate protons that fell outside of 3σ from the central WEPL value of the distribution. These cuts are effective in minimizing errors from protons that undergo hadronic interactions in the scanner or phantom.

The resulting sinogram was then passed through a Shepp-Logan filter and was used as input to the Feldkamp-Davis-Kress (FDK) algorithm for a 3-dimensional filtered back projection (FBP) [74]. The FBP image was used both for boundary detection and as a starting point for the subsequent iterative reconstruction. For the iterative reconstruction, the diagonally relaxed orthogonal projection (DROP) onto convex sets was used. This DROP algorithm was further enhanced by interleaved superiorization of the total variation (TV) of the reconstructed image. Details of this DROP-TV superiorization (TVS) algorithm have been described elsewhere [65].

Determination of MTFs

The MTF was determined from the edge spread function (ESF) using a modified oversampling method first described by Judy [75], for characterizing the spatial resolution of an early-generation CT scanner, and later by Fujita et al. for determining spatial resolution of digital radiography [76]. In this work, further modifications and improvements to the method described in papers by Mori and Machida [77], and by Watanabe et al. [78] were employed.

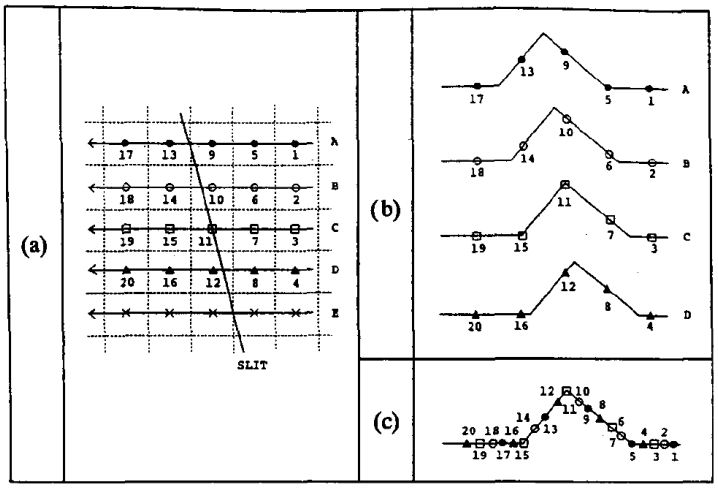


Figure 7.3: Schematic illustrating the oversampling technique using an example line spread function [76].

The oversampling method is well described by fig. 7.3. In (a), a slit passes through a pixel grid where each pixel is numbered 1-20. The slit passes through several pixels at a slightly different position in each one. One sample per pixel is evaluated in order to obtain the LSF, however it is clear from (b) that the data do not well-represent the true shape of the LSF. However, if the center of mass of each LSF is shifted so that the data can be overlaid, the oversampled data now much better represents the true shape of the LSF as shown in (c).

The high-contrast materials were juxtaposed to produce a sharp edge with a slight angle, α, with respect to the principal axes $(x\,, y)$. In general $\tan\alpha$ should not be equal to an integer, or to the ratio of two small integers, in order to ensure that all possible regions of the pixel can be sampled, as the oversampling method requires. The present phantom was constructed using $\alpha = 5°$, $35°$, $65°$, etc. The determination of the MTFs was accomplished using Python version 2.7 with the NumPy and SciPy modules imported [79]. A 2-dimensional bilinear interpolation was applied to the reconstructed image to allow for subpixel-spaced sampling.

For each slice, sampling was performed orthogonal to the edge along a central 10 mm segment of the insert edge at 0.1 mm intervals. The resulting ESFs were overlaid by aligning the 50% values of each ESF in order to yield an oversampled ESF. For noise suppression, the oversampled ESF was rebinned using a bin size equal to the sampling pitch. A non-interpolating (smoothing) cubic spline was fitted to the oversampled and rebinned ESF by minimizing the χ^2 per degree of freedom goodness-of-fit value of the spline to the data, so as not to force the ESF to obey a particular functional form. A line spread function (LSF) was obtained by numerically differentiating the splined ESF. The discrete MTF for each slice was obtained by Fourier-transforming the LSF using a discrete Fourier transform. The discrete MTFs for each slice were then averaged and another smoothing spline was fitted to the data to obtain the final MTF curves.

Experimental scans

The proton CT scanner was installed on the fixed horizontal beamline at the Northwestern Medicine Chicago Proton Center (NMCPC), and the edge phantom was placed on the rotation stage with the drill holes aligned with the alignment lasers, as shown in fig. 7.4. In the first scan, the phantom was rotated in 4 degree intervals and about 4 million proton events were acquired at each rotation angle. In the second scan, the phantom was rotated continuously at a rate of one rotation per minute over a period of 7 minutes, acquiring about 320 million proton events.

The data from the stepped scan were reduced such that the total number of proton histories after applying 3σ cuts was approximately the same for the two scans. This was done by randomly selecting a fixed percentage of the events from each projection angle. For the continuous scan, the incoming angles of the protons

Figure 7.4: Prototype proton CT scanner installed on the fixed horizontal beamline at NMCPC with the edge phantom mounted on the rotation platform.

were calculated using the time stamps of the events and the known angular speed of the rotation stage; the angles were then binned into 1° bins for the FBP.

Six iterations of DROP-TVS were performed on each of the two resulting FBPs, which was decided upon based on previous studies (unpublished) which have indicated that the RSP of the image converges to a satisfactory solution with an acceptable amount of noise after this number of iterations. Two hundred blocks and a relaxation parameter of 0.20 was used. Images were reconstructed using a pixel size of 0.625 mm and a slice thickness of 2.5 mm.

Radial and azimuthal MTFs were calculated at the location of each edge as described above (section 7.3), and the $\text{MTF}_{10\%}$ was determined. A longitudinal MTF was not calculated since the spatial resolution would be primarily dominated by the slice thickness, and because the phantom was primarily designed for in-plane measurements of spatial resolution.

Monte Carlo simulations

The edge phantom was also simulated using TOPAS 2.0, which is based on Geant4 version 10.01.p02, using the standard physics activation for TOPAS [80–82]. The simulation was performed using two different levels of idealization of proton CT scanning: In the first simulation, a 200 MeV uniform parallel beam was incident on an idealized proton CT system where the silicon tracking detectors were replaced with sensitive areas composed of air, thus eliminating MCS in the tracking detectors, and the entrance and exit coordinates were determined exactly. The multistage scintillator was eliminated and instead the energy of the protons was evaluated at the front and the rear trackers, respectively; the energy loss was converted to WEPL via a calibration procedure that correlated energy loss with the water equivalent thickness (WET) of a series of degraders with varying WET placed in the simulated proton CT scanner. In the second simulation, silicon trackers equivalent to those used in the prototype scanner were restored; however, exact measurement of tracking coordinates and all other idealized components remained the same. The measurement uncertainty due to the strip pitch of the SSD was simulated by adding random Gaussian noise with a width of $228\,\mu m/\sqrt{12}$ to the tracking coordinates in the output from the simulation with the silicon tracker. The simulated scans were performed in 4° steps and the resulting data sets contained approximately 140 million histories of which approximately 95 million passed the 3σ cuts in reconstruction.

Theoretical model of spatial resolution

The 2008 paper by Schulte et al. [28] described the formalism for determining the MLP and its transverse position uncertainty for protons traversing a water slab of

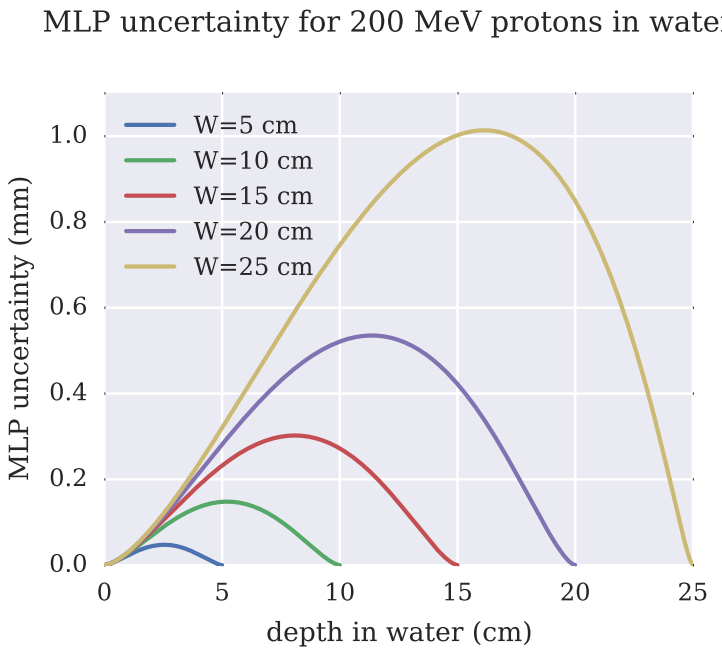


Figure 7.5: Uncertainty $\sigma_W(u)$ in the MLP for 200 MeV protons passing through 5, 10, 15, 20 and 25 cm of water.

constant thickness. The uncertainty curves for 200 MeV protons traversing $W =$ 5, 10, 15, 20 and 25 cm of water are shown in fig. 7.5, which plots the uncertainty $\sigma_W(u)$ versus the depth u in the object.

In order to evaluate a theoretical limit on spatial resolution due to MCS in the phantom, we considered a simplified model of a homogeneous, cylindrical water phantom in a parallel beam. The MLP uncertainty at a depth, u, along the proton's trajectory was obtained from the function $\sigma_W(u)$. As shown in fig. 7.6, the proton tracks that contribute to the azimuthal resolution traverse the same thickness of water, $W = 2\sqrt{R^2 - (w/2)^2}$, where w accounts for the width of the inserts in the physical phantom, regardless of the position of evaluation. On the other hand, protons contributing to the radial resolution traverse varying thicknesses of water depending on the position of evaluation such that $W = 2\sqrt{R^2 - r^2}$ where r is the radial displacement of the edge from the center of the

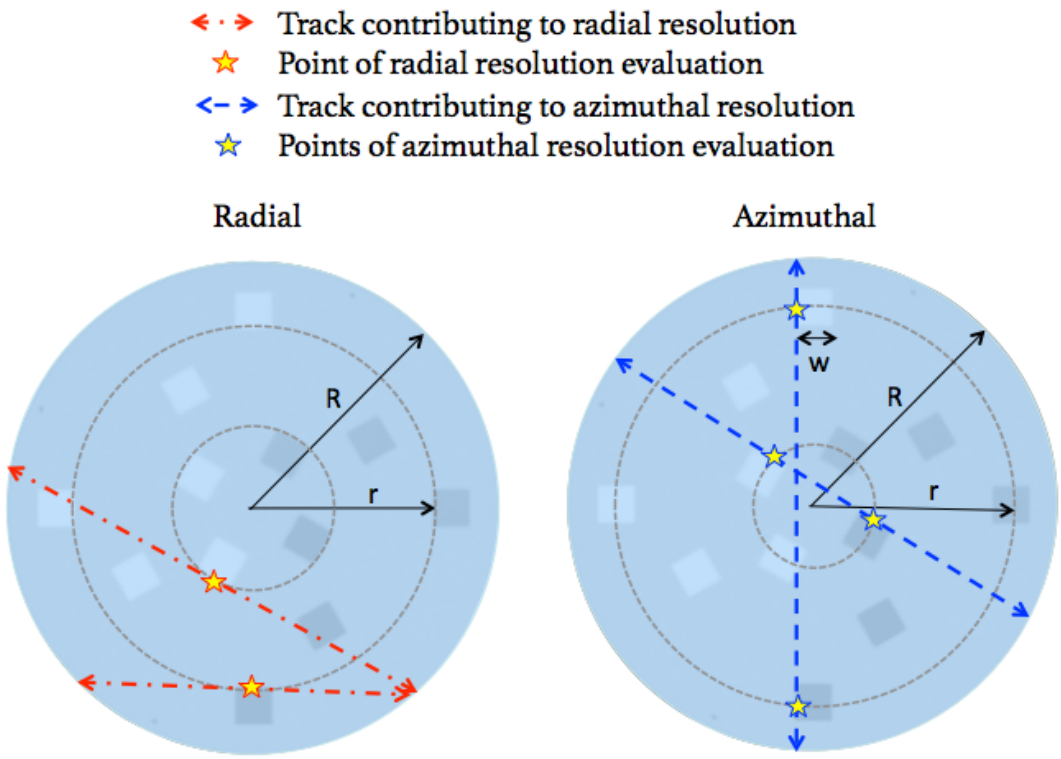


Figure 7.6: Schematics illustrating the derivation of radial (left) and azimuthal (right) limits of spatial resolution using a theoretical model of MCS in a simplified water phantom without inhomogeneities but the same size as the edge phantom. R is the radius of the phantom, r is the radius of evaluation (two are shown but only one is labeled for clarity), and w is the width of the inserts in the physical phantom, which must be accounted for in the derivation of the azimuthal limits. Protons contributing to the azimuthal resolution have approximately the same W, which is longer than those contributing to radial resolution. Azimuthal tracks must be evaluated at both points where the track intersects the radius of evaluation.

phantom.

Although the proton tracks that contribute to azimuthal resolution are typically longer than those contributing to radial resolution, the points at which the MLP uncertainty is evaluated are typically closer to the periphery of the phantom where the MLP is less uncertain, while the tracks contributing to radial resolution are evaluated at $W/2$ near the maximum of the MLP uncertainty curve for any

W. For the azimuthal measurement, the MLP uncertainty curve is evaluated at the two points, $u = \sqrt{R^2 - (w/2)^2} \pm \sqrt{r^2 - (w/2)^2}$ since the cylindrical symmetry permits protons to traverse the phantom in either direction. Although the sum of two Gaussian distributions is not strictly Gaussian, it was found by using a random Gaussian number generator that for the values of uncertainty relevant to this study, the resulting distributions were close to Gaussian, with a width approximately equal to the arithmetic mean of the the two uncertainties, and so this was used to approximate the average uncertainty in the azimuthal direction.

The evaluated uncertainty, σ_0, was taken to be the width of the point spread function (PSF) at the position r. Since the MTF is the Fourier transform of the PSF, the width of the MTF is related to σ_0 by:

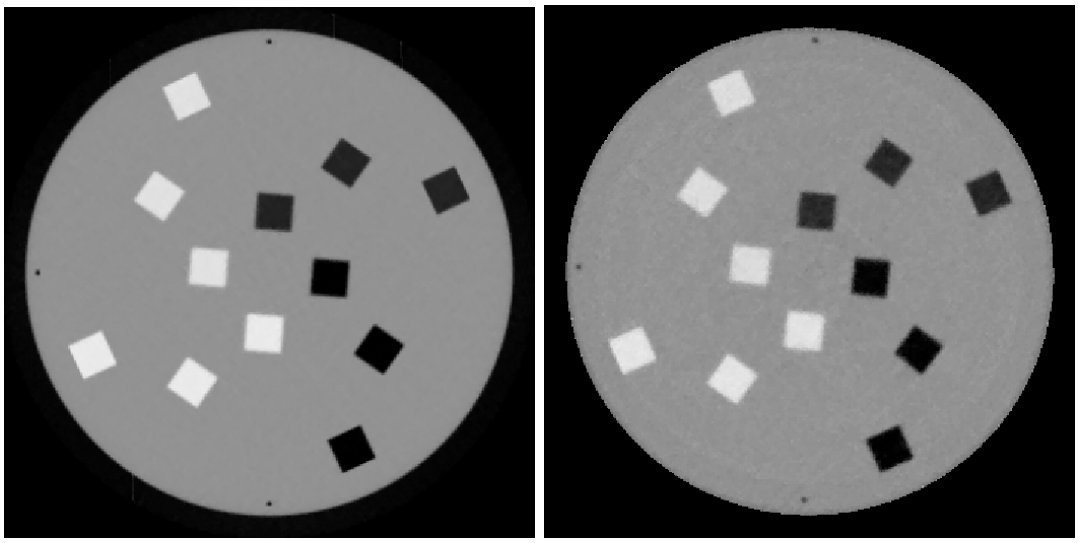
$$\sigma^2_{\text{MTF}} = \frac{1}{4\pi^2\sigma_0^2}. \tag{7.7}$$

Therefore, the theoretical limit on the spatial resolution, defined as the spatial frequency at which the MTF has fallen to a value of 10%, may be obtained by evaluating the 10% point of the Gaussian function $f(k) = \mathrm{e}^{-2\pi^2\sigma_0^2 k^2}$.

7.4 Results and Discussion

Figures 7.7a and 7.7b show slices of the reconstructed edge phantom from the most ideal simulation and experiment respectively. Each of these images was reconstructed from roughly 95 million proton histories, using six iterations of DROP-TVS. A slice thickness of 2.5 mm and a pixel size of 0.625 mm were used.

The $\text{MTF}_{10\%}$ of the ideal simulation was compared with the limits imposed by MCS as predicted by theory. The resolution was evaluated for images with pixel sizes of 1.0 mm, 0.625 mm, and 0.5 mm. The results of this analysis are presented in fig. 7.8. The dashed curves in this plot represent the theoretical limits of spatial



(a) Simulation (b) Experiment

Figure 7.7: Reconstructions of 4 degree stepped scan of the edge phantom in (a) the most idealized simulation, and (b) the stepped experimental scan. The pixel size is 0.625 mm and the slice thickness is 2.5 mm.

resolution imposed by MCS and pixel size. It is evident from the left hand panel that for the 1 mm pixel size the theoretical limit of the resolution (5 lp/cm) is determined by the pixel size. This limit is the equivalent of $1/2\omega_N$, where ω_N is the Nyquist frequency. For the pixel size of 0.625 mm (central panel) the limiting spatial frequency due to pixelation is 8 lp/cm. In this case, MCS limits spatial resolution for r <40 mm. For the 0.5 mm pixel size, the limiting spatial frequency due to pixelation is 10 lp/cm, and MCS limits the spatial resolution for $r < 60$ mm.

The solid lines in fig. 7.8 represent the spatial resolution of images reconstructed from the data from the most ideal simulation for pixel sizes of 1.0 mm, 0.625 mm, and 0.5 mm. The data points were obtained by taking the average of all the edges at the same radial displacement. The error bars represent the range of the values. Evidently, the reconstructed spatial resolution is significantly degraded compared with the theoretical limits, even in this idealized simulation in which the influence of the silicon detectors on spatial resolution has been entirely elim-

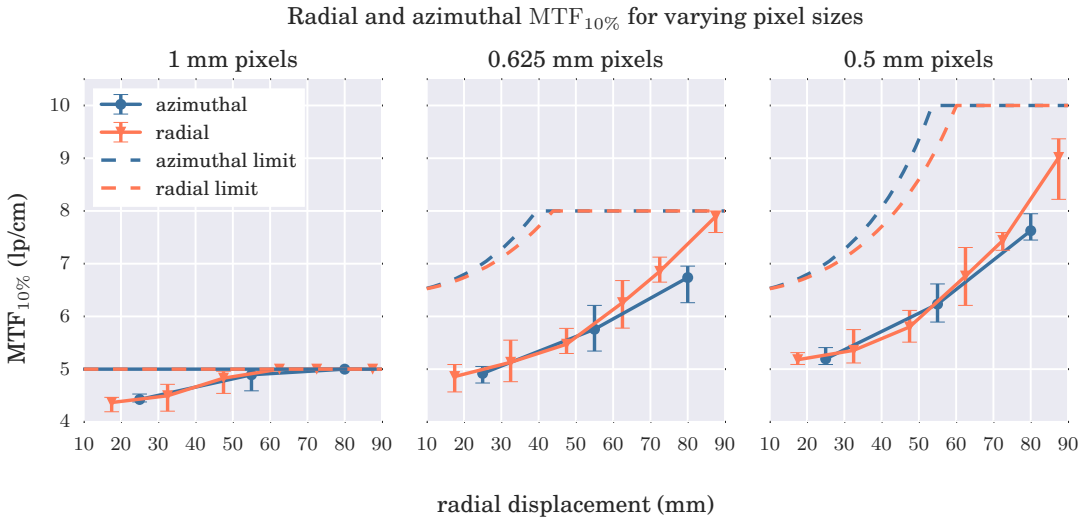


Figure 7.8: Plots of theoretical limits for azimuthal and radial spatial resolution (dashed lines), and the evaluated $MTF_{10\%}$'s for the ideal simulation (solid lines) as a function of radial displacement of the evaluated edge, reconstructed with three different pixel sizes: 1 mm (left), 0.625 mm (center), and 0.5 mm (right).

inated. This result suggests additional degradation of spatial resolution, which may be introduced by inhomogeneities in the phantom and by the reconstruction process. There are at least two explanations for this but there may be other contributing factors: first, the MLP model used in image reconstruction assumes a uniform material (water), which is a reasonable approximation for soft tissues in the absence of large inhomogeneities such as bone and air. However, when high contrast inserts are present along the proton path, the MLP of the proton is incorrectly approximated and errors are propagated along the proton's path through the phantom causing an overall reduction in spatial resolution. Second, the theoretical model discussed above ignores the effects of iterative reconstruction algorithms on spatial resolution, which are difficult to model. It has been observed that performing more iterations of DROP-TVS increases the spatial resolution, but does so at the expense of image noise, which also increases. Consequently, after six iterations, the current DROP-TVS algorithm does not achieve the spa-

tial resolution that is possible with a larger number of iterations, which explains, perhaps, the largest part of the discrepancy.

The resolution of the azimuthal edges tended to be more degraded than that of the radial edges towards the periphery of the phantom. It is likely that this unexpected result is a combination of several geometric factors that distinguish the azimuthal from the radial edges. First, because of the asymmetry of the MLP uncertainty curve, the azimuthal resolution differs depending on whether the proton intersected the edge upstream or downstream from the center of the phantom. Near the periphery of the phantom, the downstream uncertainty can be two to three times larger than that of the upstream uncertainty. The theoretical model takes the average of these two values, but this is an oversimplification. Second, there are two edges along proton tracks in azimuthal direction, while there is only one edge for radial proton tracks. Lastly, the resolution of the azimuthal edges are averaged over a 10 mm range of radial distances along the edges, whereas the radial distance of the radial edges is practically constant.

It is clear from fig. 7.8 that there is still room for an improvement in spatial resolution of up to 50%, which may be achieved by improving the the MLP model. This should include knowledge of the position of high and low density materials after the initial FBP reconstruction has been performed. Furthermore, future versions of the reconstruction algorithm should focus on improving the quality of the initial iterate for DROP-TVS and improved noise suppression in order to achieve better spatial resolution.

A plot of the $\mathrm{MTF}_{10\%}$'s versus radial displacement for all edges in fig. 7.7b, reconstructed from the experimental proton CT data, is given in fig. 7.9. The radial plot has data at each of six radial displacements, while the azimuthal plot has data at three radial displacements. Since there are two azimuthal edges at each

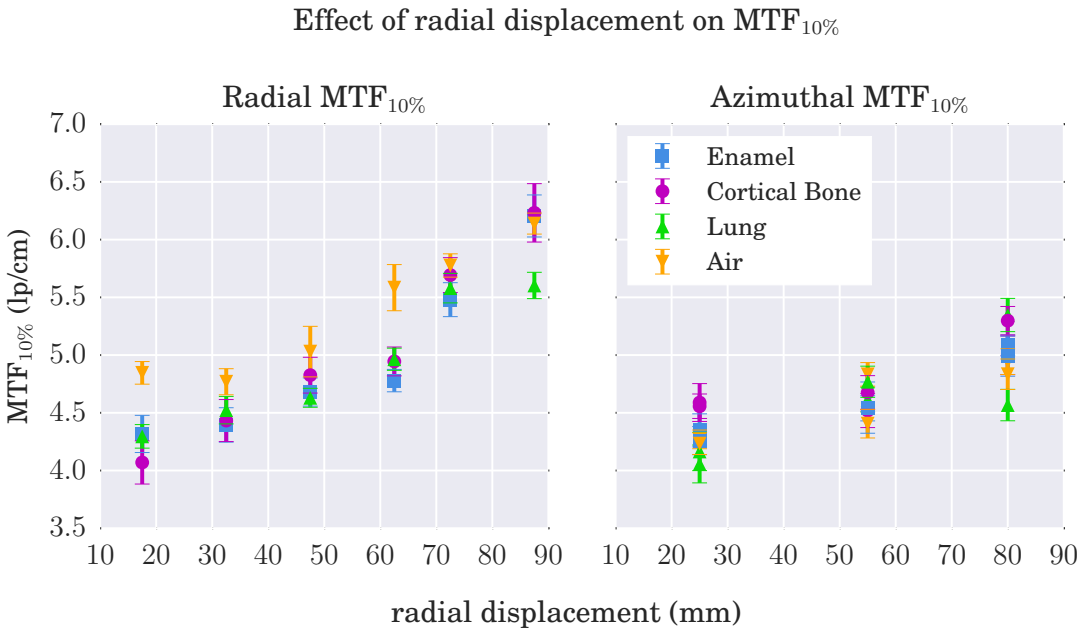
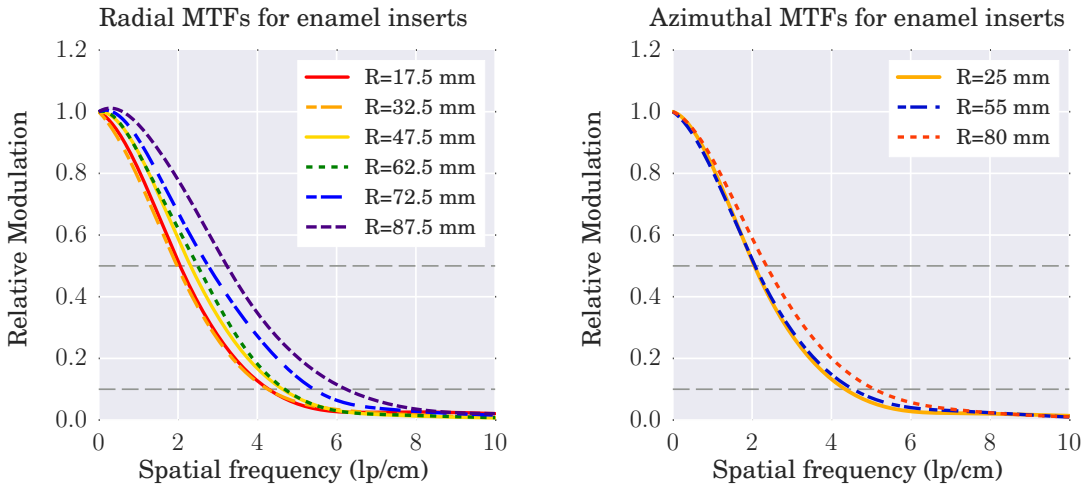


Figure 7.9: Plots of the $MTF_{10\%}$'s for the radial MTFs at the six radial edges for each of the four material inserts (left), and for the six azimuthal edges at three radial displacements for each of the four material inserts (right) for the image reconstructed from experimental data (fig. 7.7b) using 0.625 mm pixels and 2.5 mm slices.



(a) (b)

Figure 7.10: (a) The radial MTFs for the six radial enamel edges. (b) The azimuthal MTFs for three azimuthal enamel edges. The gray dashed lines in each figure indicate the $MTF_{50\%}$ and the $MTF_{10\%}$.

of the three radial displacements, each value of r in the azimuthal plot contains twice as many data points as in the radial plot (eight data points instead of four). The plots show that the spatial resolution in the radial direction ranges from 4.07 ± 0.19 lp/cm at a radial displacement of r =17.5 mm to 6.23 ± 0.25 lp/cm at r =87.5 mm. In the azimuthal direction, the spatial resolution ranges from 4.06 ± 0.16 lp/cm at r =25 mm to 5.35 ± 0.14 lp/cm at r =80 mm. The error bars in this plot represent the statistical uncertainty (1 standard deviation) on the average value of the $\text{MTF}_{10\%}$ and were obtained by bootstrapping the set of discrete MTFs for each slice. In addition, fig. 7.10 shows the radial and azimuthal MTFs for the enamel inserts. The dashed gray lines indicate the $\text{MTF}_{50\%}$ and the $\text{MTF}_{10\%}$. The MTFs of the three other materials behave similarly.

As expected $\text{MTF}_{10\%}$'s shows a strong radial dependence. The material dependance of the radial MTF may be understood in terms of the amount of scattering the different density materials cause, which affects the MLPs in different ways. Scattering will be largest for the high-density materials (bone and enamel) and for the lower proton energies at smaller radial displacements. This is explains why the air inserts have consistently higher edge resolution then the bone and enamel inserts, in particular for the central locations. One would expect the low density lung inserts to perform similar to air, but it is believed that due to manufacturing difficulties, the lung edges are inherently of less sharp quality compared to the air interfaces, which are formed by the background material of solid water.

The material dependence of the azimuthal edge resolution is less pronounced than that of the radial edge distribution, despite the fact that the radial dependence is conserved. The azimuthal edge resolution is further compounded by the other edges that are intersected by the azimuthal proton tracks. This is clearly seen by the difference in the resolution of pairs of edges from the same insert (e.g.

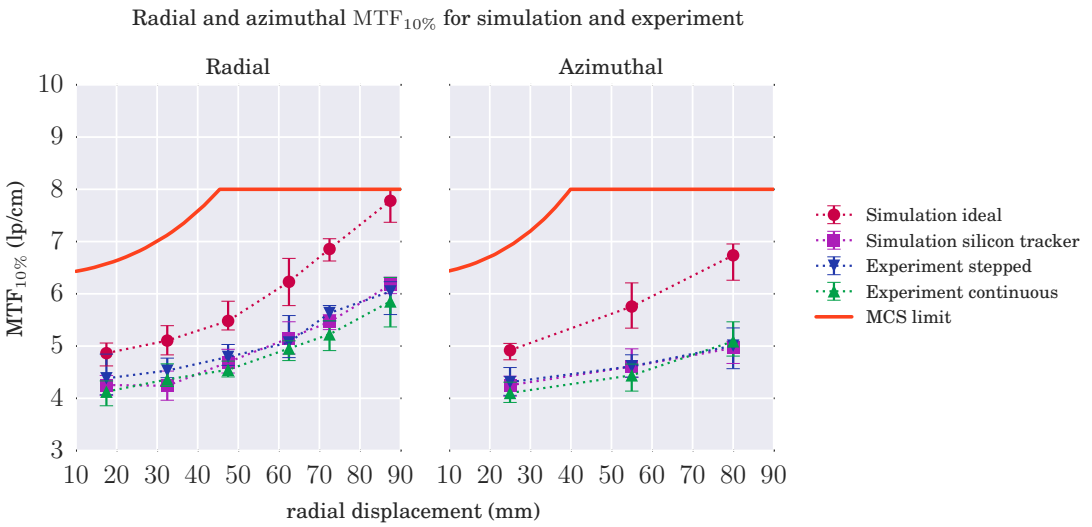


Figure 7.11: A comparison between the most ideal simulation, the simulation with the realistic silicon tracker, the stepped experimental scan and the continuous experimental scan. The average $MTF_{10\%}$ of enamel, cortical bone, lung and air are plotted as points, with the error bars indicating the range of measurements. The pixel size in the images was 0.625 mm.

air and lung at 55 mm and 80 mm radial distance, respectively). The protons probing the edge on one side also pass through a high-density insert while those probing the opposite edge pass through a low-density insert, which affects them less.

Figure 7.11 shows a comparison between the $MTF_{10\%}$ values of the most ideal simulation, the simulation with the realistic silicon tracker, the stepped experimental scan and the continuous experimental scan. The plot shows strong agreement between the idealized simulation with the silicon tracker and the two experimental scans. Evidently, replacing the ideal tracking detectors with silicon detectors significantly reduced the spatial resolution overall. This is because the MLP model assumes exact knowledge of the entry and exit points from the phantom, whereas due to MCS and point resolution uncertainty in the silicon trackers, the MLP endpoints are fairly uncertain. This increases the uncertainty at every point along the

MLP curve, with the most significant increase occurring at the endpoints, where the uncertainty is zero in the ideal case. This results in the largest reduction in spatial resolution occurring near the periphery of the phantom, as is indicated by the plots.

In addition, it appears that at large radial displacements, the azimuthal resolution is reduced by a greater amount than the radial resolution (about 25% compared with 20%). This is expected because these protons have much larger W than their radial counterparts, resulting in the azimuthal protons having less residual energy when they enter the rear tracker. Because of this, they are subject to a greater amount of MCS in the rear tracker than protons with higher residual energy.

As indicated by fig. 7.11, the range of MTF values from continuous experimental scan, and simulated and experimental stepped scans overlapped for both radial and azimuthal resolutions. This means that a stepped scan with 4 degree steps will give equivalent results in terms of spatial resolution, which is important when the proton beam is pulsed (as is the case for synchrotrons) and thus continuous scans are not possible.

7.5 Chapter Summary

We have presented an analysis of the spatial resolution of a prototype proton CT scanner. Using a simplified theoretical model based on the MLP formalism we have established an upper limit on the spatial resolution that is physically achievable by the system and have compared this limit with simulated systems at varying levels of idealization, and with two experimental scans.

It has been shown that a nearly 50% increase in spatial resolution at some

radial displacements may be possible by using an improved MLP estimate. MCS in the silicon detectors is the second largest contributor to the degradation of spatial resolution which poses a convincing case for a more transparent proton tracking system. Since the silicon wafers were left over from a previous and unrelated project [83], their design was not optimized for application in proton CT. Single-sided detectors of half the thickness (200 µm) could be easily manufactured and we expect these would result in at least a 5-10% improvement in spatial resolution. Using double-sided SSDs could improve this further. The strip pitch of the SSDs results in a small reduction in spatial resolution due to uncertainty in the coordinate measurement. This could be reduced by using SSDs with a finer pitch, but is a fairly insignificant effect compared with the other factors discussed, and would increase the complexity of the detector readout. Hence, we have demonstrated an understanding of how all the components of the prototype system affect the spatial resolution.

In order to further improve spatial resolution of proton CT, we suggest to develop more transparent tracking detectors, and to implement improved path estimation algorithms together with noise-reducing but edge preserving iterative reconstruction techniques. This could include starting from a more advanced FBP algorithm and including heterogeneities that become apparent from the initial reconstruction in the MLP formalism of the subsequent iterative reconstruction. This is expected to lead to spatial resolutions that are closer to the theoretical limit. Lastly, with the developing of higher-energy medical accelerators for protons or other ions (e.g. helium), spatial resolution of charged particle imaging is expected to further improve over time.

Chapter 8

Present and Future Directions in Proton CT

8.1 Introduction

At this point, we have discussed several successful experiments that utilized our prototype proton CT head scanner. We have achieved many of our goals in this phase, such as:

- 1 MHz data acquisition rate
- a sensitive area large enough to cover the width of a human head
- minimal loss of events due to protons passing through gaps in the tracker and the energy/range detector.

In addition, we have achieved:

- WEPL resolution close to the theoretical limit of 2.8 mm

- hit recovery of protons that traverse gaps in the tracking detector further reducing the number of lost events
- the ability to continuously scan an object in an isochronous cyclotron beamline.

We have evaluated the error of reconstructed RSP and have found better than 3% agreement for most materials, and better than 1% error for many, and we have thoroughly evaluated the spatial resolution of the system. Regardless, fundamental limitations of the system, specifically the inability of the device to handle multiple coincident protons, range straggling in the calorimeter, and the amount of time required to reconstruct even a single image must be overcome in the next stages of development. In this chapter we will discuss possible hardware and software upgrades for a proposed phase-III prototype proton CT head scanner.

8.2 Phase-III Hardware Upgrade Possibilities

The biggest limitation of our current system is the inability of the energy detector to distinguish between coincident protons. Overcoming this issue could lead to significant improvement in performance. With our present device, a lack of lateral segmentation in the calorimeter makes it impossible to discriminate between two or more events that are detected at roughly the same time. Because we operate at low intensity such that proton events are Poisson-distributed in the accelerator RF buckets, this happens a relatively low percentage of the time, however, it still requires that 5-10% percent of events acquired by the detector (this quantity varies depending on the data rate) be eliminated. This is not ideal because any events that are not used in reconstruction are "wasted" dose to the patient. In

addition, if coincident protons could be detected, this could pave the way to the ability to scan at higher intensities, and further reduce scanning time.

Potential Replacements for the MSS

The MSS was a fairly inexpensive and simple device to build, as it is constructed from relatively few components, however the trade-off for this was that developing adequate electronics for the device, and establishing a satisfactory calibration for the system was very time consuming and complicated. The complexity of the calibration was largely a result of range straggling which posed a significant problem when protons stopped in or near the interfaces between the stages of the MSS. In the future, we hope to replace this detector with one that is not only more sophisticated but hopefully easier to calibrate. We've considered several possibilities to replace the 5-stage MSS.

The most attractive contender, at this time, is a range counter. This was proposed early on in the present phase, but was ultimately rejected because of the large number of physical components and the complexity of assembling the device. A range counter with adequate resolution for our needs requires upwards of 60 scintillating stages, all of which need to be integrated with readout electronics. Although the range counter requires a bit more overhead than the MSS, the range resolution is completely independent of proton range, and the calibration is remarkably straightforward, since the range of the proton is defined by the center position of the stage in which the last energy deposition is detected. Additionally, it would be fairly simple to laterally segment the first few layers in order to break ambiguities in track reconstruction as a result of multiple coincident protons. We would propose a range counter that includes ten to twenty lateral segments in the first few layers of the range counter. Signals in the lateral segments could then be

matched to proton tracks in the particle telescopes. The signals in the remainder of the detector would have to be separated in time, but since the scintillator can be quite a bit faster than the tracker, this would not be a limiting property.

The second limitation is range straggling in the energy/range detector. Although we expect a certain amount of range straggling in the tracker and phantom, using an energy or range detector composed of less material than an MSS or range counter would eliminate a significant amount of uncertainty in the range measurement. A magnetic spectrometer or a time-of-flight detector are two alternatives which would eliminate a large amount of range straggling in the calorimeter.

Use of a magnetic spectrometer for proton CT is a fairly old idea and was used by Takada et al. as early as 1988 [84]. A magnetic spectrometer utilizes the Lorentz force acting on a charged particle in a magnetic field in order to measure the momentum of the particle. Knowing the magnitude of the magnetic field, the momentum of a proton can be obtained from the radius of the curvature of the proton path within the field.

The energy resolution of a magnetic spectrometer is related to the spatial resolution by $\mathrm{d}E/E = 2\,\mathrm{d}r/r$, that is, for a given energy resolution, the deflection must be measured twice as well. For example, the 1σ resolution for 80 µm SSDs is roughly 23 µm. For 1% position resolution (2% energy resolution) a deflection of at least 2.3 mm must be measured, and therefore, for a 1 m long detector, a uniform B field of roughly 0.015 T is required. A field of this strength is reasonably achievable with permanent magnets, however achieving uniformity of the field over an area of approximately 1.0 m $\times$ 0.5 m would be a difficult and expensive undertaking.

Another solution to range straggling would be a time-of-flight detector (TOF). A TOF is composed of two slabs of fast plastic scintillating material, one upstream and one downstream from the phantom, and it measures the time it takes for a

particle to travel between the two detectors. This time should be directly proportional to the particle's velocity, which is proportional to the square root of energy. Like the magnetic spectrometer, the energy resolution is twice the velocity resolution. Good signal-to-noise ratio (SNR) requires slabs to be relatively thick, for adequate light collection; however, the thicker the slabs, the more difficult the timing calibration becomes. State-of-the-art TOFs have time resolution on the order of 100 ps, but at the scale required for a CT detector system, we would require at least 50% better timing resolution. That said, TOF technology is still maturing and it is very plausible that in the not-so-distant future, the required timing resolution could be achieved. Sadrozinksi et al. have proposed an ultra-fast silicon detector (UFSD) which would allow for ∼10 ps and ∼10 µm resolution simultaneously [85].

In any case, the most simple upgrade would require adding a sixth layer to the present MSS, in order that the beamline at NMCPC could be used at its nominal energy of 230 MeV. Because our system is built to operate with 200 MeV incident protons at LLUMC, when operating at NMCPC the beam energy must be reduced using degraders upstream from the beam aperture. This is not ideal since range straggling in the degrader causes a fairly large energy spread of the incident beam, which results in increased WEPL calibration uncertainty and increased image noise. If we were able to operate at 230 MeV, not only would these issues be eliminated, but spatial resolution would also increase due to reduced MCS in the phantom and trackers (although density resolution would decrease).

Tracker Upgrades

There are several tracking technologies that have been proposed for application in proton CT. In general, each of these detectors come with pros and cons and

the choice of one or the other typically results in the sacrifice of either speed or efficiency.

Scintillating fibers (Sci-Fis) are presently being used by several groups studying proton CT [25]. Sci-Fis are composed of a polystyrene core containing a fast fluorescent dye with a matched excitation energy, and one or more thin layers composed of polymers with a lower refractive index than polystyrene. When excited by a signal, light is transported via total internal reflection along the fiber. Benefits of scintillating fibers include high sensitivity, high gain, and a signal decay time of $< 10\,\text{ns}$. One major drawback, however, is that the spatial resolution is directly correlated with the fiber diameter, which is inversely correlated with light yield. This problem can be overcome by clever multi-layered arrangements, but this too comes at a cost in terms of the required number of readout channels and the overall budget for the device [86]. In addition, Sci-Fis possess a SNR less than 10, which we believe is not sufficient for application in proton CT.

Complementary metal oxide semiconductor (CMOS) pixelated detectors are a class of active pixel sensors that have a high radiation tolerance, nearly 100% detection efficiency and high time resolution. CMOS sensors are typically used to detect visible light and are on the verge of replacing CCDs as state-of-the-art in biological imaging and medical applications, however they also work as particle detectors [87]. Although CMOS pixel sensors have high spatial resolution and 2-D sensitivity, the readout of the detectors is less than optimal for proton CT, requiring that data packets be shifted along rows of pixels until they arrive at the readout chip. As a result of this, a single particle detection rate of 1 MHz or more would not be feasible. Pixel detectors with individual readouts are in development, and would eliminate this major obstacle, however these will likely be restrictively expensive (at least initially) and would likely afford little benefit

compared with SSDs.

Plasma panel sensors (PPS) and gas microstrip detectors (GMDs) are two additional technologies that could be applied in proton CT tracking. PPSs are a new application of plasma display panel technology (PDP) in the R&D phase. PDP technology is currently used in television and other graphical display devices. PDPs are composed primarily of two glass plates with lines of electrodes on the internal surfaces. The gap between the two plates is filled with gas at relatively low pressure. A pixel is made of an electrode intersection and a gas gap. In PDP technology, plasma discharge is actively induced at an addressed pixel, whereas in PPS technology, the discharge would be caused by ionizing radiation entering a cell [88]. GMDs are a type of lithographically fabricated multi-wire proportional chamber that use an alternating array of wires held at high voltage, and strips of conductive material held at ground potential printed on a substrate. This arrangement is used in order that a uniform electric field draws negative ions toward the annode wires. The chamber is filled with gas such that any ionizing particle that passes into the chamber will ionize the surrounding gas resulting in a localized cascade. This cascade collects on the nearest wire and results in a charge proportional to the ionisation of the detected particle [89]. Of all the above mentioned technologies, GMDs are probably the most mature and viable technology at this time; however, the advantages over SSDs, if any, are not clear.

Careful consideration may or may not result in a change in tracking technologies since SSDs are reliable and well understood, and to this point, appear to be a very good choice for this application. However, it has been made clear in chapter 7 that the thickness of the present tracking detector is one of the primary factors contributing to the degradation of spatial resolution. The simplest upgrade would be to use thinner SSDs, which would be more appropriate for ap-

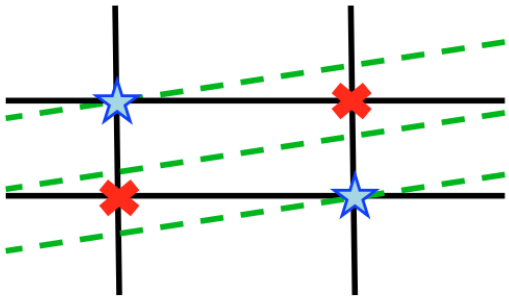


Figure 8.1: Schematic illustrating the utility of a rotated detector layer. The black lines indicate T and V strips. The blue stars represent hits and the red x's represent ambiguities in the coordinates of these hits. Including an additional layer of SSDs disambiguates multi-proton events by adding a third strip ID to define the positions of these hits. The angle of rotation of the third tracking layer (green) is exaggerated for clarity.

plication in proton CT than the detectors that are presently being used. SSDs of 200 µm thickness or less are commonly available and would result in a significant improvement in spatial resolution due to reduced MCS in the tracking detectors. A further upgrade would be to use double-sided SSDs, which could reduce the thickness of each tracking plane by a factor of four or greater compared with the present thickness.

Another improvement would be to incorporate an additional tracker plane oriented at a slight angle with respect to the t-v axes, which would help to disambiguate some tracks in the case of multi-proton events, as in fig. 8.1. Two hits in the t-v plane are indicated by blue stars, whereas red x's indicate the coordinate ambiguities associated with these two hits. By including a third layer (green) rotated at a slight angle (typically about 3°) with respect to the t-v plane, the additional strip ID makes clear at which intersections the hits lie. This would improve the accepted-proton efficiency of the detector system, resulting in less wasted dose, and may also improve spatial resolution due to improved definition

of proton paths.

8.3 Software Upgrades

Chapter 7 indicates that there are several improvements in the reconstruction code that can be made in order to improve spatial resolution of reconstructed images. The DROP-TVS algorithm did not achieve the theoretical limit after six iterations, and although spatial resolution increased with every additional iteration, so did the amount of image noise, which required an optimal image to be selected based on a balance between the two. Although we would like to increase noise suppression at the same time as maintaining the same degree of edge preservation, this may not be possible and must be investigated further. Therefore, the primary focus should be on making the MLP better by incorporating information about inhomogeneities, and improving the initial iterate for DROP-TVS. The latter may be accomplished using backprojections along MLPs, which has been shown to significantly reduce noise in projection images (work in progress).

With a fully functional prototype capable of collecting high quality data, it seems the time is ripe to reevaluate all the options for image reconstruction algorithms. Many improvements to the reconstruction software used in this thesis are currently in progress. Some of these improvements deal directly with the reconstruction algorithm while others deal with implementation of reconstruction and workflow. In addition, other research groups have recently been developing their own algorithms for proton CT which must also be evaluated in terms of accuracy and spatial resolution.

Algorithms

Karbasi, Schultze and Schubert at Baylor University are working on a complete overhaul of the original Penfold reconstruction code, which is what was used to produce the images in this thesis. Major updates include complete parallelization of the code, which results in at least an order of magnitude speed increase[1], using space carving techniques for hull detection [90] and use of robust algebraic reconstruction methods for suppression of image artifacts [68]. Schultze and Censor (University of Haifa), are also improving the performance and increasing the usefulness of the TVS algorithm by using random updates of the perturbation parameter in order to more fully realize its noise-suppressing and feasibility-seeking features. Other algorithms being studied include a modified FBP algorithm which uses "distance-driven-binning" rather than a straight-line projection algorithm in order to achieve a higher-quality image with a fast, FBP-based algorithm [91], and a proton-attenuation algorithm which uses the MLP of the proton and the ratio of protons emitted to detected, similar as in x-ray CT [92].

Presently, the calculation of the MLP is the most computationally expensive component of the reconstruction process. Replacing this process with a simpler algorithm could have significant effects on the speed of reconstruction. Fekete et al. [93] have suggested an algorithm for finding an optimal cubic splined path of the proton. They have shown that this optimal cubic spline is equivalent to the MLP formalism discussed in chapters 2 and 7 in terms of the RMS deviation of the estimated trajectory from the Monte Carlo trajectory (fig. 7.5), but it is much less computationally expensive.

We would also like to be able to update the MLP of the proton "on-the-fly" in

[1]Image reconstruction typically takes several hours depending on the reconstruction parameters selected and the size of the data set.

order to account for inhomogeneities in the phantom. The assumption of a homogeneous water phantom is a reasonable approximation in imaging of soft tissues which have RSP similar to water, however it fails in regions containing materials with high contrast compared to water, as discussed in chapter 7. In chapter 5, however, it was shown that radiographs could be be constructed using only the relative scattering angle of the proton, therefore, one could reasonably use the scattering power of the proton $T \equiv \mathrm{d}\theta^2/\,\mathrm{d}x$ to produce 3D reconstructions of proton scattering power. Scattering power theory was developed first from Moliere's theory of multiple scattering [94, 95] and improved by others [96–99] and was most recently described in detail by Gottschalk [100]. By using alternating reconstructions of RSP and scattering power, the scattering power information from each voxel could be applied in order account for different amounts of scattering in each voxel of the MLP.

Implementation and Workflow

Consideration has also been given to a variety of implementations for reconstruction, in order to achieve a goal of "real-time" preprocessing and reconstruction, which has yet to be realized. Our colleagues at Northern Illinois University (NIU) have focused significant energies on a distributed-GPU implementation of the (original) Penfold code. In addition to multi-threading all components of the original code that run on a CPU, they have also parallelized the code such that it can run on up to 60 GPU nodes on the Gaea cluster at NIU. There are several essential differences between the original version and the distributed GPU implementation. First, while the original version loads all data sequentially onto a single CPU, the NIU version loads the data in many parallel processes (one per MPI). The second key difference is that the NIU version only computes the

MLPs once, and then stores them in memory, eliminating the need to compute them every iteration. In the original version this was not possible due to memory limitations of a single CPU/GPU instance. Finally, the NIU version uses a reconstruction algorithm called component-averaged row projections (CARP) rather than DROP. Using this parallel-distributed implementation, the group at NIU has been able to achieve reconstruction of 500 million proton histories in roughly 3 minutes using 60 compute nodes each with two CPUs and two GPUs.

It is unlikely that a typical cancer center would have the resources to build or maintain a computing cluster of the same size and quality of NIU's Gaea, however, the Amazon web services (AWS) cloud makes GPU-enabled cluster resources accessible to any individual or institution. The NIU group has performed studies comparing the performance of the Gaea cluster with AWS GPU-enabled cluster resources and has found that using the AWS resources, proton CT images could be produced in under ten minutes for less than $10 per image if on-demand provisioning strategies are employed [101]. Even without on-demand provisioning, the cost of producing an image on AWS compared with the overall cost of radiation therapy would be negligible.

The AWS has evidently caught the attention of many researchers in medical imaging and radiotherapy planning due to its flexibility, convenience and power, however it has long been noted that the processing of identifiable health information on commercial clouds is not possible due to data privacy requirements. Presently, compliance and privacy are areas of active research in the cloud community. Some progress has been made recently by cloud providers such as the AWS to offer compliance frameworks [102]. If the current trends persist, cloud computing may become integrated into standard radiotherapy planning and imaging procedures, and may be an essential tool for clinical proton image reconstruction.

8.4 Chapter Summary and Conclusions

A prototype proton CT imaging system, which is capable of tracking upwards of one million individual proton histories per second, and can acquire a 360 degree scan in six minutes or less has been presented in this thesis. The primary limitations of the detector system are due to MCS in the tracking detectors, range straggling in the range detector and inability to resolve coincident proton histories. In this chapter, we have proposed several upgrades for the detector system, which would address these issues.

There is still much to be done on the image reconstruction front in terms of optimizing the image reconstruction process in time, spatial resolution and RSP reconstruction. It is important in the near term that we rigorously define what is meant by "image convergence", and a heuristic to automatically determine this must be implemented in the reconstruction algorithm. In addition, a variety of patterns need to be characterized, such as the relationships between the number of histories in each projection, block size, value of the relaxation parameter, and how these affect image convergence. The interplay between these ought to be well understood so that reconstruction parameters can be efficiently selected.

As we work toward an overhauled version of our present algorithm, we also hope to work alongside other groups studying proton CT image reconstruction in order to compare the performance of our techniques with other algorithms. The data and tools presented in this thesis will be useful for these studies.

One of the most important upcoming tasks will be to experimentally validate the fundamental hypothesis of proton CT: that direct reconstruction of proton RSP will lead to significantly more accurate proton treatment plans than the standard practice of transforming diagnostic x-ray CT scans into RSP maps. There

is especially interest in comparing the accuracy of treatment plans made with proton CT with those made with dual-energy CT (DECT), an imaging modality that utilizes both keV and MeV x-rays, which has recently received a lot of attention from the CT community. This can be accomplished by developing treatment plans for a phantom designed for this purpose using the typical workflow for both modalities. The treatment plan can then be delivered to the phantom and the results can be evaluated by comparing the measured dose distribution to that of the treatment plan. With this result, we hope that proton CT can make its way into the treatment room and become a more mainstream imaging modality.

Bibliography

[1] A. A. Mustafa and D. F. Jackson. "The relation between x-ray CT numbers and charged particle stopping powers and its significance for radiotherapy treatment planning". In: *Physics in medicine and biology* 28.2 (1983), p. 169.

[2] M. Yang et al. "Comprehensive analysis of proton range uncertainties related to patient stopping-power-ratio estimation using the stoichiometric calibration". In: *Physics in medicine and biology* 57.13 (2012), p. 4095.

[3] Tami Freeman. "Will protons gradually replace photons". In: *Medical Physics Web* 22 (2012).

[4] D. Wang, T. R. Mackie, and W. A. Tomé. "Bragg peak prediction from quantitative proton computed tomography using different path estimates". In: *Physics in medicine and biology* 56.3 (2011), p. 587.

[5] Robert R. Wilson. "Radiological use of fast protons". In: *Radiology* 47.5 (1946), pp. 487–491.

[6] C. A. Tobias et al. "Pituitary irradiation with high-energy proton beams a preliminary report". In: *Cancer Research* 18.2 (1958), pp. 121–134.

[7] J. M. Slater et al. "The proton treatment center at Loma Linda University Medical Center: rationale for and description of its development". In: *International Journal of Radiation Oncology* Biology* Physics* 22.2 (1992), pp. 383–389.

[8] Particle Therapy Co-Operative Group. *Particle therapy facilities in operation*. 2016. URL: `http://www.ptcog.ch/index.php/facilities-in-operation` (visited on 01/13/2016).

[9] The Clatterbridge Cancer Center. *Proton therapy*. Jan. 2016. URL: `http://www.clatterbridgecc.nhs.uk/patients/treatment-and-support/proton-therapy`.

[10] Proton Therapy Center. *Proton Therapy Around the World*. URL: `http://www.proton-cancer-treatment.com/proton-therapy/proton-therapy-around-the-world/` (visited on 01/13/2016).

[11] Allan M. Cormack. “Representation of a function by its line integrals, with some radiological applications”. In: *Journal of applied physics* 34.9 (1963), pp. 2722–2727.

[12] A. M. Koehler. “Proton radiography”. In: *Science* 160 (1968), pp. 303–304.

[13] V. W. Steward and A. M. Koehler. “Proton beam radiography in tumor detection”. In: *Science* 179.4076 (1973), pp. 913–914.

[14] V. W. Steward and A. M. Koehler. “Proton radiographic detection of strokes”. In: *Nature* 245 (1973), pp. 38–40.

[15] V. W. Steward and A. M. Koehler. “Proton radiography in the diagnosis of breast carcinoma.” In: *Radiology* 110.1 (1974), pp. 217–221.

[16] Michael Goitein. “Three-dimensional density reconstruction from a series of two-dimensional projections”. In: *Nuclear Instruments and Methods* 101.3 (1972), pp. 509–518.

[17] K. M. Crowe et al. “Axial scanning with 900 MeV alpha particles”. In: *IEEE Transactions on Nuclear Science* 3.22 (1975), pp. 1752–1754.

[18] A. M. Cormack and A. M. Koehler. “Quantitative proton tomography: preliminary experiments”. In: *Physics in medicine and biology* 21.4 (1976), p. 560.

[19] Kenneth M. Hanson. “Proton computed tomography”. In: *Nuclear Science, IEEE Transactions on* 26.1 (1979), pp. 1635–1640.

[20] K. M. Hanson et al. “Computed tomography using proton energy loss”. In: *Physics in medicine and biology* 26.6 (1981), p. 965.

[21] K. M. Hanson et al. “Proton computed tomography of human specimens”. In: *Physics in medicine and biology* 27.1 (1982), pp. 25–36.

[22] P. Pemler et al. “A detector system for proton radiography on the gantry of the Paul-Scherrer-Institute”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 432.2 (1999), pp. 483–495.

[23] Uwe Schneider and Eros Pedroni. “Proton radiography as a tool for quality control in proton therapy”. In: *Medical physics* 22.4 (1995), pp. 353–363.

[24] U. Schneider et al. “First proton radiography of an animal patient”. In: *Medical physics* 31.5 (2004), pp. 1046–1051.

[25] G. Poludniowski, N. M. Allinson, and P. M. Evans. “Proton radiography and tomography with application to proton therapy”. In: *The British journal of radiology* 88 (2015).

[26] D. C. Williams. “The most likely path of an energetic charged particle through a uniform medium”. In: *Physics in medicine and biology* 49.13 (2004), p. 2899.

[27] J. Beringer et al. “Review of Particle Physics”. In: *Phys. Rev. D* 86 (1 July 2012), p. 010001. DOI: 10.1103/PhysRevD.86.010001. URL: http://link.aps.org/doi/10.1103/PhysRevD.86.010001.

[28] R. W. Schulte et al. “A maximum likelihood proton path formalism for application in proton computed tomography”. In: *Medical physics* 35.11 (2008), pp. 4849–4856.

[29] C. Tschalär. “Straggling distributions of large energy losses”. In: *Nuclear Instruments and Methods* 61.2 (1968), pp. 141–156.

[30] C. Tschalär. *Straggling distributions of extremely large energy losses by heavy particles*. Tech. rep. Rutherford High Energy Lab., Chilton (England), 1968.

[31] R. W. Schulte et al. “Density resolution of proton computed tomography”. In: *Medical physics* 32.4 (2005), pp. 1035–1046.

[32] M. Goitein and M. Jermann. “The relative costs of proton and X-ray radiation therapy”. In: *Clinical Oncology* 15.1 (2003), S37–S50.

[33] M. Lodge et al. “A systematic literature review of the clinical and cost-effectiveness of hadron therapy in cancer”. In: *Radiotherapy and Oncology* 83.2 (2007), pp. 110–122.

[34] J. Lundkvist et al. “Cost-effectiveness of proton radiation in the treatment of childhood medulloblastoma”. In: *Cancer* 103.4 (2005), pp. 793–801.

[35] J. Lundkvist et al. “Economic evaluation of proton radiation therapy in the treatment of breast cancer”. In: *Radiotherapy and oncology* 75.2 (2005), pp. 179–185.

[36] M. Pijls-Johannesma, P. Pommier, and Y. Lievens. “Cost-effectiveness of particle therapy: current evidence and future needs”. In: *Radiotherapy and oncology* 89.2 (2008), pp. 127–134.

[37] J. Unkelbach. *Proton therapy – close to becoming mainstream*. World Congress of Medical Physics and Bioengineering. Toronto, 2015.

[38] M.J. Berger et al. *ESTAR, PSTAR, and ASTAR: Computer Programs for Calculating Stopping-Power and Range Tables for Electrons, Protons, and Helium Ions (version 1.2.3)*. 2005. URL: http://physics.nist.gov/Star (visited on 01/12/2016).

[39] Michael F. L’Annunziata. *Radioactivity: Introduction and history*. Elsevier, 2007.

[40] R. F. Hurley et al. “Water-equivalent path length calibration of a prototype proton CT scanner”. In: *Medical physics* 39.5 (2012), pp. 2438–2446.

[41] U. Schneider and E. Pedroni. “Multiple Coulomb scattering and spatial resolution in proton radiography”. In: *Medical physics* 21.11 (1994), pp. 1657–1663.

[42] Mara Bruzzi et al. “Prototype tracking studies for proton CT”. In: *Nuclear Science, IEEE Transactions on* 54.1 (2007), pp. 140–145.

[43] Stephen M. Stigler. “Thomas Bayes’s bayesian inference”. In: *Journal of the Royal Statistical Society. Series A (General)* (1982), pp. 250–258.

[44] Leonard Eyges. “Multiple scattering with energy loss”. In: *Physical Review* 74.10 (1948), p. 1534.

[45] Jiang Hsieh. “Computed tomography”. In: *Bellingham: SPIE* (2003), pp. 1–12.

[46] Gengsheng Lawrence Zeng. *Medical image reconstruction*. Springer, 2010.

[47] Avinash C. Kak and Malcolm Slaney. *Principles of computerized tomographic imaging*. Society for Industrial and Applied Mathematics, 2001.

[48] Yair Censor and Stavros Andrea Zenios. *Parallel optimization: Theory, algorithms, and applications*. Oxford University Press, 1997.

[49] Gabor T. Herman. *Fundamentals of computerized tomography: image reconstruction from projections*. Springer Science & Business Media, 2009.

[50] Scott N. Penfold. “Image Reconstruction and Monte Carlo Simulations in the Development of Proton Computed Tomography for Applications in Proton Radiation Therapy”. PhD thesis. University of Wollongong, 2010.

[51] Y. Saraya et al. “Study of spatial resolution of proton computed tomography using a silicon strip detector”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 735 (2014), pp. 485–489.

[52] M. Scaringella et al. “A proton Computed Tomography based medical imaging system”. In: *Journal of Instrumentation* 9.12 (2014), p. C12009.

[53] J. T. Taylor et al. “Proton tracking for medical imaging and dosimetry”. In: *Journal of Instrumentation* 10.02 (2015), p. C02015.

[54] Robert P. Johnson et al. “A fast experimental scanner for proton CT: Technical performance and first experience with phantom scans”. In: *Transactions on Nuclear Science* 63.1 (2015), pp. 52–60.

[55] Robert P. Johnson et al. “Tracker Readout ASIC for Proton Computed Tomography Data Acquisition”. In: *Nuclear Science, IEEE Transactions on* 60.5 (2013), pp. 3262–3269.

[56] A. Artikova et al. “Properties of the Ukraine polystyrene-based plastic scintillator UPS 923A”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 555 (2005), pp. 125–131.

[57] V. A. Bashkirov et al. “Novel scintillation detector design and performance for proton radiography and computed tomography”. In: *Medical Physics* 43.2 (2016), pp. 664–674.

[58] A. Zatserklyaniy et al. “Track Reconstruction with the Silicon Strip Tracker of the Proton CT Phase 2 Scanner”. In: *IEEE 2014 NSS/MIC Conference Record*. Nov. 2014.

[59] V. Giacometti et al. “Geant4-based simulation platform for a prototype proton CT scanner”. Currently unpublished. 2016.

[60] S. Chauvie et al. “Geant4 low energy electromagnetic physics”. In: *Nuclear Science Symposium Conference Record, 2004 IEEE*. Vol. 3. IEEE. 2004, pp. 1881–1885.

[61] Uwe Schneider et al. “Patient specific optimization of the relation between CT-Hounsfield units and proton stopping power with proton radiography”. In: *Medical physics* 32.1 (2005), pp. 195–199.

[62] R. Schulte et al. “Conceptual design of a proton computed tomography system for applications in proton radiation therapy”. In: *Nuclear Science, IEEE Transactions on* 51.3 (2004), pp. 866–872.

[63] Nicolas Depauw and Joao Seco. “Sensitivity study of proton radiography and comparison with kV and MV x-ray imaging using GEANT4 Monte Carlo simulations”. In: *Physics in medicine and biology* 56.8 (2011), p. 2407.

[64] M. G. Stabin. *RADAR Medical Procedure Radiation Dose Calculator*. 2012.

[65] S. N. Penfold et al. “Total variation superiorization schemes in proton computed tomography image reconstruction”. In: *Medical physics* 37.11 (2010), pp. 5887–5895.

[66] Yair Censor et al. “On diagonally relaxed orthogonal projection methods”. In: *SIAM Journal on Scientific Computing* 30.1 (2008), pp. 473–504.

[67] Theodore Geoghegan. “Interface events within the calorimeter of the proton computed tomography head scanner”. B.S. Thesis. University of California, Santa Cruz, 2015.

[68] P. Karbasi et al. “Incorporating robustness in diagonally-relaxed orthogonal projections method for proton computed tomography”. In: *Proceedings of the 22nd Annual Symposium on Room-Temperature Semiconducctor X-Ray and Gamma-Ray Detectors and Medical Imaging Conference*. San Diego, CA, USA: IEEE, 2015.

[69] Scott Penfold et al. “Block-iterative and string-averaging projection algorithms in proton computed tomography image reconstruction”. In: (2010).

[70] Barbara Schaffner, Eros Pedroni, and Antony Lomax. “Dose calculation models for proton treatment planning using a dynamic beam delivery system: an attempt to include density heterogeneity effects in the analytical dose calculation”. In: *Physics in medicine and biology* 44.1 (1999), p. 27.

[71] Joao Seco et al. “Characterizing the modulation transfer function (MTF) of proton/carbon radiography using Monte Carlo simulations”. In: *Medical physics* 40.9 (2013), p. 091717.

[72] Uwe Schneider et al. “Spatial resolution of proton tomography: methods, initial phase space and object thickness”. In: *Zeitschrift für Medizinische Physik* 22.2 (2012), pp. 100–108.

[73] Bob Atkins. *Modulation Transfer Function – what is it and why does it matter?* 2007. URL: `http://photo.net/learn/optics/mtf/` (visited on 01/29/2016).

[74] L. A. Feldkamp, L. C. Davis, and J. W. Kress. “Practical cone-beam algorithm”. In: *JOSA A* 1.6 (1984), pp. 612–619.

[75] P. F. Judy. “The line spread function and modulation transfer function of a computed tomographic scanner”. In: *Medical physics* 3.4 (1976), pp. 233–236.

[76] H. Fujita et al. “A simple method for determining the modulation transfer function in digital radiography”. In: *Medical Imaging, IEEE Transactions on* 11.1 (1992), pp. 34–39.

[77] I. Mori and Y. Machida. “Deriving the modulation transfer function of CT from extremely noisy edge profiles”. In: *Radiological physics and technology* 2.1 (2009), pp. 22–32.

[78] H. Watanabe, E. Honda, and T. Kurabayashi. “Modulation transfer function evaluation of cone beam computed tomography for dental use with the oversampling method”. In: *Dentomaxillofacial Radiology* (2010).

[79] Python Software Foundation, ed. *Python Language Reference, version 2.7*. `http://www.python.org`. 2016.

[80] S. Agostinelli et al. “GEANT4: A Simulation toolkit”. In: *Nucl. Instrum. Meth.* A506 (2003), pp. 250–303. DOI: `10.1016/S0168-9002(03)01368-8`.

[81] Joseph Perl et al., eds. *User Guide for TOPAS Version 2.0*. `http://www.topasmc.org/user-guide`. 2015.

[82] Joseph Perl et al. “TOPAS: an innovative proton Monte Carlo platform for research and clinical applications”. In: *Medical physics* 39.11 (2012), pp. 6818–6837.

[83] W. B. Atwood et al. “The Large Area Telescope on the Fermi gamma-ray space telescope mission”. In: *The Astrophysical Journal* 697.2 (2009), p. 1071.

[84] Y. Takada et al. “Proton computed tomography with a 250 MeV pulsed beam”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 273.1 (1988), pp. 410–422.

[85] H. F.-W. Sadrozinski et al. “Ultra-fast silicon detectors”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 730 (2013), pp. 226–231.

[86] C. Joram, G. Haefeli, and B. Leverington. “Scintillating Fibre Tracking at High Luminosity Colliders”. In: *Journal of Instrumentation* 10.08 (2015), p. C08005.

[87] Ivan Perić et al. “High-voltage pixel detectors in commercial CMOS technologies for ATLAS, CLIC and Mu3e experiments”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 731 (2013), pp. 131–136.

[88] R. Ball et al. “Development of a plasma panel radiation detector”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 764 (2014), pp. 122–132.

[89] Jurriaan Schmitz. “The Microstrip Gas Counter and its application in the ATLAS inner tracker”. PhD thesis. Amsterdam U., 1994.

[90] Blake Schultze et al. “Performance of hull-detection algorithms for proton computed tomography reconstruction”. In: *Contemporary Mathematics* 636 (2015), pp. 211–224.

[91] Simon Rit et al. “Filtered backprojection proton CT reconstruction along most likely paths”. In: *Medical physics* 40.3 (2013), p. 031103.

[92] C. T. Quiñones, J. M. Létang, and S. Rit. “Filtered back-projection reconstruction for attenuation proton CT along most likely paths”. In: *Physics in Medicine and Biology* 61.9 (2016), p. 3258.

[93] C.-A. C. Fekete et al. “Developing a phenomenological model of the proton trajectory within a heterogeneous medium required for proton imaging”. In: *Physics in medicine and biology* 60.13 (2015), p. 5071.

[94] G. Moliere. “Theorie der Streuung schneller geladenen Teilchen II Mehrfach- und Vielfachstreuung”. In: *Z. Naturforschg* 3a (1948), pp. 78–97.

[95] H. A. Bethe. “Moliere’s theory of multiple scattering”. In: *Physical Review* 89.6 (1953), p. 1256.

[96] U. Fano. “Inelastic collisions and the Moliere theory of multiple scattering”. In: *Physical Review* 93.1 (1954), p. 117.

[97] A. O. Hanson et al. “Measurement of multiple scattering of 15.7-MeV electrons”. In: *Physical Review* 84.4 (1951), p. 634.

[98] B. P. Nigam, M. K. Sundaresan, and T. Wu. “Theory of multiple scattering: second Born approximation and corrections to Moliere’s work”. In: *Physical Review* 115.3 (1959), p. 491.

[99] W. T. Scott. “The theory of small-angle multiple scattering of fast charged particles”. In: *Reviews of modern physics* 35.2 (1963), p. 231.

[100] Bernard Gottschalk. “On the scattering power of radiotherapy protons”. In: *arXiv preprint arXiv:0908.1413* (2009).

[101] Ryan Chard et al. “Scalable pCT image reconstruction delivered as a cloud service”. In: (2014).

[102] Lifei Wei et al. “Security and privacy for storage and computation in cloud computing”. In: *Information Sciences* 258 (2014), pp. 371–386.