

Statistically-motivated, Iterative Image Reconstruction Algorithm for Proton Computed Tomography Applications

MSc Thesis

Ákos Sudár

Medical Physicist, Faculty of Nuclear Techniques,
Budapest University of Technology and Economics,
Research Engineer, Wigner Research Centre for Physics



Supervisor:

Mónika Varga-Kőfaragó, Ph.D.

Researcher, Hadron Physics Research Group, Department of High Energy Particle Physics,
Institute for Particle and Nuclear Physics, Wigner Research Centre for Physics

Consultants:

Dr. Dávid Légrády

Associate professor, Faculty of Nuclear Techniques,
Budapest University of Technology and Economics,

Dr. Gergely Gábor Barnaföldi

Group leader, Heavy-ion Physics Research Group, Department of Theoretical Physics,
Institute for Particle and Nuclear Physics, Wigner Research Centre for Physics

Budapest, 2021

Abstract

Modern proton Computed Tomography (pCT) images are usually reconstructed by algebraic reconstruction techniques (ART). The Kaczmarz-method and its variations are among the most widely used methods, which are iterative solution techniques for linear problems with sparse matrices. It is an interesting question whether statistically-motivated iterations, which have been successfully used for emission tomography, can be applied to reconstruct the novel technology of pCT images as well.

In my research, I developed a method for pCT image reconstruction, based on the Richardson–Lucy deconvolution. It treats the problem as a statistically-motivated fixed-point iteration. I implemented this algorithm as a parallel code to GPU, with spline-based trajectory calculation and on-the-fly system matrix generation. My results presented that the method works well, and it can be successfully applied in pCT applications, such as in the detector R&D of the Bergen pCT Collaboration.

Contents

1	Introduction	3
2	The Proton CT	5
2.1	The Bergen pCT Collaboration	5
2.2	My Contribution to the Bergen pCT project	8
3	Image Reconstruction Methods	10
3.1	Introduction to Iterative Image Reconstruction Techniques	10
3.2	Diagonally Relaxed Orthogonal Projection Method	11
3.3	Total Variation Superiorization of DROP Method	12
3.4	The Richardson – Lucy Algorithm	14
4	Imaging Model	15
4.1	The Physical Background of the Model	15
4.2	Modeling of the Imaging	17
4.3	The Numerical Model	19
5	Results on the Richardson – Lucy Algorithm for pCT	22
5.1	Convergence	23
5.2	Spatial Resolution	27
5.3	Relative Stopping Power Accuracy	30
5.4	The Longitudinal Position Error of the Bragg Peak	33
6	Results on PD-based Interaction Calculation	36
6.1	Test of the Convergence	36
6.2	Average Relative Stopping Power Accuracy	38
7	Summary	42
	Bibliography	44

1 Introduction

Cancer has one of the highest mortality compared to other diseases, especially in developed countries, so the need for effective treatments and therapies is increasingly important. One of the widely-applied treatments for cancer is radiotherapy. Recently there is growing interest for a novel technology, hadron therapy, what uses mainly proton beams for the therapy. Applying high-energy protons, it is possible to focus the delivered dose more precisely on the tumor than with the standard radiotherapy methods using MeV-energy photons.

One of the main directions in the development of proton therapy is to increase the accuracy of the treatment planning, which is important in reducing the dose and in focusing the damage on the target volume. The limitation of the treatment planning is the accuracy of the relative stopping power (RSP) distribution map of the patient. Today these maps are generally converted from X-ray CT images with the use of conversion tables between Hounsfield unit (HU) and RSP. This conversion is not trivial in the typical range of Hounsfield units of the human body, therefore the conversion can significantly increase the error of the final RSP map, which can cause up to 1.7% statistical error. This error can be reduced to 0.5 % with the use of proton CT (pCT) imaging [1].

The pCT scanner designs usually use movable treatment beams as a proton source and a detector system to measure the protons before and after the patient. A typical detector has two functionalities:

1. The first is to measure the remaining energy of the protons behind the patient. If it is built of energy sensitive detectors it is called a proton calorimeter (as it measures the energy of protons), while if it is built of alternating detector and energy absorber layers it is called a range counter (as it measures the remaining water equivalent range of the protons).
2. The second functionality is the measurement of the incoming and outgoing directions of the protons, which information is necessary for the reconstruction of the RSP map. This is done by pairs of tracking detectors. The tracker detector pair in front of the

patient is called the forward tracking detector, while the one behind the patient is called the backward tracking detector. Sometimes the forward tracking detector is left out. In one hand this concept is advantageous because one event can contain higher number of proton histories compared to a concept with a forward pair, thanks to the lack of pairing problem of the measured trajectories by the forward and the backward pairs. This design is also more advantageous from the point of view of the application in a treatment facility as there is no detector to interact with the beam during the treatment. On the other hand this concept is less accurate in terms of investigating the trajectory of protons in the patient than the concept, which contains both a forward and a backward tracker. Because of its advantages the concept without a forward tracker will be investigated in this work.

The Bergen Proton CT Collaboration is dedicated to the development of a pCT detector without a forward detector to profit from the advantages of this concept [2]. The group balances the loss of accuracy with a very accurate backward tracking detector. As a member of the Bergen pCT group I dedicated my work to a testing of an image reconstruction technique which is new in the field of proton image reconstruction.

2 The Proton CT

The idea of a proton CT was born at the same time with the idea of X-ray CT, as both of them were suggested by A.M. Cormack in 1963. One of the original motivations behind both CT imaging techniques was the dose planning for radiation therapy. Cormack realised the importance of pCT images for proton treatment planning to focus the Bragg peak into the tumor [3]. The first reasonable CT image was reconstructed by Cormack in 1964, it was about a phantom made of aluminum and Lucite cylinders and rings [4].

Motivation of the pCT is to increase the accuracy of dose delivering during proton therapy. The number of proton treatment centers is rapidly increasing [5] thanks to the fact that hadron beams deliver more focused dose distribution compared with MeV-photon therapy [6, 7]. The need for more accurate proton treatment is evolving in parallel with the number of treatment facilities, as more and more hospitals could gain from a novel device, technology or treatment method. There are a lot of possible directions for improvement in this field, including but not limited to patient position and movement monitoring, use of Monte Carlo simulations in dose planning [8], pCT imaging for the more accurate dose planning [3, 5, 9] and *in vivo* dosimetry for monitoring the dose distribution during and after the treatment [8, 10, 11].

2.1 The Bergen pCT Collaboration

The Bergen pCT Collaboration is an international collaboration based at the University of Bergen, Bergen, Norway [2]. It was established to develop a pCT system based on the detector technologies used at high-energy particle detectors at CERN or in other international collaborations. This applied study aims to reach the clinical testing phase, which has never been reached before. From the beginning, the Wigner Research Centre for Physics has been contributing to this effort, and I – as a member of the group – could personally participate in the development of the Bergen pCT detector prototype.

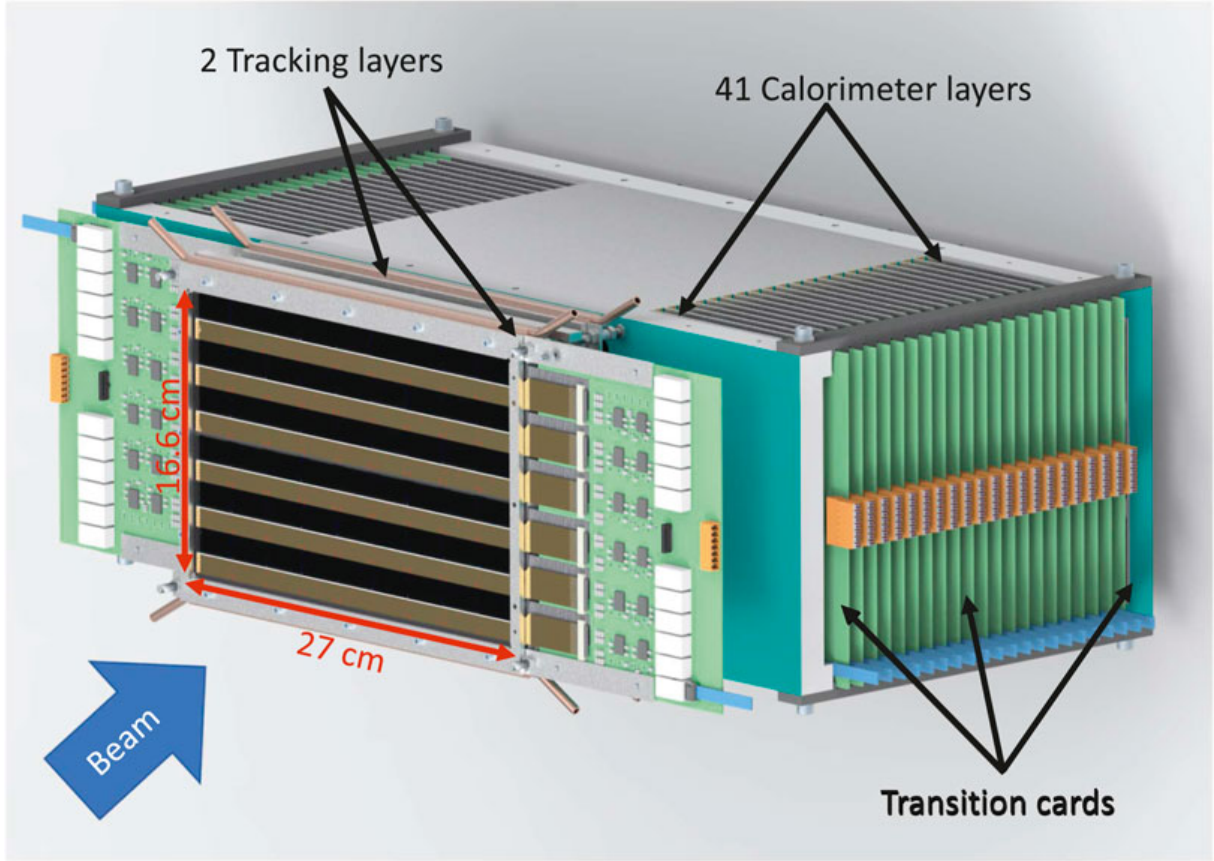


Figure 1: The design of the Bergen pCT detector system. The beam comes from the left hand side, the tracker pair is visible in the foreground, the 41 layers of the Digital Tracking Calorimeter (DTC) is visible in the background and the transition cards of the readout electronics are shown on the sides [2].

The main concept of the pCT detector system is to use silicon pixel tracker detectors for the measurement of the direction and the remaining energy of the protons. The challenge of this is that the energy range here is a few magnitudes lower, than what this detector was originally designed for in high-energy physics. The direction of the protons is measured only after the patient, which makes it possible to measure 100 protons in each frame in the area of a pencil beam with 7 mm full width at half maximum (FWHM). The lack of information coming from the lack of a forward tracker is partially compensated by the very accurate rear tracker [12]. The accuracy of the rear tracker pair comes from the high spatial resolution of

the tracker detector ($\sim 5 \mu\text{m}$) and the low material budget of the tracker layers: the thickness of the detector and the support structure are $50 \mu\text{m}$ and $200 \mu\text{m}$, respectively. The energy measurement is based on a range counter design, where alternating tracker and absorber layers are implemented. The number of firing pixels in the tracker detector, when a proton goes through it, contains some information about the deposited energy in the tracker detector. The energy measurement, based on the fit of a Bragg curve to these energy deposits, makes possible a more accurate energy measurement, than the number of traversed layers would allow [13].

The pCT detector system is based on the ALPIDE chip, which is a monolithic active pixel sensor (MAPS) type silicon pixel detector [14]. This contains the sensitive sensors together with the readout electronics on the same silicon layer. The ALPIDE was developed for the Inner Tracking System (ITS) of ‘A Large Ion Collider Experiment’ (ALICE), which is one of the four large experimental facilities of the Large Hadron Collider (LHC) at the European Organization for Nuclear Research (CERN), Geneva, Switzerland.

The Bergen pCT Collaboration had a previous range counter design, which was the proof-of-concept Digital Tracking Calorimeter (DTC) in the pCT application. The measurements confirmed the concept as it reached 4% accuracy in terms of the measurement of the remaining water equivalent thickness (WET) of the protons. The same concept could reach 10^6 proton per second performance based on MC simulations, which is in the higher end of the sensitivity and time resolution of the pCT prototypes [13]. The group continued the work with the design study of a next, more optimized pCT detector [9]. The new concept contains two tracker layers and 41 alternating aluminium absorbers and tracker detector layers, which is an optimum between accuracy and cost. The detector is designed to measure 10^7 individual proton tracks in each second with the use of a 7 mm FWHM pencil beam.

The research and development work is still in progress, and the new prototype is in the construction phase now [2], while a detailed investigation of the detector concept is ongoing. In parallel, the track reconstruction and data analysis software development has started, in which I am also participating.

2.2 My Contribution to the Bergen pCT project

I was actively involved in the measurement of the relation between the particle energy (which determines the expected energy deposition) and the number of firing pixels. We took the data in two test beam measurements with proton, helium, and carbon ion beams in the Heidelberg Ion-beam Therapy Center (HIT). In Figure 2 the points represent the mean value of a normal distribution fitted to the measured cluster size distributions of the silicon pixel detector. The bars are for the standard deviation of the fitted distribution, so these stand for the representation of the width of the distribution and not for the uncertainty of the mean value. We found a clear relation between the cluster size and the deposited energy shown by the following equation [15, 2]:

$$N = 2.40 \cdot \Delta E^{0.338}, \quad (1)$$

where N is the number of firing pixels (called cluster size) and ΔE is the expected deposited energy. The constant factor 0.338 was determined by fitting the measured data as presented in Figure 2. My main contribution was the data analysis of the measurement, which resulted in the data points and appeared in publications of the collaboration [15, 2] and in my previous OTDK thesis [16].

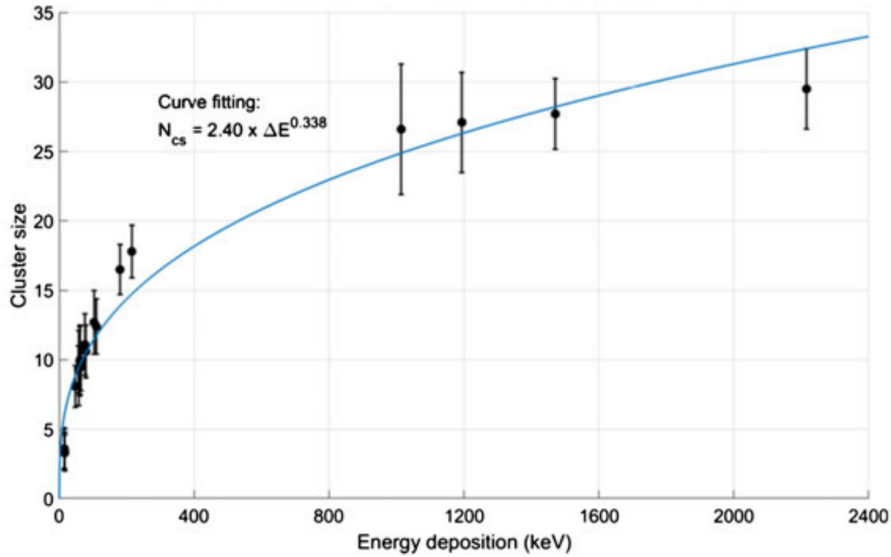


Figure 2: The measured cluster size as a function of the expected deposited energy [2].

My further contribution is the study of two cooling system concepts for the range detector. I used analytical calculations and finite element simulations for modeling the temperature distribution inside the range detector. My results appears in one of the status reports of the collaboration [2], in one of my previous TDK thesis [17], and in my BSc thesis [18].

My current contribution is the investigation of the application of statistically motivated iterative methods for image reconstruction. This has been first discussed in this present thesis. So far there is no such analysis known publicly available in the international literature.

3 Image Reconstruction Methods

During proton CT imaging, we collect information about the relative stopping power (RSP) of the phantom or patient by the trajectories. The role of image reconstruction is to determine the original RSP distribution of the object. There are two main directions to reconstruct CT images. The first is based on integral transformations. The first representative of this kind is the Radon and the Inverse Radon transformation, which was published by Johann K. A. Radon in 1917 [19]. The CT reconstruction techniques introduced by A. M. Cormack [3, 4] belong to this group. The second main direction is to model the problem as a linear equation system, and to solve it with the use of iterative algorithms. This latter direction was introduced and called Algebraic Reconstruction Technique (ART) in image reconstruction by Gordon, Bender and Herman in 1970 [20]. The Richardson–Lucy deconvolution [21, 22] used in this work is also an iterative reconstruction technique that can be applied for the reconstruction of CT images. This algorithm is mainly used in emission tomography, and is first applied for pCT image reconstruction in one of my previous works [23].

3.1 Introduction to Iterative Image Reconstruction Techniques

If we reconstruct the image in discrete voxels then the image reconstruction can be written in the following general algebraic form:

$$\mathbf{y} = \mathbf{A} \cdot \mathbf{x}, \tag{2}$$

where $\mathbf{y}$ is an m -dimensional vector (typically $10^8 - 10^9$ elements), which contains the water equivalent path length (WEPL) reduction of the protons in the reconstruction area, $\mathbf{x}$ is an n -dimensional vector (typically $10^5 - 10^7$ elements) containing the RSP of the voxels, and finally $\mathbf{A}$ is an $n \times m$ matrix (also called system matrix, typically $10^{13} - 10^{16}$ elements) which contain the interaction coefficient between protons and voxels. The information in matrix $\mathbf{A}$ can be described as the length of the proton path in the voxel. In practice, for the real cases n is larger than m . The goal of the image reconstruction is to investigate the values of $\mathbf{x}$ with the knowledge of $\mathbf{y}$ and $\mathbf{A}$. Typically, to obtain high-enough resolution, the size of the matrix is too

large to use standard matrix inversion techniques e.g. the pseudo inverse method. The matrix however is sparse, which property can be used to find an effective way for the solution of the linear problem. In the literature, orthogonal projection based iterative algorithms are widely used [5, 20, 24, 25, 26]. A short review is given in the following subsections.

3.2 Diagonally Relaxed Orthogonal Projection Method

In the literature, usually different variations of the method appears for pCT image reconstruction. The predecessor of these methods in linear algebra is the Kaczmarz [27] iteration, which was rediscovered in the field of image reconstruction by Richard Gordon, Robert Bender, and Gábor Herman under the name of Algebraic Reconstruction Technique (ART) [20]:

$$\rho_{i,j}^{q+1} = \max \left[\rho_{i,j}^q + \left(R_{k,\theta} - R_{k,\theta}^q \right) / N_{k,\theta}, 0 \right], \quad (3)$$

for (i, j) in ray (k, θ) , where q stands for the number of iteration, (i, j) is the index of the voxels, (k, θ) is the index of the k^{th} ray in projection θ , ρ is the density matrix of the voxels, R is the matrix of the integrated density along the rays:

$$R_{k,\theta}^q = \sum_{(i,j) \text{ in ray } (k,\theta)} \rho_{i,j}^q \quad (4)$$

and finally $N_{k,\theta}$ is the number of voxels crossed by ray (k, θ) .

A more recent, diagonally relaxed orthogonal projection (DROP) iteration scheme is also known [25]:

$$x^{k+1} = x^k + \frac{\lambda_k}{m} \sum_{i=1}^m \frac{b_i - \langle A_i, x^k \rangle}{\|A_i\|_2^2} A_i, \quad (5)$$

where k is the number of iteration, x is the vector of the voxel values, λ^k are the relaxation parameters, m is the number of trajectories, i is the index of trajectories, b is the vector of integrated values along trajectories, A is the interaction coefficient between the trajectories and the voxels and $\|\dots\|_2$ is the Euclidean norm. The DROP iteration defined in equation (5) has several advantages over the original ART iteration defined in equation (3), since it can take into account different interaction coefficients between the voxels and the trajectories. One can reach better convergence with the use of relaxation parameters, which can be changed

during the iteration and one can easily parallelise the evaluation of equation (5) thanks to the independent elements of the sum.

3.3 Total Variation Superiorization of DROP Method

This subsection is a brief summary of the work of Penfold *et al.* [24]. The main idea of the total variation superiorization of DROP (TVS-DROP) method is to try to find a solution, which minimizes better a merit function:

$$x^{k+1} = \mathbf{P}_{t(k)}(x^k) = x^k + \lambda_k U_{t(k)} \sum_{i \in I_{t(k)}} \frac{b_i - \langle a_i, x^k \rangle}{\|a_i\|_2^2} a_i, \quad (6)$$

where $U_{t(k)} = \text{diag}\left(\left(1, 1/h_j^t\right)\right)$, where h_j^t is the number of proton trajectories which intersects the j^{th} voxel in the t^{th} block.

The merit function can be given as

$$\phi(p^k) = \sum_{g=1}^{J-1} \sum_{l=1}^{J-1} \sqrt{\left(p_{g+1,l}^k - p_{g,l}^k\right)^2 + \left(p_{g,l+1}^k - p_{g,l}^k\right)^2}. \quad (7)$$

One can see the schematic steps of the total variation superiorization of DROP Method in Algorithm 1 [24].

Algorithm 1 Total Variation Superiorization of Diagonally Relaxed Orthogonal Projections

```
1: set  $k = 0$ , set  $x^k = x_{FBP}$  the initial Filtered Back Projection reconstruction, and  $\beta_k = 1$ 
2: for 10 cycles do
3:   set  $s$  to a subgradient of  $\phi$  at  $x^k$ 
4:   if  $\|s\| > 0$  then
5:     set  $v^k = -s/\|s\|$ 
6:   else
7:     set  $v^k = s$ 
8:   end if
9:   set  $\text{continue} = \text{true}$ 
10:  while  $\text{continue}$  do
11:    set  $y^k = x^k + \beta_k v^k$ 
12:    calculate the merit function (total variation)
13:    if  $\phi(y^k) \leq \phi(x^k)$  then
14:      apply sequentially  $M$  times the projection operator  $P_{t(k)}$  to  $y^k$ 
15:      calculate the feasibility proximity with using values from all earlier  $M$  blocks
16:      if  $Pr(\mathbf{P}_M y) < Pr(x^k)$  then
17:        set  $x^{k+1} = \mathbf{P}_M y$ 
18:        set  $\text{continue} = \text{false}$ 
19:      else
20:        set  $\beta_k = \beta_k/2$ 
21:      end if
22:    else
23:      set  $\beta_k = \beta_k/2$ 
24:    end if
25:  end while
26: end for
```

3.4 The Richardson – Lucy Algorithm

There is a different solution strategy which approaches the problem from the probability theory direction called Maximum Likelihood – Expectation Maximization (ML-EM) method. This theory was successful in the case of emission tomography, and it results also an iteration cycle for the image reconstruction called Richardson – Lucy deconvolution [21, 22]. This iteration cycle can be deduced from ML-EM theory in case of emission tomography, but is known as a fixed-point iteration as well. The iterative solution is based on the formula,

$$x_i^{k+1} = x_i^k \frac{1}{\sum_j A_{i,j}} \sum_j \frac{y_j}{\sum_l A_{l,j} x_l^k} A_{i,j} , \quad (8)$$

for every i in N , where N is the length of vector x , which contains the RSP of the voxels, k is the number of iterations, matrix A contains the interaction coefficients between the proton trajectories and the voxels, j is the index of the trajectories in M , which is the number of the trajectories, y contains the integrated RSP along the trajectories, which is equivalent with the WEPL reduction of the protons travelling along the trajectories. The $y_j / \sum_l A_{l,j} x_l^k$ term is usually called Hadamard ratio, and it represents the ratio of the integrated RSP along the proton path and its estimate based on the voxel values calculated in the previous iteration.

The ML-EM problem for proton CT imaging has not been solved yet. As long as the iteration is not introduced, the Richardson – Lucy deconvolution can be used as a fixed-point iteration for pCT image reconstruction as a linear problem. Therefore, my work is the first step to use statistical based iterations in pCT image reconstruction, but it can motivate further developments in this research direction.

4 Imaging Model

In this section, I will introduce the main details of the pCT image reconstruction model developed in this work. The first part of this section will summarize the details of the model of the physics implementation in my work. I applied the Richardson–Lucy algorithm to restore the images, in which the elements of the system matrix, that represent the physical interactions between the protons and the voxels, was set based on this model. The second part of this section details the computational model of the problem, which contains the on-the-fly calculation of the system matrix elements and the Richardson–Lucy iteration itself. In this model, I implemented on-the-fly matrix calculation, because even the zero suppressed system matrix could reach the size of 10^{11} elements. If each element requires to store a single precision value and a double precision matrix index, the matrix can be stored in $12 \times 10^{11} = 1.2 \times 10^{12}$ bytes ≈ 1.2 TB. It is expensive to use a so large memory volume in the computer and it is too slow to read this much data from the hard disk in every iteration step.

4.1 The Physical Background of the Model

The proton imaging is based on the energy deposition of the protons through the material, described by the theory of Bethe [28] and discussed with corrections in Ref. [29]. The Bethe or also called Bethe–Bloch equation describes the mean rate of energy loss by moderately relativistic charged heavy particles traversing through material:

$$-\left\langle \frac{dE}{dx} \right\rangle = K z^2 \frac{Z}{A} \frac{1}{\beta^2} \left[\frac{1}{2} \ln \frac{2m_e c^2 \beta^2 \gamma^2 T_{max}}{I^2} - \beta^2 - \frac{\delta(\beta\gamma)}{2} \right], \quad (9)$$

where dE/dx is the mean energy transfer rate, also called stopping power in the pCT literature, $K = 4\pi N_A r_e^2 m_e c^2$, N_A is the Avogadro's number, r_e is the classical electron radius, m_e is the mass of the electron, c is the speed of the light, Z and A are the atomic and mass numbers of the ion, respectively. β is the velocity of the ion in the units of the speed of the light (v/c), $\gamma = 1/\sqrt{1-\beta^2}$, T_{max} is the possible maximal energy transfer between the ion and a free electron, I is the mean excitation energy and $\delta(\beta\gamma)$ is the density correction.

The path of the protons is described by multiple coulomb (MC) scattering [30]:

$$\theta_0 = \frac{13.6 \text{ MeV}}{\beta c p} z \sqrt{\frac{x}{X_0}} \left[1 + 0.038 \times \ln \left(\frac{x}{X_0} \right) \right] , \quad (10)$$

where θ_0 is the standard deviation of the Gaussian distribution of the scattering, p , βc and z are the momentum, velocity and charge (in units of elementary charge) of the particle and x/X_0 is the thickness of the material compared with its radiation length.

The tangential displacement of the protons as a result of the change in its direction by MC scattering while passing through the material is investigated in Ref. [31], and is given by

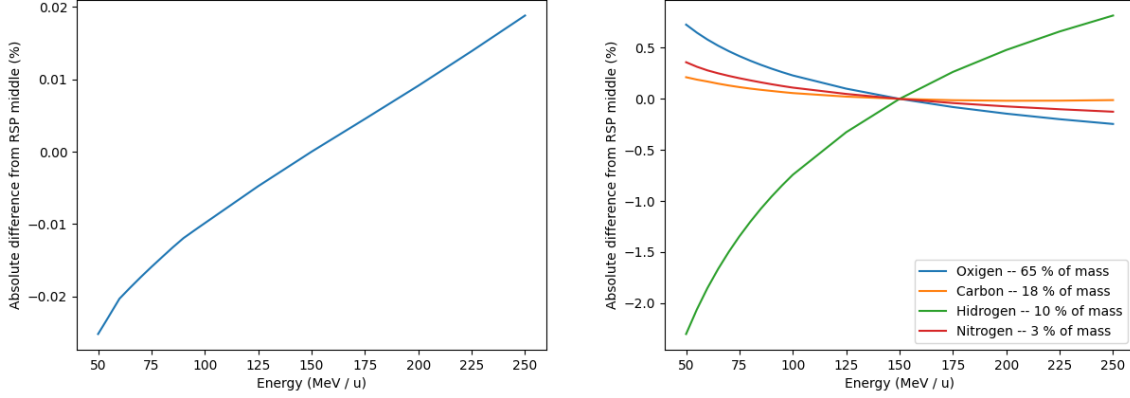
$$\sigma_y = \frac{1}{\sqrt{3}} x \sigma_\theta , \quad (11)$$

where y and θ are the tangential displacement and the scattering angle, respectively after passing a certain thickness of material represented by x . y and θ are correlated, with a coefficient of $\sqrt{3}/2$. σ_y and σ_θ are the standard deviation of the tangential displacement and the scattering angle, respectively. σ_θ is calculated by Equation (10).

Nuclear interactions can also happen while the particle passes through the material, and they can have an impact on the proton imaging as explained in details in Ref. [5].

The stopping power of the materials change with the energy of the proton, described by equation (9). Only energy independent material properties can be reconstructed by the image reconstruction algorithm. If one measures the stopping power of a material in units of stopping power in water with the same energy protons, one can reach the relative stopping power (RSP) of the material, which is approximately independent of the energy in the range of 50-250 MeV/u. I compared the RSP in the previously mentioned energy range with the RSP at 150 MeV/u. I found that the energy dependency of the RSP of the human body is less than 0.025 % (Figure 3(a)), which is good enough for image reconstruction. I investigated the same properties for the four most likely chemical components of the human body. I found that the energy dependency of RSP is less than 2.5 %, in case of the hydrogen and less than 0.75 % in case of oxygen, nitrogen and carbon. Furthermore the RSP of hydrogen changes opposite with oxygen, nitrogen and carbon, so the average can have a much smaller energy dependence for a typical human tissue, which is likely to contain carbon and hydrogen (and other chemical

elements). Lets note that the energy dependency of RSP in case of water is zero, as water is the reference material. The energy dependence of the chemical elements themselves is higher than the acceptable 1 %, but as they are unlikely to stand on their own, the RSP dependence of human tissues is likely to be lower than 1 %, as was seen in case of the whole human body.



(a) Average human body, based on its four most common components.

(b) Most common components of the human body: oxygen, carbon, hydrogen and nitrogen.

Figure 3: The constancy of the relative stopping power (RSP) of the main atoms of the human body as a function of the proton energy in the relevant 50-250 MeV/u energy range. The reference RSP at 150 MeV/u is set as a reference. CTP404 phantom – 100 iterations.

4.2 Modeling of the Imaging

I use list-mode proton imaging, which means, that the detector measures: the energy loss, and the outgoing position and direction of each proton individually. For this data, there are two important steps to reconstruct the image: the first is to investigate the trajectory of the protons in the patient, the second is to reconstruct the image from the proton paths and the energy losses.

The trajectory investigation is based on the most likely path (MLP) of the protons, which can be calculated based on the maximization of the probability of the scattering of the protons

during their path. A more computation efficient and not much less accurate solution is to approximate the MLP with a third-order spline, which was applied in this work as well [31]. The calculation of the elements of the system matrix (which are the interaction coefficients) is based on the probability distribution of the appearance of the real proton path around the estimated MLP of the protons. This method is new in the literature, but is motivated by the published investigations about the MLP calculation techniques [31]. I expect smoother and less noisy images from this method, than the interaction coefficient calculation based on the sections of the MLP in the voxels which is widely used in the literature [31]. The cost of the calculation per iteration is much higher in case of the probability density based method compared with the section based method, therefore results in a slower convergence.

To investigate the performance of the image reconstruction algorithm I developed an imaging model, in which I can set the value of the different types of uncertainties e. g. the error of the measurement of the direction of protons or the beam characteristics. This model approximated the path of the protons with third-order cubic splines as the imaging track, but I can set uncertainty on the beam position and for the measurement of the proton behind the patient. This practically means that the third-order spline will be fitted to slightly different incoming and outgoing information.

The energy deposition of the protons is modeled with the integration of the RSP of the phantom along the path of the protons. It integrates the RSP of voxels around the MLP during the image reconstruction phase. On one hand, this solution makes it possible to test the image reconstruction with exactly the same approximations as during data taking – if I set the uncertainties to zero. On the other hand, if I tend with the standard deviation around the MLP to zero, I can simulate the realistic situation, when the protons passing through their MLP, so their energy loss can be calculated from the track integral of the RSP along the MLP. By setting the realistic uncertainties, the model approximate a realistic data taking considering without nuclear interactions. The main properties of the real imaging, image reconstruction and modeling of imaging with and without uncertainties is listed in Table 1.

We note, the nozzle exit-isocenter and the detector-isocenter distance was set to be smaller

than in the real experimental setup case to compensate the fact that the model uses splines between the nozzle exit and the detector. In a later version this can be improved with straight lines applied outside the phantom, and these distances can be set to realistic values.

	Real imaging	Imaging Model 1	Imaging Model 2	Image
		(as reconstruction)	(realistic)	reconstruction
Trajectory	MC scattering	Spline	Spline	Spline
Beam FWHM	7 mm	0 mm	7 mm	0 mm
Beam divergence	2.8 mrad	0 mrad	2.8 mrad	0 mrad
Direction error	0.89 mrad	0 mrad	0.89 mrad	0 mrad
Multiply scattering	yes	yes	yes	yes
Scattering angle	22.0 mrad	22.0 mrad	22.0 mrad	22.0 mrad
Nuclear interactions	yes	no	no	no
σ around MLP	-	1 mm	0.25 & 0.125 mm	1-10 mm
Nozzle exit - isocenter	500 mm [12]	150 mm	150 mm	150 mm
Detector - isocenter	250 mm [12]	150 mm	150 mm	150 mm
Resolution ($n \times n$)	-	64	2048	64-512

Table 1: The main physical parameters that were applied in the image modelling and reconstruction for simulating the interactions between protons and the material of the phantom.

4.3 The Numerical Model

In my model, I assumed that the interaction between the protons and voxels is based on the probability density (PD) of the real path that follows a normal distribution around the MLP of the protons. I suspected that this consideration makes the linear problem less well-conditioned, but I also expected, that it works better with the relatively small number of protons typically measured in a pCT detector.

In order to speed up the calculations, fast parallel computing methods were applied. Graphical card, also known as video card including graphical processing units (GPU) is widely used for such computations. Recently program libraries are openly available for general purpose graph-

ical processing units (GPGPU). I implemented my model in C++ and in CUDA, which latter is a GPU programming language based on C++ and developed by Nvidia [32]. The main focus of the code development is the parallelization of the problem. There are two parallelization levels in the code and both are used during the image reconstruction.

1. The first level is the parallel computation of the protons, this solution is used for Hadamard ratio calculation. The current implementation runs on a single CPU core, which manage all the 14336 GPU threads on four Nvidia 1080Ti GPU cards. This structure can be fast enough, since the machines' CPU reads only the data from the storage and synchronizes the GPU processes. If one CPU core would not be enough during further optimization, the code can be parallelized more on the CPU level based on the parallel processing of protons.
2. The second parallelization level is the parallelization of the proton voxel interactions. This parallelization has advantageous memory use, because each GPU thread multiply the same Hadamard ratio, but each with different element of the system matrix, belongs to different voxels. Finally each thread add the result of the multiplication to the sum vector of the voxel. This multiplication is safe from race conditioning (when more than one thread write the same time the same variable), as every thread process different voxels, so write different elements of the sum vector.

Here, I apply both parallelization methods for my work, of which the structural outline is shown in Algorithm 2.

Algorithm 2 GPU implementation of the Richardson–Lucy algorithm for pCT

```
1: set  $N$ ,  $M$ ,  $L$  and  $T$ : number of iterations, trajectories, voxels and treads, respectively
2: set  $X$ , the vector of the voxel values to be a unit vector
3: set  $dT$  and  $normT$  to  $L$  length zero vectors and  $Y$  is the vector of the measured data
4: for  $N$  cycles do
5:   set  $i = 0$ 
6:   while  $i < M/T$  do
7:      $H_t = Y_{i+t} / \sum_j^L A_{i+t,j} \cdot X_j$ , for all  $t$ , where  $t$  is the index of the thread,  $H$  is the
       Hadamard ratio and  $A_{i+t,j}$  is the on the fly calculated interaction coefficient between the
        $i + t^{\text{th}}$  voxel and the  $j^{\text{th}}$  proton trajectory
8:     set  $j = 0$ 
9:     while  $j < T$  do
10:       $dX_k += H_j \cdot A_{k,j}$  and  $normX_k += A_{k,j}$ , for all  $k$ , where  $k$  is the index of voxels
11:    end while
12:  end while
13:   $X_i = X_i \cdot dX_i / normX_i$  for all  $i$ , where  $i$  is the index of voxels
14:  set  $dT$  and  $normT$  to zero vectors
15: end for
16: save  $X$  vector as the resulted image
```

I applied the 1st parallelization in the step 7, then I used the 2nd parallelization in the step 9-11. One can notice that the elements of the system matrix are calculated in this structure two times. This is necessary to preserve the computer memory storing the rows of the system matrix for each proton evaluated in an independent thread during the 1st parallelization.

5 Results on the Richardson – Lucy Algorithm for pCT

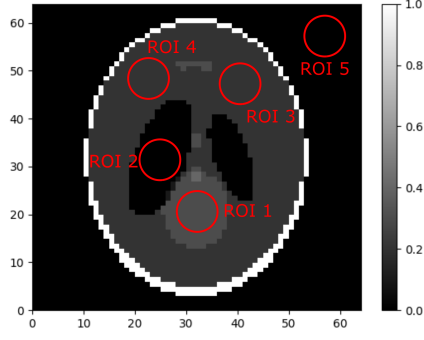
In this section, I will investigate the performance of the reconstruction algorithm with the use of three different phantoms. The first of them is the classical Shepp – Logan phantom [33], which I used to test the convergence of the algorithm in case of imaging with Model 1 and 2 defined in Table 1. Model 1 does not consider any uncertainties of the measurement, in contrary to Model 2, which considers the spatial and angular uncertainties of the incoming and outgoing positions and directions as well. The second phantom is the Derenzo phantom [34], which I applied to investigate the resolution of the modeled imaging process with the Richardson – Lucy reconstruction algorithm. The third phantom is the CTP404 [35], which is used primarily to investigate the RSP accuracy of the imaging. Moreover, this latter phantom was also for investigating the longitudinal position errors of the Bragg peak if when irradiating a certain point of the CTP404 phantom and planing the dose based on the map taken by the modelled imaging system. I keep the standard deviation of the distribution in 1 mm during this Section.

The first phantom I used is the Shepp – Logan phantom, which is presented in Figure 4(a). It is a commonly used, two dimensional phantom recommended by Larry Shepp and Benjamin F. Logan [33]. This model was used to test the convergence of the Richardson – Lucy deconvolution if I set the uncertainties of the imaging to zero. In this case I expect that the algorithm will tend to the exact solution.

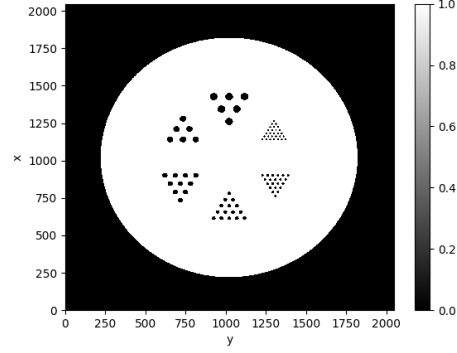
The second phantom was used to test the spatial resolution of the imaging process is the Derenzo phantom in Figure 4(b) developed by Stephen E. Derenzo in Ref. [34]. This phantom has more variants, I used the so called cold rod variation, which is a water equivalent cylinder that contains cylindrical holes filled with air, which are placed in a triangular structure such that their distance from each other is the same as the diameter of the holes. This phantom contains six regions with hole diameters of 1.5, 2, 3, 4, 5 and 6 mm.

The third phantom is called CTP404 presented in Figure 4(c), which is produced by The Phantom Laboratory [35]. The CTP404 phantom is designed to measure how accurately an imaging system measures the properties of different materials, which is in my case the Relative Stopping Power (RSP). This phantom is an epoxy cylinder, which contains 8 inserts with

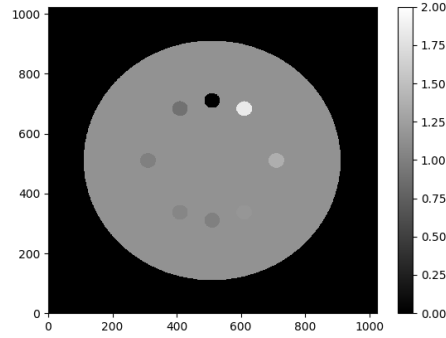
different materials and each with a diameter of 12.2 mm. The RSP of the inserts and the epoxy was investigated by Alme et al. [2]. I evaluated the reconstructed average RSP in an 8 mm diameter circle in the middle of each inserts.



(a) Shepp–Logan phantom.



(b) Derenzo phantom.



(c) CTP404 phantom.

Figure 4: Phantoms used to investigate the convergence (a), spatial resolution (b) and RSP accuracy (c) of the image reconstruction.

5.1 Convergence

In this section I investigated the convergence of the image reconstruction with and without measurement uncertainties of the incoming and outgoing positions and directions. I analyzed the reconstruction of the Shepp–Logan phantom without and with measurement uncertainties based on the parameters of Model 1 and 2, respectively, as summarized in Table 1. I used

simulation with 3.6×10^4 protons for the imaging both without and with the measurement uncertainties. This latter was also tested with 10 times more statistics as well.

Figure 5(a) shows the root mean square (RMS) error, which is the L2 norm of the difference between the reconstructed and the original phantom. It is clearly visible that the RMS error of the image, which was taken without the measurement uncertainties tends to zero. The RMS error of the images, which was taken with the measurement uncertainties converges to the non-zero RMS error limit. This is certainly lower for the case of the higher-statistics case.

Figure 5(b) shows the error of the average RSP of the whole image. This error approaches to zero within 1 000 iteration in case of imaging without measurement uncertainties. This is not true for the case with measurement uncertainties, but the image with higher statistics reach a lower error.

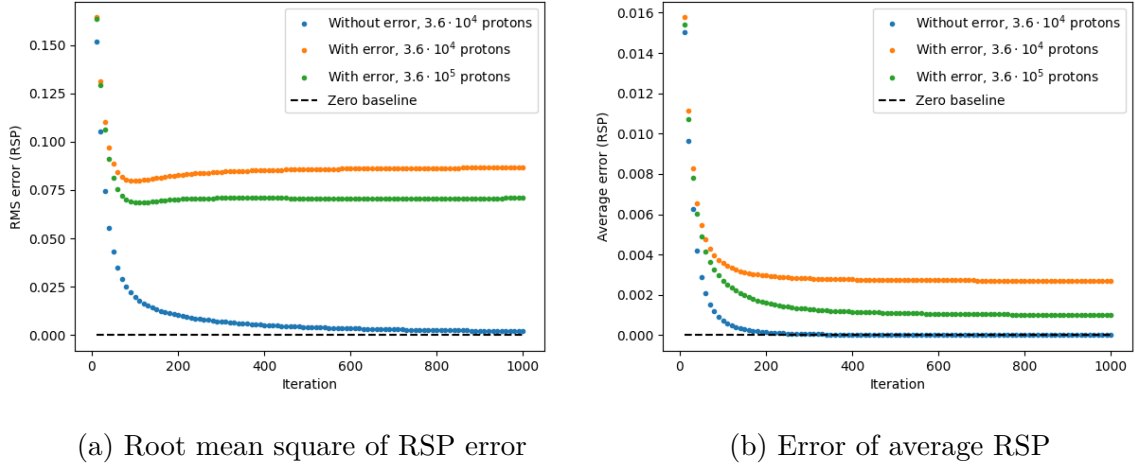


Figure 5: Root mean square of the RSP error (a) and the error of the average RSP of the image (b). A 64×64 Shepp–Logan Phantom was reconstructed with 1000 iterations and $\sigma = 1$ mm.

I defined five regions of interest (ROIs) in the Shepp–Logan phantom as it is visible in Figure 4. The 'ground truth' RSP value of the ROIs are 0.3, 0.0, 0.2, 0.2 and 0.0, respectively. The average noise of the five ROIs is visible in Figure 6(a). The noise of the image that was taken without measurement uncertainties is increasing in the beginning of the iterations, has a maximum around iteration 60 and decreasing between iteration 60 and iteration 1000.

The noise of the images that were taken with measurement uncertainties is increasing in the beginning and saturating. This is significantly lower in case of the image with higher-statistics (3.6×10^5) proton case than the one with lower (3.6×10^4) number of protons. One can notice that the noise of the images with 3.6×10^4 protons are increasing in the same way during the first 30 iterations. Figure 6(b) shows the average difference between the reconstructed and the 'ground truth' RSP value of the ROIs. This difference will be negligible for the image without measurement uncertainties, and result in a non-zero limit for the image with measurement uncertainties. One can conclude that this limit for the case, where the image was taken with 3.6×10^5 protons is much lower than it is for the image that was taken with 3.6×10^4 protons. Reconstructed images after 100 and 1000 iterations is visible in Figures 7(a-c), and 7(d-f), respectively.

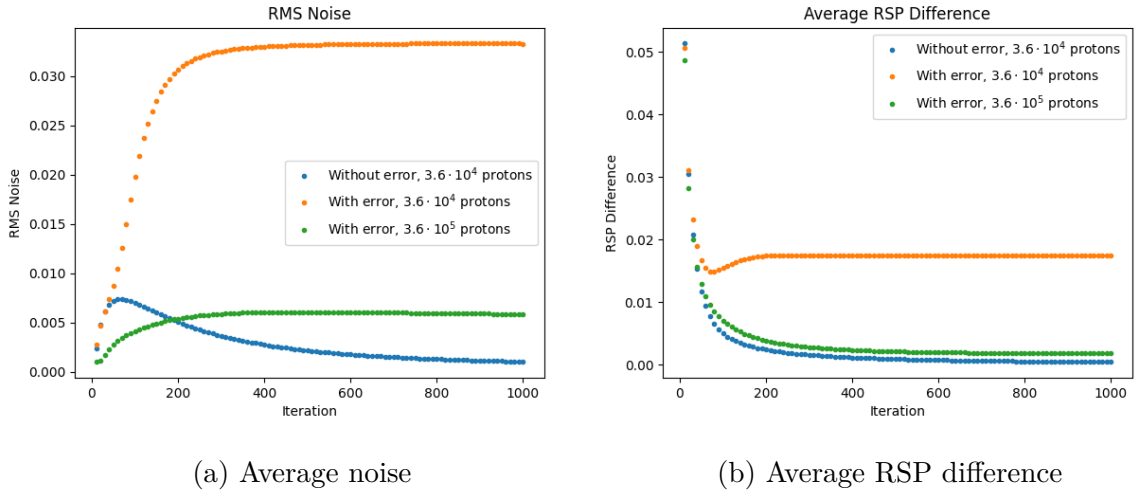
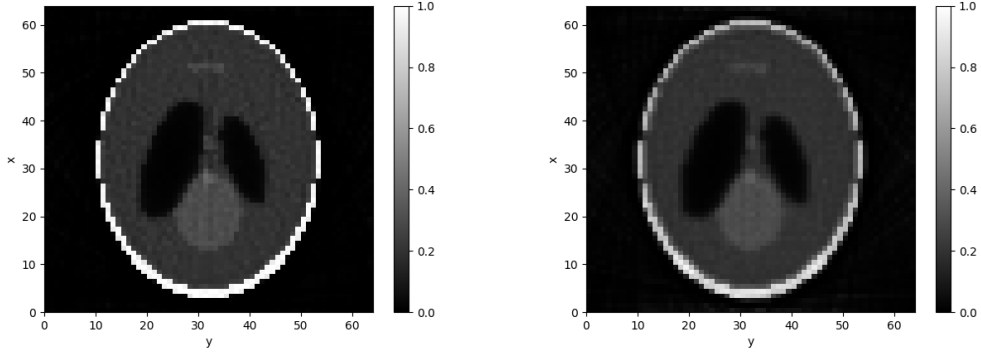
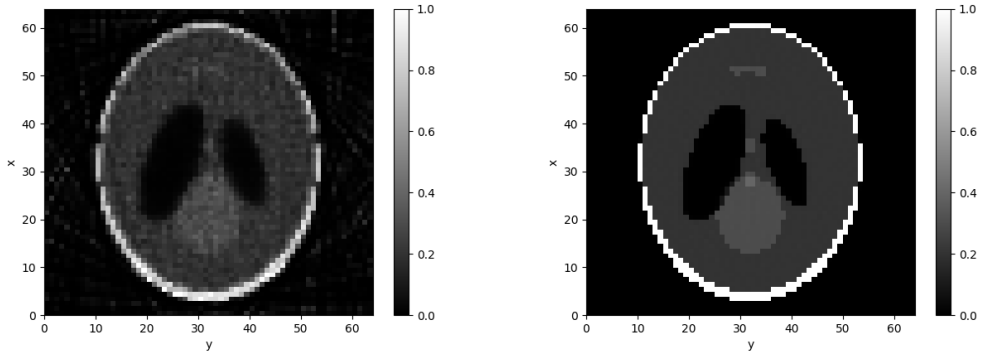


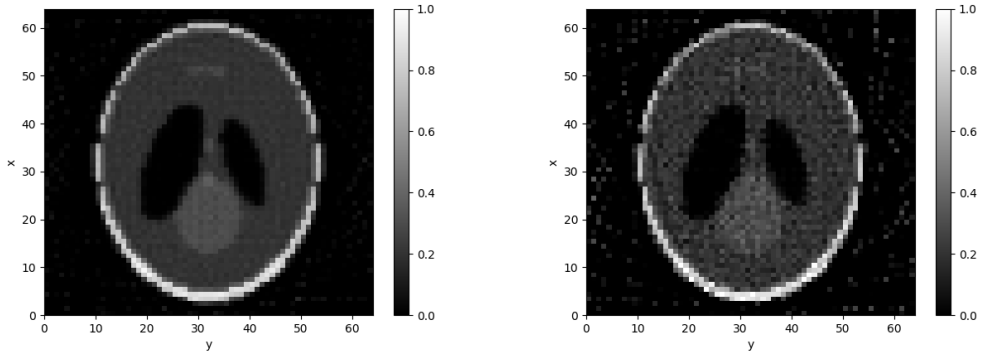
Figure 6: The average noise (a) and the average RSP difference (b) measured in the ROIs defined in Figure 4(a). The 64×64 Shepp–Logan Phantom was reconstructed with 1000 iterations and $\sigma = 1$ mm.



(a) No uncertainties after 100 iterations (b) 3.6×10^5 protons after 100 iterations



(c) 3.6×10^4 protons after 100 iterations (d) No uncertainties after 1000 iterations



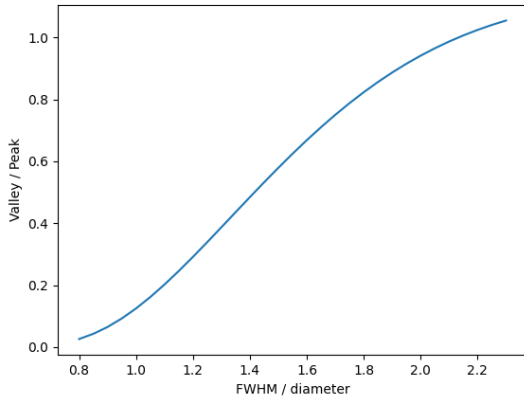
(e) 3.6×10^5 protons after 1000 iterations (f) 3.6×10^4 protons after 1000 iterations

Figure 7: The 64×64 Shepp–Logan Phantom which was reconstructed with $\sigma = 1$ mm. (a) and (d) were taken without measurement uncertainties, while (b), (c), (e), and (f) were taken with measurement uncertainties.

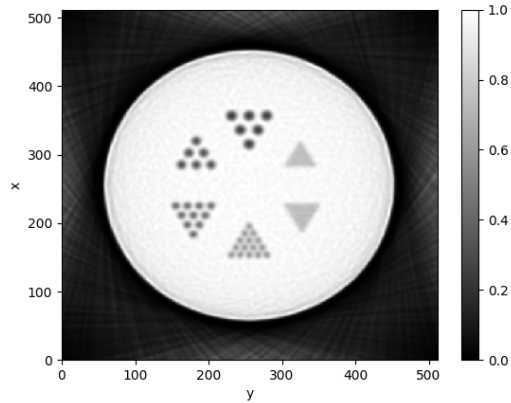
5.2 Spatial Resolution

The goal of this section is to analyse the spatial and the RSP resolution (called as RSP accuracy) of the reconstructed images. In this work, I defined the spatial resolution of the image as the full width at half maximum (FWHM) of the point spread function (PSF), which convolved with the original distribution would give the same valley-to-peak ratio as the reconstructed distribution. The image of two rods after convolution with the point spread function is approximately the sum of two normal distributions. With calculate the valley-to-peak ratio for many FWHM values, one can plot the valley-to-peak ratio as a function of the FWHM of the PSF (Figure 8(a)).

I tested the spatial resolution with the imaging of the Derenzo phantom, with 3.6×10^6 protons, 512×512 image size and 250 iterations. Other parameters were set according to imaging Model 2 in Table 1. The reconstructed phantom can be seen in Figure 8(b). The diameter of the rods are 6, 5, 4, 3, 2, and 1.5 mm.



(a) Valley-peak and FWHM.

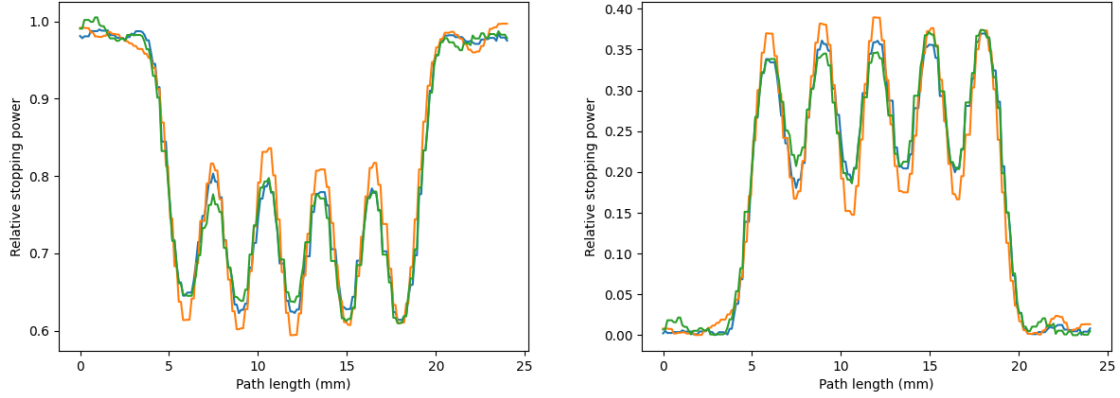


(b) Reconstructed Derenzo phantom.

Figure 8: The valley-to-peak ratio as a function of the FWHM of the PSF (a). The reconstructed Derenzo Phantom (b).

I evaluated the resolution based on the valley-to-peak ratio in Figure 9(a). First, I subtracted the background, which is equal to the RSP value in the edges of Figure 9(a). I took the absolute

value of the distribution from Figure 9(b). Second, I read the valley and peak values of the distribution in Figure 9(b). I took the average of them and calculated their ratio, for which I obtained 0.521. Finally, I compared the valley-to-peak ratio to the FWHM in units of diameter in Figure 8(a), which is 1.47; this ratio multiplied by the 3 mm rod diameter results 4.4 mm FWHM as the resolution of the reconstructed image.

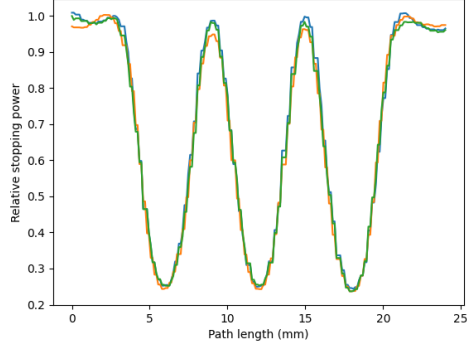


(a) RSP in straight lines.

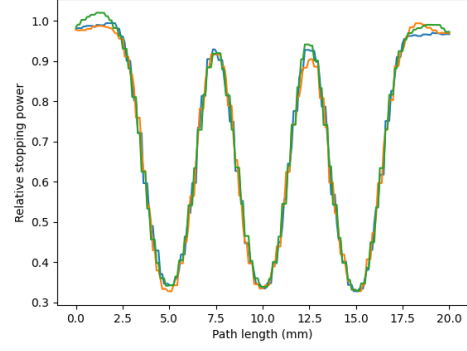
(b) Absolute RSP after background subtraction.

Figure 9: RSP in straight line through 3 mm rods (a). The absolute value of the RSP distribution after the division by the background (b).

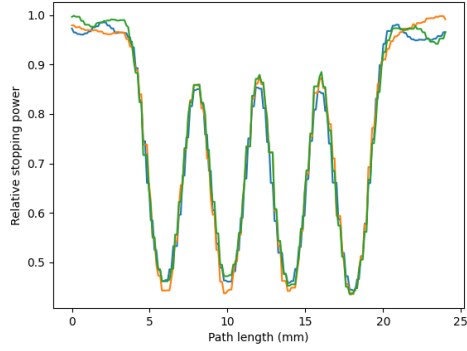
I tested the resolution, which was calculated based on the first four sectors (6-3 mm rod diameter) of the Derenzo phantom, because the valleys and the peak were separable in these sectors. The valleys and the peaks was merging in the last two sectors (2-1.5 mm rod diameter), which makes them unusable in the resolution calculation process. The RSP along straight lines through the rods of the six sectors are visible in the Figure 10(a-f). The resolution as a function of the iteration is visible in Figure 11. In the beginning of the iteration the four sectors result significantly different resolutions, meanwhile the resolutions estimated by the four sector are in the narrow range between 4.4 and 4.65 mm after 250 iterations. Figure 11 also shows that the reconstruction has not reached the best possible resolution after 250 iterations.



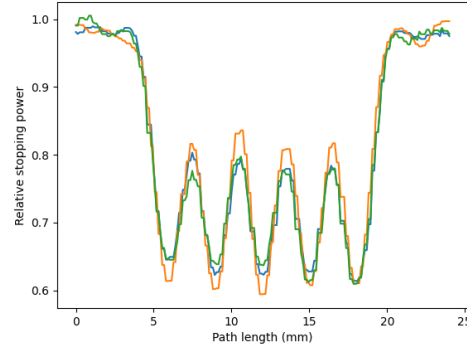
(a) 6 mm rod diameter



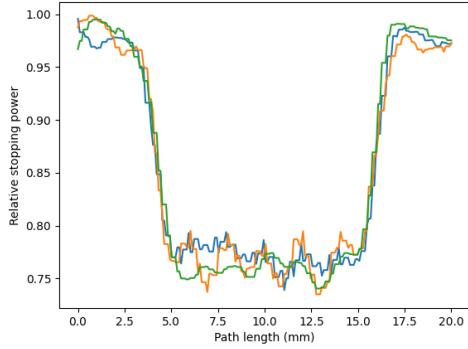
(b) 5 mm rod diameter



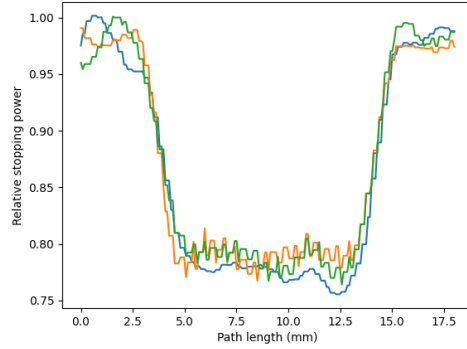
(c) 4 mm rod diameter



(d) 3 mm rod diameter



(e) 2 mm rod diameter



(f) 1.5 mm rod diameter

Figure 10: Relative stopping power along straight lines through the middle of the rods with different diameters in Figure 8(b). The imaging of the Derenzo phantom was taken with realistic measurement uncertainties (see Model 2 in Table 1) and 3.6×10^6 proton histories. The number of iterations was 250 during image reconstruction.

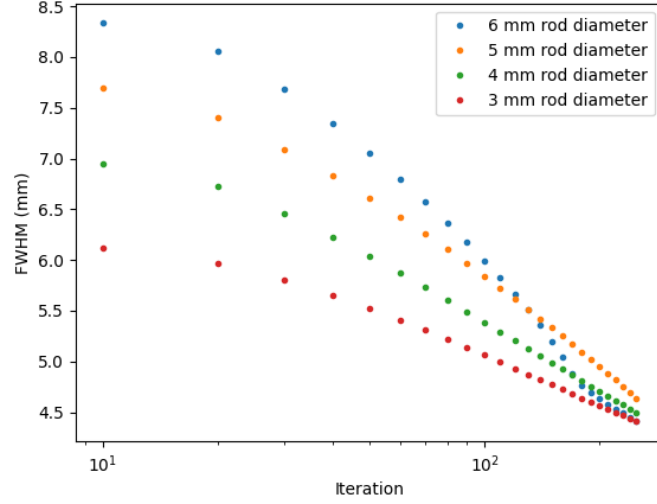


Figure 11: Resolution of the imaging estimated based on four sectors (6-3 mm rod diameter) of the Derenzo phantom which is visible in Figure 8(b).

5.3 Relative Stopping Power Accuracy

The RSP accuracy was tested with a phantom based on the CTP404 phantom. I used the parameters of Imaging Model from 2 from Table 1 for the imaging. The parameters of the image reconstructions were the following:

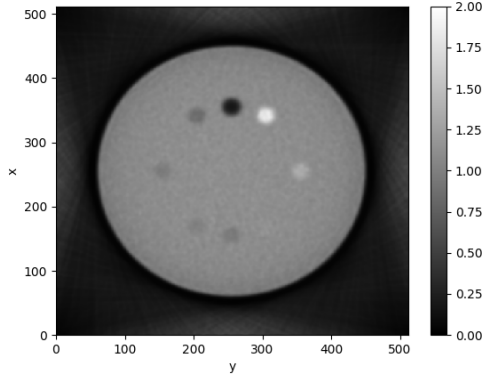
- Number of voxels: 512,
- Voxel edge size: 0.5 mm,
- Number of protons: 3.6×10^6 ,
- Number of iterations: 100, 200, and 300.

The list of the inserts and their 'ground truth' and reconstructed RSP values are visible in Table 2. I defined the reconstructed RSP value as the average RSP in an 8 mm diameter circle in the middle of the inserts.

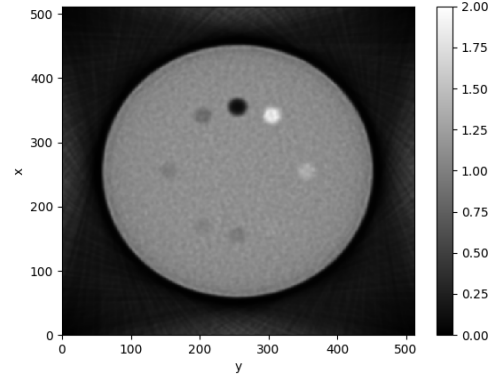
Material	'Ground truth' RSP	Reconstructed average RSP	Standard deviation of average RSP	Relative difference (%)
Air	0.0	0.102	0.027	-
Teflon	1.833	1.803	0.058	-1.62
Delrin	1.363	1.338	0.039	-1.84
Acrylic	1.179	1.155	0.034	-2.04
Water	1.0	0.972	0.029	-2.81
Polystyrene	1.022	1.031	0.030	-2.49
LDPE	1.003	0.985	0.027	-1.78
PMP	0.886	0.866	0.038	-2.22
Epoxy	1.143	1.130	0.039	-1.14

Table 2: Inserts of the CTP404 phantom. The 'ground truth' RSP value was taken from one of the previous works of the Bergen pCT Collaboration [2].

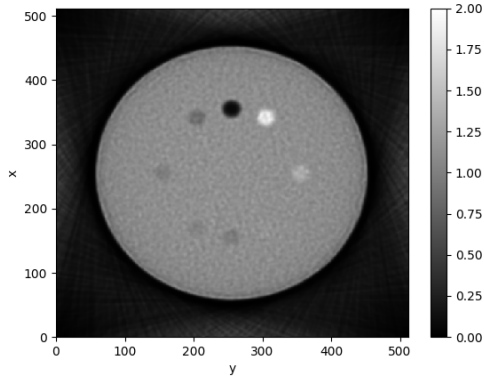
The accuracy of the RSP reconstruction of an object with a typical size of 12.2 mm is better than 3%, which is far from the 0.4% published earlier by our group based on simulations with the same phantom [2]. The RSP accuracy is increasing with the number of iterations between 100 and 300 iterations as Figure 12(f) shows. At the same time the noise of the image is increasing as presented in Figure 12(e) which is also visible in the reconstructed images in Figures 12(a-c). The 'ground truth' RSP value and the reconstructed RSP values after 100, 200 and 300 iterations is shown by Figure 12(d). The change in the reconstructed RSP of the air is much more significant between 100 and 300 iterations than the change in the RSP of the other inserts.



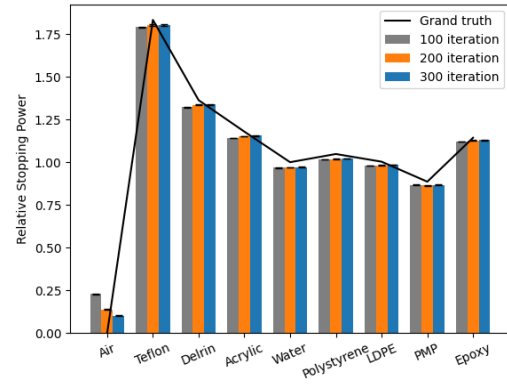
(a) CTP404 phantom – 100 iterations.



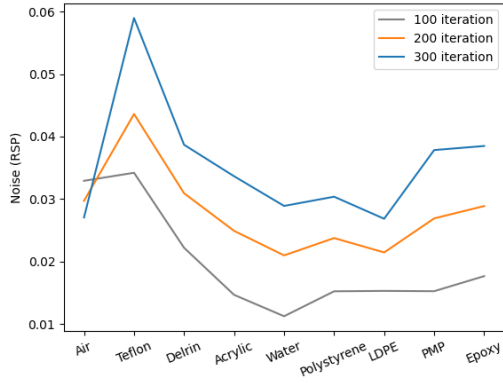
(b) CTP404 phantom – 200 iterations.



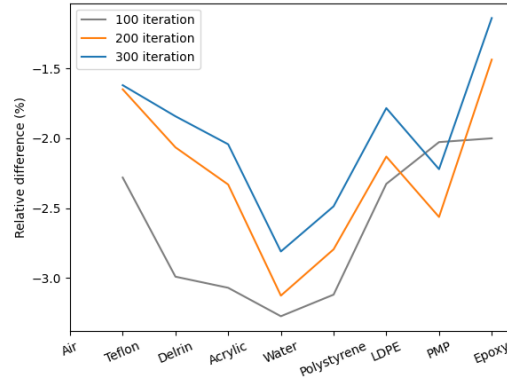
(c) CTP404 phantom – 300 iterations.



(d) Reconstructed and real RSP.



(e) RMS noise.



(f) Relative error of RSP.

Figure 12: Reconstructed CTP404 Phantom (a), (b) and (c), reconstructed and real RSP value (d), RMS noise of the inserts (e) and relative difference of RSP values (f).

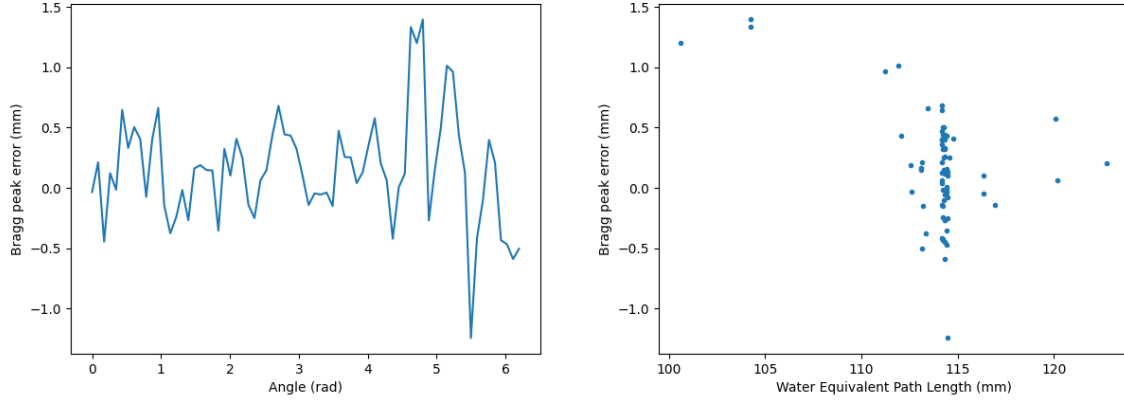
5.4 The Longitudinal Position Error of the Bragg Peak

Here, I introduce a clinically-motivated investigation of the image quality. The final purpose of the pCT images is to represent an accurate RSP map about the patient for dose planning. Following this, I tried to find a measure which characterizes the image quality from the point of view of clinical usefulness. I decided to compare the integrated RSP along a straight line to a certain point of the phantom from a certain direction.

The first step is to find a point of the phantom, which represents the location of the tumor. The second step is of this measure is to calculate the integrated RSP value along the straight line to the selected point, based on the reconstructed RSP map. This calculation of the energy for the proton beam, represented in water equivalent thickness (WET). The third step is to calculate the stopping point (which is almost the same with the maximum point of the Bragg peak) of the beam with the same energy in the same straight line, but based on the original RSP map of the phantom. The distance of this point and the selected point characterizes the accuracy of the reconstructed RSP map. This measure can be automated and quickly evaluated for a numerous set of selected points in an arbitrary phantom.

This measure was investigated on the CTP404 phantom. The first selected point was the central point of the phantom. The direction of the straight line was changed between 0° and 360° in 5° steps. The direction is measured from 12 hour position in Figure 4(c). The position error of the Bragg peak (measured in the direction of the line) in every step is visible in Figure 13(a) and stays within ± 1.5 mm. The position error of the Bragg peak is not correlated with the direction, which result can be interpreted as the reconstruction of the CTP404 is symmetrical. The WEPL along the lines is the same from most of the directions, as the phantom is a homogeneous cylinder containing only eight inserts, which are small compared to the size of the phantom. If we calculate the WEPL based on the 'ground truth' RSP distribution most of the measure points belongs to a small WEPL range as it is shown in Figure 13(b). In this case it is not possible to investigate the dependency of the Bragg peak position error as a function of the WEPL of the line.

The second selected point was the middle of the water insert (6 hour position in Figure 4(c)).

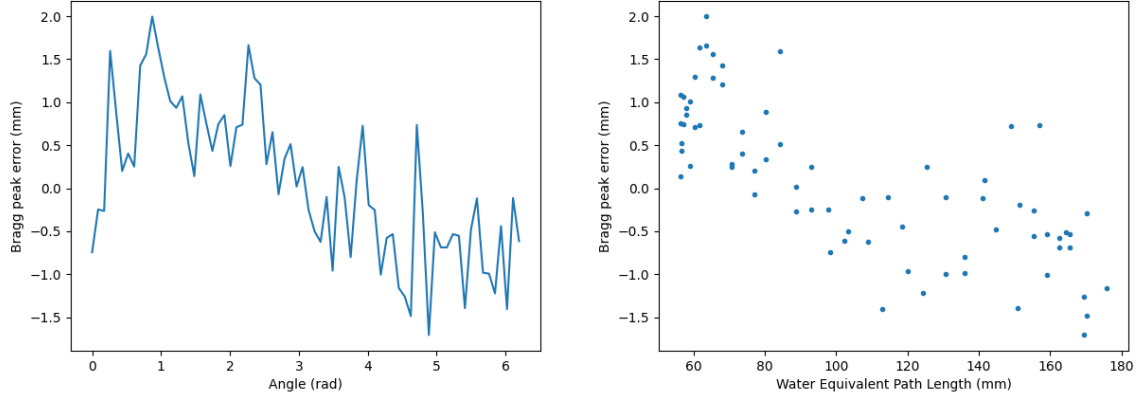


(a) Bragg Peak error in function of path angle. (b) Bragg Peak error in function of WEPL.

Figure 13: The error of the Bragg Peak position as a function of the beam direction (a). The error of the Bragg Peak position as a function of the water equivalent path length of the proton before stopping (b). The beam aims at the middle of the CTP404 phantom.

The error of the Bragg peak is less than 2 % and is visible as a function of the direction in Figure 14(a). One can notice that there is some correlation between the direction and the error, which is confirmed by the Figure 14(b), where a clear negative correlation is visible between the WEPL of the line and the Bragg peak position error. The WEPL of the line is smaller than the average if it does not go through the middle of the phantom; in this case the WEPL is overestimated based on the reconstructed RSP map. The WEPL of the line is higher than the average if it goes through the middle of the phantom; in this case the WEPL is underestimated based on the reconstructed RSP map. This result suggested that the RSP of the middle of the phantom is underestimated, which agrees with the result that the RSP of all the inserts except air is underestimated.

I fixed a direction, and points were selected on a grid. I generated a map about the Bragg peak position errors of the selected points (Figure 15). The Bragg peak error outside the phantom is higher than 10 mm. Because the RSP of the air was set to zero in this investigation, so the protons could not stop in the air. We note, that the error in the position of the air insert is also higher than in other regions of the phantom due to the same reason.



(a) Bragg Peak error in function of path angle. (b) Bragg Peak error in function of WEPL.

Figure 14: (The error of the Bragg Peak position as a function of the beam direction (a). The error of the Bragg Peak position as a function of the water equivalent path length of the proton before stopping (b). The beam aims at the middle of the water insert.

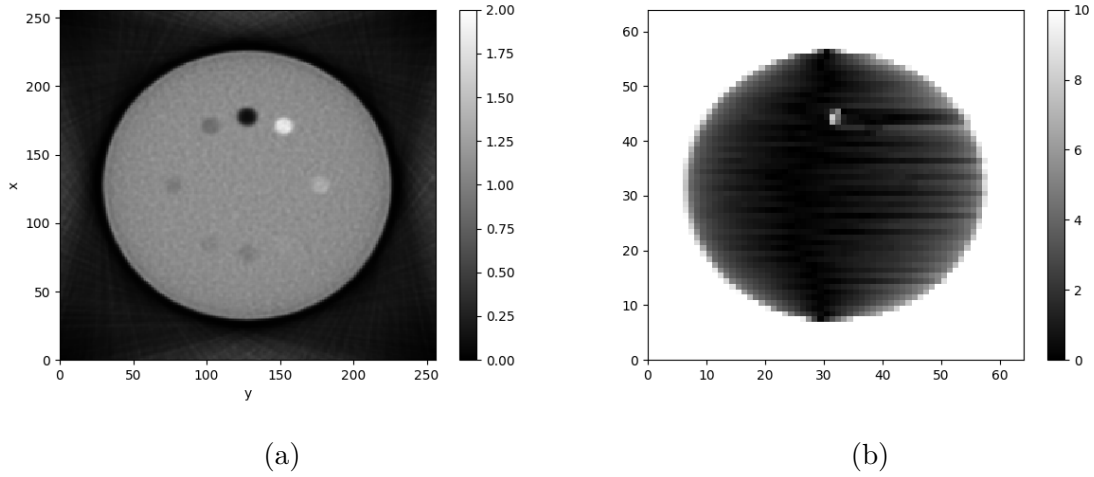


Figure 15: Reconstructed CTP404 phantom (a). Bragg Peak position error map (b).

I have showed that this type of measure can be useful for the evaluation of the image quality, so it is worth the further investigation to develop, test and standardise this measure.

6 Results on PD-based Interaction Calculation

In this section I investigated the performance of the interaction modeling based on the probability density around the MLP of the protons. I tested the image reconstruction algorithm with different standard deviations of the probability density (PD) around the MLP of the protons.

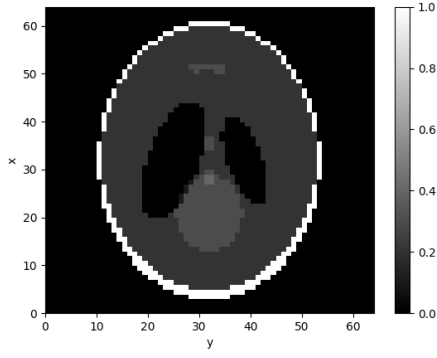
6.1 Test of the Convergence

I tested the reconstruction algorithm with imaging Model 1, which is similar with the physical model of the image reconstruction. I used the 64×64 voxel Shepp–Logan phantom for this purpose and I measured the L2 or Euclidean norm of the difference between the reconstructed image and the phantom. Note, I used the same number of voxels in the phantom and in the image. I changed the standard deviation of the normal distribution around the MLP between 1 mm and 10 mm.

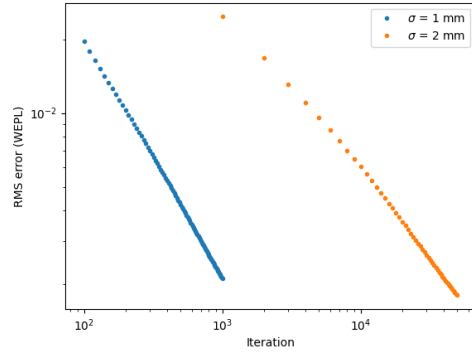
The L2 norm tends to zero within 1 000 and 50 000 iterations for the 1 mm and 2 mm cases, respectively as it can be seen in Figure 16(b). The reconstructed images are presented in Figures 16(c) and 16(d), which are almost indistinguishable from the phantom in Figure 16(a). The difference becomes only visible if zoomed into the images, scaled into the RSP range of one region of the phantom in Figures 16(e) and 16(f).

Theoretically the L2 norm tends to zero, in case of 5 mm and 10 mm as well, but it would require too many iterations to show this convergence based on Figure 17.

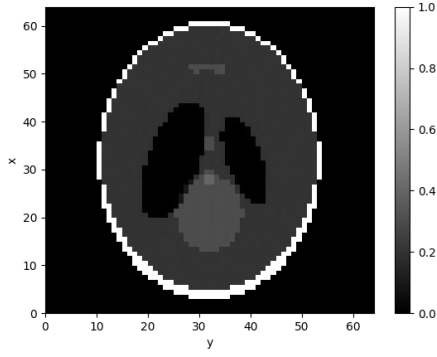
The first conclusion of this convergence test is that the image reconstruction works well, and the problem is well-conditioned – at least in the case of $\sigma = 1$ mm and 2 mm. The second conclusion is that the number of required iterations for the same convergence is quickly increasing with the value of the σ parameter. We note, this represents the standard deviation of the probability distribution of the proton path around the MLP of the proton, which is the most likely the sign of the determination of the linear equation system's condition. The main question is if the conditioning of the problem depends on the value of the σ in the image



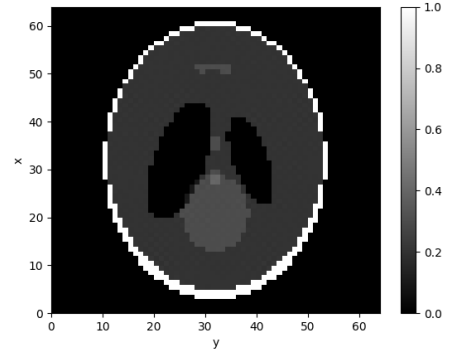
(a) Shepp-Logan Phantom



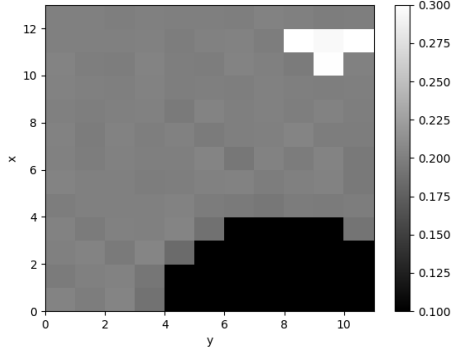
(b) Convergence



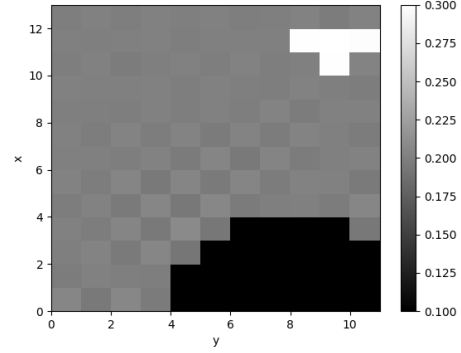
(c) $\sigma = 1 \text{ mm}$



(d) $\sigma = 2 \text{ mm}$



(e) Zoomed, $\sigma = 1 \text{ mm}$



(f) Zoomed, $\sigma = 2 \text{ mm}$

Figure 16: Shepp -Logan Phantom (a), convergence of the iteration (b), reconstructed image with $\sigma = 1 \text{ mm}$ and 2 mm (c) and (d), after 1000 and 50000 iterations, respectively, (e) and (f): zoom into the image and into 0.1-0.3 RSP range of the images in (c) and (d), respectively.

reconstruction. This can be easily optimized, however this depends on the uncertainties of the most likely path (MLP) determination, which is more difficult to improve.

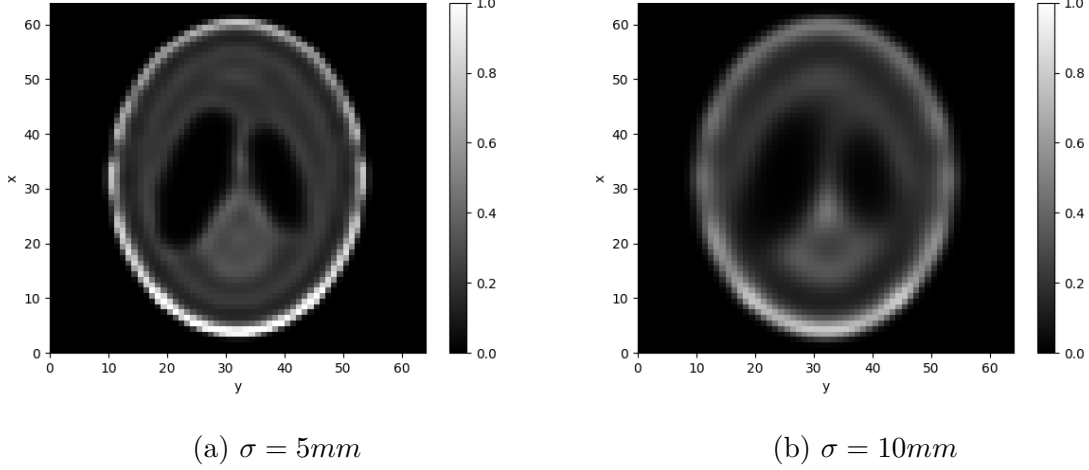


Figure 17: Reconstruction of 64×64 Shepp–Logan Phantom, after 1000 iterations, $\sigma = 5$ mm (a) and $\sigma = 10$ mm (b).

6.2 Average Relative Stopping Power Accuracy

I analyzed the RSP accuracy as a function of the the standard deviation, σ_{pd} of the probability density distribution around the MLP. I reconstructed the CTP404 phantom with both statistics of 3.6×10^5 and 1.0×10^5 protons, respectively. I applied 512×512 voxels and 300 iterations. Other parameters were set from Table 1, based on Model 2.

Imaging with realistic statistics

The root means square (RMS) error of the reconstructed image (or with other words the L2 norm between the reconstructed and the original image) is visible in the left panel of Figure 18. One can conclude that the RMS error is reaches the minimum after 300 iterations in case of $\sigma_{pd} = 2$ mm, which means that this σ_{pd} is optimal in this terms. The error of the average RSP is visible in the right panel of Figure 18, where a clear relation is visible between: the higher σ_{pd} is lower the error of the average RSP.

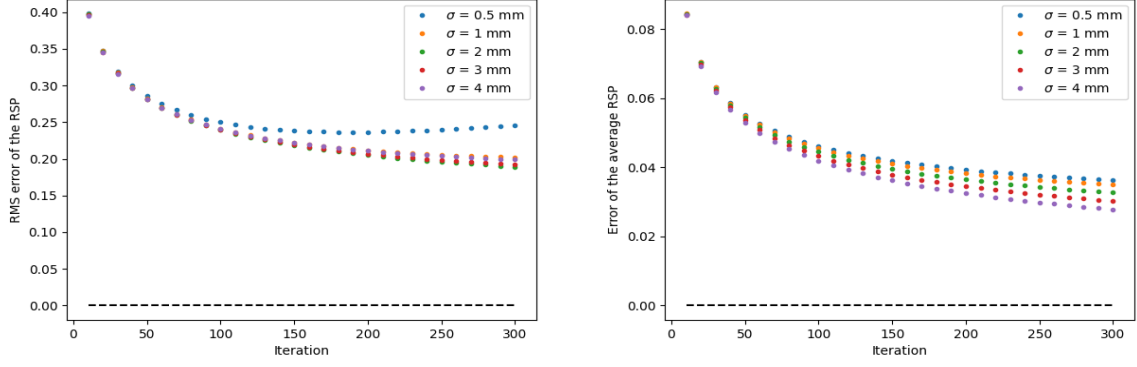


Figure 18: The RMS error of the image (left) and the error of average RSP (right).

Left panel of the Figure 19 shows the noise measured in the inserts as a function of the iteration. The noise is increasing with the iteration in case of every σ_{pd} . The the lower the σ_{pd} the higher the noise after the same number of iterations. The average difference of the RSP of the inserts is visible in the right panel of the Figure 19. One can notice that the difference has a minimum between 40 and 60 iterations in case of every σ_{pd} . If the number of iterations is higher than 60 the difference is increasing slower in case of 0.5 and 1 mm σ_{pd} than in case of the others. The difference seems to reach a limit in case of $\sigma_{pd} = 0.5, 1$, and 2 mm.

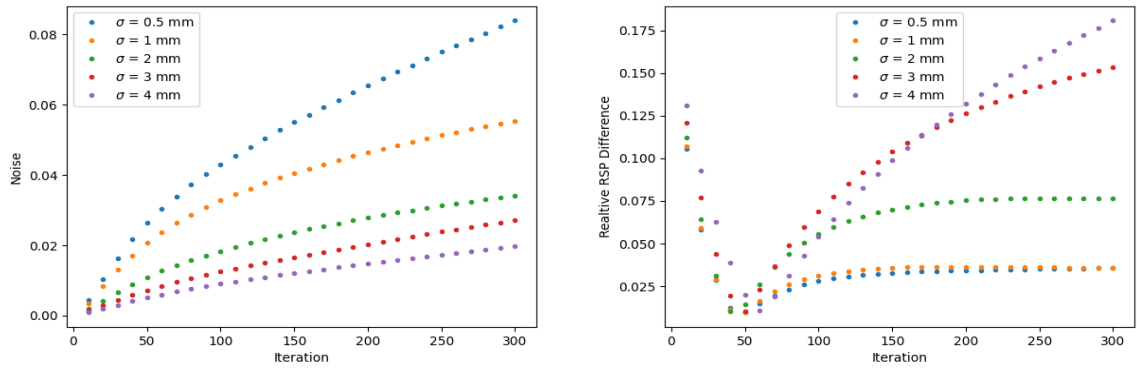


Figure 19: The noise of the reconstructed image (left) and the average difference between the reconstructed and the 'ground truth' RSP of the inserts (right).

Imaging with low statistics

The root means square (RMS) error of the reconstructed image (or with other words the L2 norm between the reconstructed and the original image) is visible in the left panel of Figure 20. One can conclude that the RMS error is the lowest after 300 iterations in case of $\sigma_{pd} = 3$ mm, which means that this σ_{pd} is optimal in this terms. The error of the average RSP is visible in the right panel of Figure 20, where a clear relation is visible between: the higher is the σ_{pd} the lower is the error of the average RSP.

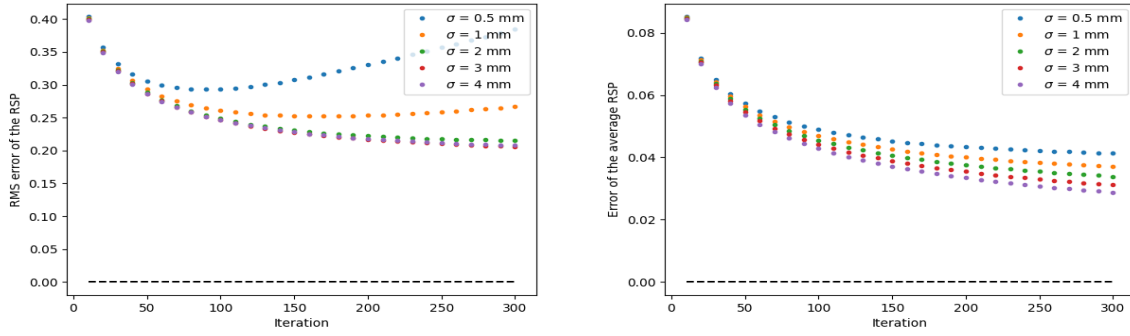


Figure 20: The root mean square error of the reconstructed image (left) and the error of the average RSP of the whole image (right).

The left panel of Figure 21 shows the noise measured in the inserts as a function of the iteration. The noise is increasing with the iterations for every σ_{pd} values. The the lower the σ_{pd} the higher the noise after the same number of iterations. The average difference of the RSP of the inserts is visible in the right panel of Figure 21. One can notice that the lower the σ_{pd} the lower the minimum difference reached during the iteration. As the σ_{pd} becomes lower, the number of iterations required to reach the minimum difference is increasing, but every iterations reached their minimum difference between 40 and 70 iterations.

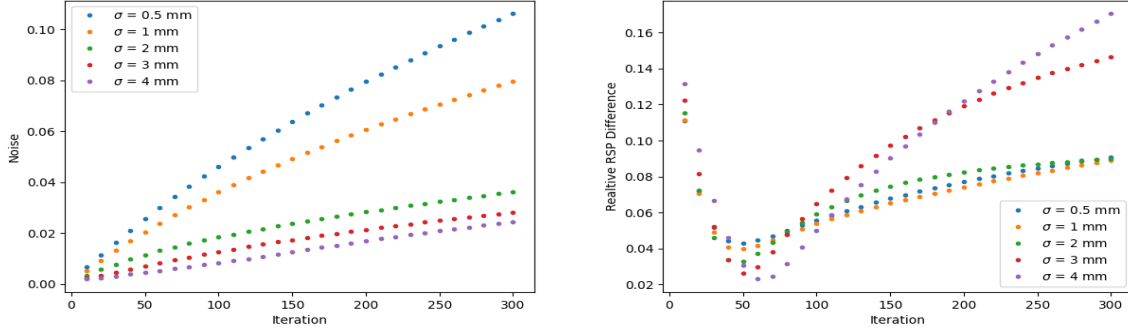


Figure 21: The noise of the image (left) and the average RSP error of the inserts (right).

Conclusion

The first conclusion of the analysis of the RSP accuracy as a function of the number of iterations evaluated with different σ_{pd} constants is that in case of every analyzed reconstruction in this subsection the reconstruction reached the highest RSP accuracy between 40 and 70 iterations. The second conclusion is that the L2 norm seems to have an optimal σ_{pd} , which represents a balance between the high noise in case of reconstructions with low σ_{pd} and the appearing artefacts in case of high σ_{pd} . The artefact appears around 100 iterations in case of high σ_{pd} is seen in Figure 22. The third conclusion is that in case of low statistics (in this case means 1×10^5 proton histories) the lower σ_{pd} can results higher RSP accuracy.

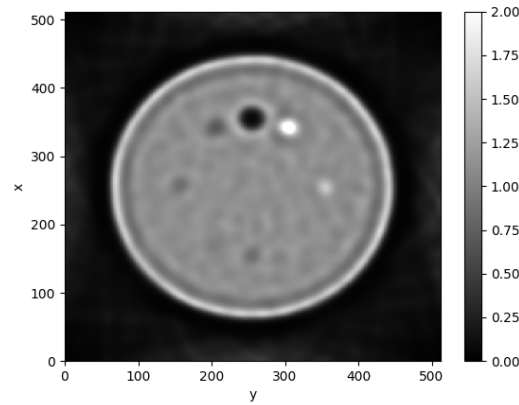


Figure 22: Artefacts in a Derenzo phantom reconstructed with 3.6×10^5 protons (300 iterations).

7 Summary

Thanks to the novel particle acceleration techniques, the role of proton therapy has been becoming more-and-more emphasized in the clinical practice in recent years. The number of hadron therapy centers has increased as well, indeed a variety of hadron beams is now available. One of the key questions of this emerging field today is how would it be possible to reduce the safety margin around the tumor and reduce the dose to the healthy tissues? As I presented, the most promising direction is the use of proton CT technology during dose planning. This can provide a more accurate relative stopping power (RSP) map of the patient than the standard radiology methods by X-ray. My recent thesis is related to this research topic as a member of the Bergen Proton CT international R&D collaboration.

My focus was on the pCT imaging and the investigation of the image reconstruction with pCT, as it has a direct effect on the image quality. Nowadays, an improved version of the Diagonally Relaxed Orthogonal Projection method is used for image reconstruction in general. In my research, I investigated the Richardson–Lucy algorithm for pCT image reconstruction for the first time, since it performs well in the field of emission tomography. To reach reasonable speed with this algorithm, a well structured, computational and memory efficient algorithm was implemented on four GPU cards provided by the Wigner GPU Laboratory.

My simulation results confirmed the applicability of the Richardson–Lucy algorithm. I tested this by using various phantoms: Shepp–Logan, Derenzo, and CTP404. The algorithm exactly restored the image if the imaging errors are turned off during the simulation. I have investigated the effect of measurement uncertainties on the reconstruction quality and performed a parameter study to maximize the image accuracy. I also found that the determination of the Bragg-peak position is quite precise even in this stage of the development.

The presented results contributed to the image reconstruction developments of the proton CT project of the Bergen pCT collaboration. Furthermore, the developed code, which can be also applied in hadron therapy dose planning, in general, after further optimization and speed improvements.

Acknowledgements

First of all I would like to thank Gergely Gábor Barnaföldi for believing in this work even in the most critical times. Without his advices and help I couldn't have walked the path that leads to these results. I would like to thank Mónika Varga-Kőfaragó for investing a huge amount of time and energy to launch my career. Last but not least I would like to thank Dávid Légrády for my knowledge about image reconstruction techniques.

I would like to thank the support of the Hungarian National Research, Development and Innovation Office (NKFIH) grants under the contract numbers OTKA K135515 and 2019-2.1.6-NEMZ_KI-2019-00011. Computational resources were provided by the Wigner GPU Laboratory.

References

- [1] D. C. Hansen *et al*, A simulation study on proton computed tomography (CT) stopping power accuracy using dual energy CT scans as benchmark, *Acta Oncologica*, **54**, 2015. doi: 10.3109/0284186X.2015.1061212.
- [2] J. Alme *et al*, A High-Granularity Digital Tracking Calorimeter Optimized for Proton CT, *Frontiers in Physics*, **8**, 2020, doi: 10.3389/fphy.2020.568243.
- [3] A. M. Cormack, Representation of a Function by Its Line Integrals, with Some Radiological Applications, *Journal of Applied Physics* **34**, 1963. doi: 10.1063/1.1729798.
- [4] A. M. Cormack, Representation of a Function by Its Line Integrals, with Some Radiological Applications. II, *Journal of Applied Physics* **34**, 1964. doi: 10.1063/1.1713127.
- [5] R. P. Johnson, Review of medical radiography and tomography with proton beams, *Reports on Progress in Physics*, **81**, doi: 10.1088/1361-6633/aa8b1d.
- [6] S. Damjanovich, J. Fidy and J. Szöllösi, *Sugárzások és kölcsönhatásuk az "élő" anyaggal* (in Hungarian), *Orvosi Biofizika, Medicina Könyvkiadó RT.* 2007.
- [7] J. F. Fowler, What can we expect from dose escalation using proton beams?, *Clin Oncol (R Coll Radiol)*, **15**, 2003, doi: 10.1053/clon.2002.0182.
- [8] H. Paganetti, Range uncertainties in proton therapy and the role of Monte Carlo simulations, *Phys. Med. Biol.* **57**, 2012, doi: 10.1088/0031-9155/57/11/R99.
- [9] H. E. S. Pettersen *et al*, Design optimization of a pixel-based range telescope for proton computed tomography, *Physica Medica*, **63**, 2019, doi: 10.1016/j.ejmp.2019.05.026.
- [10] C. Golnik *et al*, Range assessment in particle therapy based on prompt γ -ray timing measurements, *Phys Med Biol*, **59**, 2014, doi: 10.1088/0031-9155/59/18/5399.
- [11] C. Richter *et al*, First clinical application of a prompt gamma based in vivo proton range verification system, *Radiother Oncol*, **118**, 2016, doi: 10.1016/j.radonc.2016.01.004.

- [12] J. R. Sølvi et al, Image quality of list-mode proton imaging without front trackers, *Phys. Med. Biol.* **65**, 2020. doi: 10.1088/1361-6560/ab8ddb.
- [13] H. E. S. Pettersen *et al*, Proton tracking in a high-granularity Digital Tracking Calorimeter for proton CT purposes, *Nuclear Instruments and Methods in Physics*, **860**, 2017, doi: 10.1016/j.nima.2017.02.007.
- [14] B. Abelev *et al*, Technical Design Report for the Upgrade of the ALICE Inner Tracking System, *Journal of Physics G Nuclear and Particle Physics*, **41**, 2014, doi: 10.1088/0954-3899/41/8/087002.
- [15] G. Tambave *et al*, Characterization of monolithic CMOS pixel sensor chip with ion beams for application in particle computed tomography, *Nuclear Instruments and Methods in Physics Research Section A*, **958**, 2020, doi: 10.1016/j.nima.2019.162626.
- [16] Á. Sudár, *Proton CT - új diagnosztikai eszköz a rákos daganatok kezeléséhez* (in Hungarian), 34. OTDK, 2018, abstract available: <https://otdk.hu/eredmenyek/62521> (downloaded on 27.10.2021).
- [17] Á. Sudár, Investigating the temperature distribution of a hadron-tracking calorimeter, BME TDK, 2019, abstract available: <http://tdk.bme.hu/GPK/EN4/j5> (downloaded on 27.10.2021).
- [18] Á. Sudár, Measurement of the temperature distribution inside a calorimeter, BSc thesis, BME GPK Department of Energy Engineering, 2019, arXiv: <https://arxiv.org/abs/2005.02830> (downloaded on 27.10.2021).
- [19] J. Radon, *Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten*, , *Berichte über die Verhandlungen der Königlich-Sächsischen Akademie der Wissenschaften zu Leipzig, Mathematisch-Physische Klasse*, [Reports on the Proceedings of the Royal Saxonian Academy of Sciences at Leipzig, Mathematical and Physical Section], Leipzig: Teubner, 1917.

- [20] R. Gordon, Robert Bender and Gabor T. Herman, Algebraic Reconstruction Techniques (ART) for Three-dimensional Electron Microscopy and X-ray Photography, J. theor. Biol. **29**, 1970, doi: 10.1016/0022-5193(70)90109-8.
- [21] L. B. Lucy, An iterative technique for the rectification of observed distributions, The Astronomical Journal, **79**, 1974, doi: 10.1086/111605.
- [22] W. H. Richardson, Bayesian-Based Iterative Method of Image Restoration, Journal of the Optical Society of America, **62**, 1972, doi: 10.1364/JOSA.62.000055.
- [23] Á. Sudár, Statistically-motivated, Iterative Image Reconstruction Algorithm for proton Computer Tomography Applications, BME TDK, 2021, available: <https://tdk.bme.hu/TTK/DownloadPaper/Statisztikai-motivacioju-iterativ> (downloaded on 08.12.2021).
- [24] S. N. Penfold *et al*, Total variation superiorization schemes in proton computed tomography image reconstruction, American Association of Physicists in Medicine, **37**, 2010, doi: 10.1118/1.3504603.
- [25] Y. Censor *et al*, On Diagonally-Relaxed Ortoogonal Projection Methods, SIAM J Sci Comput, **30**, 2008, doi: 10.1137/050639399.
- [26] G. T. Herman, *Fundamentals of Computerized Tomography*, Second Edition, Springer-Verlag London, 2009, doi: 10.1007/978-1-84628-723-7.
- [27] M. S. Kaczmarz, *Angenäherte Auflösung von Systemen linearer Gleichungen* (in German), Int. Bull. Pol. Acad. Sci. Lett. 35 355–7, 1937.
- [28] H. A. Bethe, *Zur Theorie des Durchgangs schneller Korpuskularstrahlen durch Materie* (in German), Annalen der Physik, **397**, 1930, doi: 10.1002/andp.19303970303.
- [29] C. Amsler *et al*, Review of Particle Physics, Physics Letters B, **667**, 2008, doi: 10.1016/j.physletb.2008.07.018.

- [30] C. Patrignani *et al*, Review of Particle Physics, Chin. Phys. C **40**, 2016, doi: 10.1088/1674-1137/40/10/100001.
- [31] D. C. Williams, The most likely path of an energetic charged particle through a uniform medium, Phys. Med. Biol. **49**, 2004, doi: 10.1088/0031-9155/49/13/010.
- [32] Home page of NVIDIA CUDA: <https://developer.nvidia.com/cuda-zone> (downloaded on 27.10.2021).
- [33] L. A. Shepp, B. F. Logan, The Fourier reconstruction of a head section, IEEE Transactions on Nuclear Science, **21**, 1974, doi: 10.1109/TNS.1974.6499235.
- [34] S. E. Derenzo *et al*, High Resolution Computed Tomography of Positron Emitters, IEEE Transactions on Nuclear Science, **24**, 1977, doi: 10.1109/TNS.1977.4328738.
- [35] Catphan® 600 phantom (containing CTP404 as a section), Phantom Laboratory, available: <https://www.phantomlab.com/catphan-600> (downloaded on 27.10.2021).