



# Polynomial modelling of proton trajectories in homogeneous media for fast most likely path estimation and trajectory simulation

Nils Krah, Jean-Michel Létang, Simon Rit

## ► To cite this version:

Nils Krah, Jean-Michel Létang, Simon Rit. Polynomial modelling of proton trajectories in homogeneous media for fast most likely path estimation and trajectory simulation. *Physics in Medicine and Biology*, 2019, 64 (19), pp.195014. 10.1088/1361-6560/ab3d0b . hal-02269520

**HAL Id: hal-02269520**

**<https://hal.science/hal-02269520>**

Submitted on 2 Dec 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Polynomial modelling of proton trajectories in homogeneous media for fast most likely path estimation and trajectory simulation

N. Krah<sup>1,2</sup>, JM. Létang<sup>1</sup>, S. Rit<sup>1</sup>

<sup>1</sup>University of Lyon, CNRS, CREATIS UMR5220, Centre Léon Bérard, Lyon, France

<sup>2</sup>University of Lyon, CNRS, Institute of Nuclear Physics Lyon (IPNL) UMR 5822, Villeurbanne, France

*nils.krah@creatis.insa-lyon.fr*

## Abstract

Protons undergo many small angle deflections when traversing a medium, such as a patient. This effect, known as multiple Coulomb scattering (MCS), leads to degraded image resolution in proton radiography and computed tomography (CT) and to lateral spreading of the dose distribution in proton therapy. To optimally account for MCS in proton imaging, the most likely path (MLP) of a proton is estimated based on its position and propagation angle measured in front of and behind the object. In this work, we propose a functional which quantifies the likelihood of a proton trajectory and study how it can be used to model proton trajectories in a homogeneous medium. We focus on two aspects: first, we present an analytical method to quickly generate proton trajectories in a homogeneous medium based on the likelihood functional and validate it through Monte Carlo simulations. It could be used for fast generation of proton CT images without a full Monte Carlo simulation, or potentially to complement the components in a treatment planning Monte Carlo which simulate MCS. Second, by maximising the likelihood functional, we derive an expression for the MLP which is equivalent to the conventional ones reported in the literature yet computationally more convenient. Moreover, we show that the MLP is strictly a polynomial function if the protons' energy loss in the medium is approximated as a polynomial and that the orders of both are linked. We validate our MLP through Monte Carlo simulations and compare proton CT images reconstructed with our expression and with the conventional one. We find that an MLP polynomial of orders larger than five do not lead to increased spatial resolution compared to lower order expressions.

## 1 INTRODUCTION

Protons in the therapeutic energy range ( $< 300$  MeV) undergo a large number of small angle deflections when traversing a medium due to multiple Coulomb scattering (MCS) (Gottschalk et al., 1993). In proton therapy, this leads to a successive lateral spreading of a pencil beam as it penetrates the patient tissue and its impact on the dose distribution needs to be accurately modelled by treatment planning systems. In proton radiography and computed tomography (CT), MCS is responsible for the relatively low spatial resolution compared to X-ray imaging. In these imaging techniques, protons traverse a patient at sufficiently high energy to be captured on the downstream side by a suitable detector device (Johnson, 2018; Parodi, 2014). The most sophisticated systems measure each individual proton's position and propagation angle in front of and behind the patient with tracking detectors (Schulte et al., 2004; Scaringella et al., 2014; Civinini et al., 2013; Penfold et al., 2011; Taylor et al., 2016). Using this information as constraint, the most likely path (MLP) which the proton has followed across the medium can be estimated (Williams, 2004; Schulte et al., 2008; Li et al., 2006; Collins-Fekete et al., 2015; Collins-Fekete et al., 2016). The tomographic reconstruction is then performed using the MLP (Rit et al., 2013; Poludniowski et al., 2014; Penfold et al., 2015) and the uncertainty of the MLP estimation compared to the real trajectory is what limits the spatial resolution of the images. The conventional MLP estimation methods estimate the most likely transverse position through which a proton has passed in a certain depth and construct the MLP as a concatenation of these transverse positions by repeating the estimation in a series of depths across the medium. They exploit the fact that all MCS events are independent of each other and that the angular dispersion due to MCS is approximately Gaussian. The cited methods have been thoroughly validated with Monte Carlo simulations and one variant or another of them are commonly used in proton CT reconstruction. However, they do not provide an explicit way to quantify the likelihood of a given trajectory. This idea is elaborated from a mathematical point of view by Erdelyi (2009) who discusses the concept of random fields in this context. He describes a theoretical framework to derive an MLP estimate from an action functional, however provides few explicit calculations with a direct practical application.

In this work, we introduce a likelihood functional which quantifies the likelihood of a proton trajectory based on its weighted curvature integrated along depth and investigate how it can be used to model proton trajectories in a homogeneous medium. In particular, the likelihood functional allows us to achieve two things: first, to analytically generate random proton trajectories as they would be produced by a Monte Carlo simulation in a homogeneous medium, and second, to derive an expression for the MLP by maximising the likelihood functional, as suggested by Erdelyi (2009), which is equivalent to the conventional expressions. The functional disregards nuclear interactions and only takes into account MCS within the Gaussian approximation. Generating proton trajectories analytically would be of interest for quickly producing proton CT projections without running a full Monte Carlo simulation, e.g. for testing a reconstruction algorithm, or to partially replace the proton transport in a Monte Carlo code, e.g. in the context of treatment planning. We show that the MLP automatically becomes a polynomial as long as the protons' energy loss is approximated by a polynomial (which is common practice in proton CT reconstruction) and thereby create a link to the works of Li et al. (2006) and Collins-Fekete et al. (2015). We validate our MLP expression with Monte Carlo generated data and test it in the proton CT reconstruction of a standard image quality phantom. The advantage of our expression compared to the conventional ones is that it requires fewer floating point operations to evaluate while providing the flexibility to select the desired degree of estimation accuracy by choosing the polynomial order.

## 2 MATERIAL AND METHODS

### 2.1 Framework of the formalism

The purpose of this work is to analytically model proton trajectories in homogeneous media. When traversing such a medium, protons undergo a large number of small angle deviations due to MCS. The likelihood of each such deflection depends on the radiation length  $X_0$  of the material and on the proton energy  $E$ . We measure the depth coordinate  $u$  along the axis obtained by extending the proton's entrance vector across the object and denote with  $t$  the transverse position of a proton, i.e. projected onto a plane perpendicular to the  $u$ -axis. We define as trajectory the position  $t(u)$  as a function of  $u$ . The propagation angle  $\theta$  of a proton projected onto this plane is given relative to the  $u$ -axis, so that  $\theta = 0$  means parallel to the  $u$ -axis. The three-dimensional trajectory is given by two transverse positions  $t_x$  and  $t_y$  in the two planes perpendicular to the  $u$ -axis.

We do not attempt to describe heterogeneities transverse to the beam path. Furthermore, the developed formalism accounts for multiple Coulomb scattering (MCS), but not for discrete nuclear scattering events. The main context within which we performed this investigation is proton computed tomography where nuclear events are usually filtered out and in any case not accounted for by most tomographic reconstruction algorithms.

### 2.2 Likelihood functional of proton trajectories

Within the Gaussian approximation of MCS, the likelihood of a change in propagation angle  $\Delta\theta = \theta_2 - \theta_1$  over a depth interval  $\Delta u$  is given by

$$\mathcal{L}(\Delta\theta) = \exp\left(-\frac{(\Delta\theta/\Delta u)^2}{2T}\Delta u\right) = \exp\left(-\frac{(\Delta\theta)^2}{2T\Delta u}\right), \quad (1)$$

where  $T$  is the scattering power which depends on the proton's energy and the radiation length of the medium (Lynch et al., 1991; Gottschalk et al., 1993). The likelihood  $\mathcal{L}[t]$  of a trajectory  $t$  can be thought of as the joint probability of a series of many deflections  $\Delta\theta_i$  which have caused the proton to follow  $t(u)$ ,

$$\mathcal{L}[t] = \prod_i \mathcal{L}(\Delta\theta_i) = \prod_i \exp\left(-\frac{(\Delta\theta_i/\Delta u)^2}{2T_i}\Delta u\right) = \exp\left(-\frac{1}{2} \sum_i \frac{(\Delta\theta_i/\Delta u)^2}{T_i}\Delta u\right). \quad (2)$$

Because the propagation angle  $\theta$  relative to the beam axis is small, it can be taken as the derivative to the trajectory with respect to depth,  $\theta = \partial t/\partial u$ . Making the steps  $i$  in equation 2 above infinitesimally small and considering infinitely many of them, the sum becomes an integral and the likelihood is given as

$$\mathcal{L}[t] = \exp\left(-\frac{1}{2} \int_0^{u_{\text{out}}} \frac{(\partial^2 t(u)/\partial u^2)^2}{T(u)} du\right) = \exp\left(-\frac{1}{2} \int_0^{u_{\text{out}}} \frac{(t''(u))^2}{T(u)} du\right), \quad (3)$$

where we have replaced the finite difference  $\Delta\theta_i/\Delta u = ((\partial t/\partial u)_i - (\partial t/\partial u)_{i-1})/\Delta u$  by the second derivative and used  $t''$  as short hand notation. Without loss of generality, we chose to place the beginning of the trajectory at  $u = 0$ ;  $u_{\text{out}}$  denotes the depth at which the proton exits the volume. According to equation 3, the likelihood of a trajectory is given by its integrated curvature weighted by the depth-dependent scattering power  $T(u)$  to account for the proton's energy and the material's radiation length.

We parametrise the scattering power as  $T(u) = E_0^2/\beta^2(u)p^2(u)X_0(u)$ , with  $E_0 = 13.6 \text{ MeV}/c$  a constant,  $\beta = v/c$  the proton's velocity relative to the speed of light,  $p$  the proton's momentum, and  $X_0$  the radiation length. This is an approximate form of the expression proposed by Highland (1975). More accurately, it would contain an additional logarithmic term which is



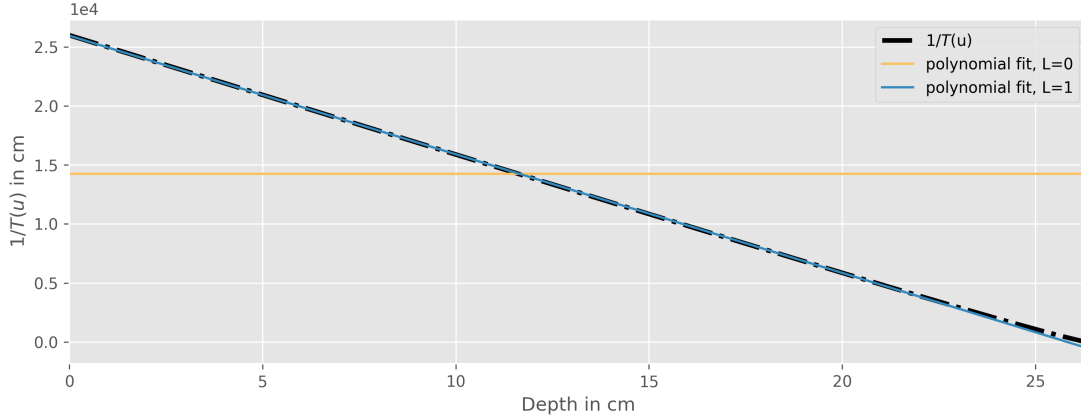


Figure 1: Depth dependent inverse scattering power  $1/T(u) = \beta^2(u)p^2(u)X_0(u)/E_0^2$  for 200 MeV protons in a homogeneous water phantom. Coloured curves are polynomial fits.

not straight forward to incorporate in the integral expression of the likelihood functional (equation 3). Some discussion on this can be found in Erdelyi (2009). For the scope of this work, we restrict ourselves to the approximate expression and leave it to future contributions to explore other parametrisations of  $T(u)$ . Furthermore, we only consider homogeneous media in this work, although parametrising  $T(u)$  appropriately would allow modelling longitudinal material heterogeneities, i.e. along the  $u$ -axis (Collins-Fekete et al., 2017).

### 2.3 Analytical generation of proton trajectories as polynomials

MCS is the most prominent physics process in a Monte Carlo simulation of protons in the therapeutic range ( $< 300$  MeV) along with electromagnetic energy loss. To generate a proton trajectory, a Monte Carlo code has to perform many steps, each time sampling the angular deviation the proton suffers. We therefore investigated how the likelihood function (equation 3) could instead be used to analytically generate random proton trajectories as a whole. Because the likelihood function only models MCS, such an analytical generation does not account for discrete nuclear scattering events. We consider mono energetic protons propagating through a box of water. For simplicity, we will first assume that all protons have the same initial position and angle, although this is not a requirement.

We approximate proton trajectories by polynomial functions,

$$t(u) = \sum_{n=0}^N a_n u^n = t_{\text{in}} + \theta_{\text{in}} u + \sum_{n=2}^N a_n u^n, \quad (4)$$

so that a trajectory  $t$  is represented by an  $N + 1$  coefficient vector  $\vec{a} = (a_0, a_1, \dots, a_N)^T$ . On the right hand side, we have used the fact that  $t(u = u_{\text{in}}) = t_{\text{in}} = a_0$  and  $t'(u = u_{\text{in}}) = \theta_{\text{in}} = a_1$ . Within the framework of polynomial functions, sampling trajectories is equivalent to sampling coefficient vectors  $\vec{a}$ . The second derivative  $t''$  in the likelihood function (equation 3) can be calculated explicitly and is

$$t''(u) = \sum_{n=2}^N n(n-1)a_n u^{n-2}. \quad (5)$$

We further expand the function  $1/T(u)$  into a polynomial,

$$\frac{1}{T(u)} \approx \sum_{l=0}^L c_l u^l. \quad (6)$$

It can be seen from figure 1 that already  $L = 1$  gives a good approximation. Using this polynomial expression, one obtains for the weighted trajectory curvature

$$\frac{(t''(u))^2}{T(u)} = \sum_{n=2, n'=2, l=0}^{N, N, L} c_l n(n-1)n'(n'-1)a_n a_{n'} u^{l+n+n'-4}. \quad (7)$$

Inserting this into the likelihood function (equation 3) and integrating the polynomial analytically yields

$$\mathcal{L}_{\text{poly}}[t] = \exp \left( -\frac{1}{2} \sum_{n=2, n'=2}^{N, N} a_n a_{n'} \Phi_{nn'} \right) \quad (8)$$

with

$$\Phi_{nn'} = \sum_{l=0}^L \frac{c_l n(n-1)n'(n'-1)}{l+n+n'-3} u_{\text{out}}^{l+n+n'-4}. \quad (9)$$

In other words, the likelihood distribution of trajectories is approximated by an (N-1)-dimensional Gaussian distribution of the coefficient vectors  $\vec{a}$ . In practice, to generate a random sample of coefficients  $\vec{a}$ , the following procedure is convenient: first diagonalise  $\Phi$  (which is always possible because the matrix is symmetric and positive definite) to obtain  $\Psi = Q^T \Phi Q$ , where the columns of  $Q$  are the unit length eigenvectors of  $\Phi$  and  $Q^T Q = \mathbb{1}$ . Defining  $W = \text{diag}(1/\sqrt{\lambda_1}, 1/\sqrt{\lambda_2}, \dots, 1/\sqrt{\lambda_N})$ , where  $\{\lambda_i, i = 1, \dots, N\}$  are the eigenvalues of  $\Phi$ , one has  $W Q^T \Phi Q W = \mathbb{1}$ . One can therefore sample vectors  $\vec{r}$  from an (N-1)-dimensional normal distribution (unity variance) and obtain the coefficient vectors  $\vec{a}$  by transformation as  $\vec{a} = Q W \vec{r}$ .

We note that the covariance matrix of the coefficients  $a_n$  is the inverse of the matrix  $\Phi$ , i.e.  $\langle a_n a_{n'} \rangle = \Phi_{nn'}^{-1}$ . This is linked to the variation of the transverse position  $t(u)$  in some depth  $u$  among an ensemble of analytically generated trajectories. Assuming for simplicity a pin-like beam impinging perpendicularly ( $\theta_{\text{in}} = 0$ ) onto a box of water at  $t_{\text{in}} = 0$ , one has  $t(u) = \sum_{n=2}^N a_n u^n$ . The average transverse position over many trajectory is zero,  $\langle t(u) \rangle = 0$ , so that the variation of  $t(u)$  is

$$\langle t^2(u) \rangle = \left\langle \left( \sum_{n=2}^N a_n u^n \right) \left( \sum_{n=2}^N a_n u^n \right) \right\rangle = \sum_{n=2, n'=2}^{N, N} \langle a_n a_{n'} \rangle u^{n+n'} = \sum_{n=2, n'=2}^{N, N} \Phi_{nn'}^{-1} u^{n+n'}. \quad (10)$$

It is important to recall that the previously derived expressions correspond to only one of the two scattering planes transverse to the beam axis. A three-dimensional proton trajectory is characterised by a pair of coefficient vectors  $\vec{a}_x$  and  $\vec{a}_y$ , one for each scattering plane and each randomly sampled from the distribution described above.

By construction, the first two coefficients represent the protons' entry position and angle:  $t(u=0) = t_{\text{in}} = a_0$  and  $t'(u) = \theta_{\text{in}} = a_1$ . It is worth noting that the likelihood does not depend on them (summation starts at  $n = 2$  in equation 8) because in a homogeneous medium, the position where a proton enters the volume and its direction does not impact the likelihood of MCS. On the other hand, the coefficients  $a_0$  and  $a_1$  can be drawn from some distribution to generate the desired beam shape. For example, setting  $a_{x,1} = a_{y,1} = 0$  and drawing  $a_{x,0}$  and  $a_{y,0}$  from Gaussian distributions  $G(a_{x,0}) = \exp(-(a_{x,0} - t_{\text{centre},x})^2/2\sigma_x^2)/\sqrt{2\pi}\sigma_x$  and  $G(a_{y,0}) = \exp(-(a_{y,0} - t_{\text{centre},y})^2/2\sigma_y^2)/\sqrt{2\pi}\sigma_y$ , with standard deviation  $\sigma_x = \sigma_y$ , produces a symmetric parallel Gaussian pencil beam of width  $\sigma_{\text{beam}}^2 = \sigma_x^2 + \sigma_y^2$  centred at  $(t_{\text{centre},x}, t_{\text{centre},y})$ .

## 2.4 Derivation of the most likely path

In the context of proton CT, an important concept is the most likely path, i.e. the trajectory which a proton registered by the proton CT scanner has most likely followed through the imaged object. The most sophisticated imaging systems measure positions and propagation angles of individual protons in front of and behind the object (Johnson, 2018). We denote these as  $(t_{\text{in}}, \theta_{\text{in}})$  and  $(t_{\text{out}}, \theta_{\text{out}})$ , respectively. In the conventional mathematical framework (Williams, 2004; Schulte et al., 2008), the MLP is constructed by estimating the most likely position  $t_{\text{MLP}}$  independently at several depths  $u$ , given the measured entry and exit data. The trajectory which connects the so-obtained series of points  $t_{\text{MLP}}(u)$  is the MLP. In the presented work, the MLP is obtained by maximising the likelihood functional  $\mathcal{L}$  (equation 3), or by minimising  $L = -\log \mathcal{L}$ , using the Euler-Lagrange formalism commonly used in physics, e.g. to derive equations of motion. Observing that  $L$  only depends on  $t''$ , but not on  $t$  or  $t'$ , we find for the infinitesimal variation of the Lagrangian (after two partial integrations)

$$\delta L = \int_0^{u_{\text{out}}} \frac{d^2}{du^2} \left( \frac{\partial F}{\partial t''} \right) \delta t \, du + \left[ \delta t' \frac{\partial F}{\partial t''} \right]_0^{u_{\text{out}}} - \left[ \delta t \frac{d}{du} \frac{\partial F}{\partial t''} \right]_0^{u_{\text{out}}} \quad \forall \delta t \in \mathcal{C}^2, \quad (11)$$

where  $\delta t$  and  $\delta t'$  is an arbitrary (continuous and differentiable) test function and its derivative with respect to  $u$ , respectively. We have further defined  $F = (t'')^2/2T$  for the sake of simpler notation. Since the proton's position and angle at  $u = 0$  and  $u = u_{\text{out}}$  are known,  $\delta t$  and  $\delta t'$  vanish there, and the last two terms in equation 11 above are zero. The minimisation requirement that  $\delta L$  be equal to zero for any test function  $\delta t$  implies

$$\left. \frac{d^2}{du^2} \frac{\partial F}{\partial t''} \right|_{t''=t''_{\text{MLP}}} = 0 \quad \Leftrightarrow \quad \left. \frac{d}{du} \frac{\partial F}{\partial t''} \right|_{t''=t''_{\text{MLP}}} = \frac{d}{du} \frac{t''_{\text{MLP}}}{T} = c_1 \quad \Leftrightarrow \quad \frac{t''_{\text{MLP}}}{T} = c_1 u + c_0, \quad (12)$$

where  $c_0$  and  $c_1$  are integration constants. An expression for the MLP can be derived from the right hand side expression in equation 12 above by integration,

$$\theta_{\text{MLP}}(u) = t'_{\text{MLP}}(u) = \theta_{\text{in}} + c_0 \int_0^u dv T(v) + c_1 \int_0^u dv T(v) v \quad (13)$$

$$t_{\text{MLP}}(u) = t_{\text{in}} + \theta_{\text{in}} u + c_0 \int_0^u dv \int_0^v dw T(w) + c_1 \int_0^u dv \int_0^v dw T(w) w, \quad (14)$$

where  $v$  and  $w$  are integration variables. The constants  $c_0$  and  $c_1$  can be determined by imposing the boundary conditions  $\theta_{\text{MLP}}(u = u_{\text{out}}) = \theta_{\text{out}}$  and  $t_{\text{MLP}}(u = u_{\text{out}}) = t_{\text{out}}$ . One finds

$$\begin{aligned} c_0 &= \frac{1}{AD - BC} (-B(t_{\text{out}} - t_{\text{in}} - \theta_{\text{in}} u_{\text{out}}) + D(\theta_{\text{out}} - \theta_{\text{in}})) \\ c_1 &= \frac{1}{AD - BC} (A(t_{\text{out}} - t_{\text{in}} - \theta_{\text{in}} u_{\text{out}}) - C(\theta_{\text{out}} - \theta_{\text{in}})), \end{aligned} \quad (15)$$

with

$$\begin{aligned} A &= \int_0^{u_{\text{out}}} dv T(v); & C &= \int_0^{u_{\text{out}}} dv \int_0^v dw T(w); \\ B &= \int_0^{u_{\text{out}}} dv T(v) v; & D &= \int_0^{u_{\text{out}}} dv \int_0^v dw T(w) w \end{aligned} \quad (16)$$

It is interesting to note that starting from our MLP formulæ (equations 13 and 14), the conventional expressions reported by Williams (2004) and Schulte et al. (2008) can be derived. We refer to appendix A for more details.

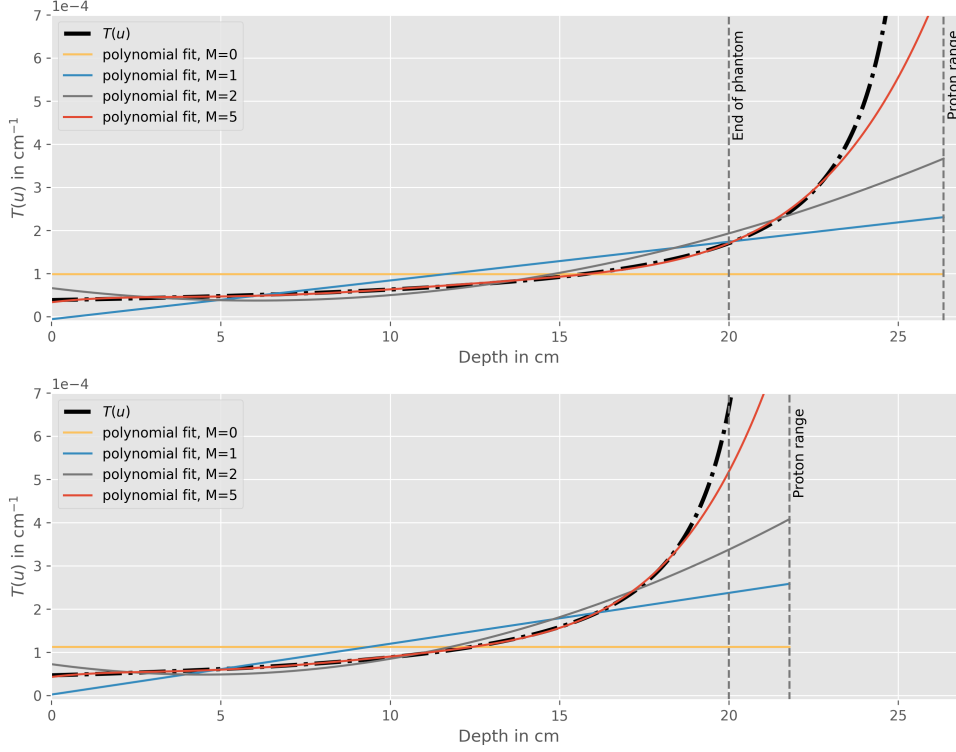


Figure 2: Depth dependent scattering power  $T(u) = E_0^2/\beta^2(u)p^2(u)X_0(u)$  for 200 MeV (upper) and 180 MeV (lower) protons. The dashed vertical line on the right side indicates the proton range at the respective energy.

### 2.5 Approximation: Describing proton energy loss by a polynomial function

In list mode proton CT reconstruction, an MLP needs to be estimated for every single recorded proton (Poludniowski et al., 2015; Johnson, 2018). In its conventional representation (equation 23), several multiplications and additions of several  $2 \times 2$  matrices need to be performed at every depth  $u$  on the reconstruction grid. The  $\Sigma$  and  $R$  matrices themselves can be precomputed and stored. The computational aspect was what motivated other authors to develop and implement an empirical cubic polynomial MLP (Li et al., 2006; Collins-Fekete et al., 2015). The advantage is that relatively fewer multiplication operations need to be performed in that case.

In the following, we show how this can be done rigorously given that the term  $\beta^2 p^2$  in  $T(u)$  is often approximated as a polynomial in terms of depth  $u$  (Williams, 2004; Schulte et al., 2008). The polynomial coefficients are easily obtained by a fit to  $T(u)$  extracted from a Monte Carlo simulation of protons impinging onto a box of water. Following this strategy, we write  $T(u)$  approximately as

$$T(u) \approx \sum_{m=0}^M b_m u^m. \quad (17)$$

An example of this approximation for different orders  $M$  is shown in figure 2 for 180 MeV and 200 MeV protons in water. The fit was performed in a depth range from 0 cm to 20 cm. It is worth observing that the goodness of the fit depends on the depth range of interest relative to the proton range because  $T(u)$  diverges too much at the end for the polynomial to represent it properly. For example, for an object of 20 cm water equivalent thickness, the discrepancy between polynomial and true scattering power is more important for 180 MeV protons (lower panel) than for 200 MeV protons (upper panel), especially beyond 15 cm.

Inserting equation 17 into equation 14, the integrals can be solved analytically and the MLP

becomes a pure polynomial expression of order  $M + 3$ :

$$t_{\text{MLP}} = t_{\text{in}} + \theta_{\text{in}}u + c_0 \sum_{m=0}^M \frac{b_m}{(m+1)(m+2)} u^{m+2} + c_1 \sum_{m=0}^M \frac{b_m}{(m+2)(m+3)} u^{m+3}. \quad (18)$$

The constant pre-factors  $c_0$  and  $c_1$  are calculated according to equation 15 using the following polynomial expressions obtained by inserting equation 17 into equation 16:

$$\begin{aligned} A &= \sum_{m=0}^M \frac{b_m}{m+1} u_{\text{out}}^{m+1}; & C &= \sum_{m=0}^M \frac{b_m}{(m+1)(m+2)} u_{\text{out}}^{m+2}; \\ B &= \sum_{m=0}^M \frac{b_m}{m+2} u_{\text{out}}^{m+2}; & D &= \sum_{m=0}^M \frac{b_m}{(m+2)(m+3)} u_{\text{out}}^{m+3}. \end{aligned} \quad (19)$$

Determining the MLP thus consists in calculating the coefficients of a polynomial of order  $M + 3$  and evaluating it in all depths across the object.

## 2.6 Special case: assuming constant proton energy

In a homogeneous medium, the scattering power  $T$  generally increases with depth because the protons continuously slow down. Neglecting the energy loss, one could replace  $T(u)$  by a constant. In terms of the polynomial expression for  $T(u)$  introduced in section 2.5, only one coefficient  $b_0$  is non-zero then and

$$t_{\text{MLP}} = t_{\text{in}} + \theta_{\text{in}}u + c_0 \frac{b_0}{2} u^2 + c_1 \frac{b_0}{6} u^3 \quad (20)$$

is of strictly cubic order. Such an expression has also been derived by Erdelyi (2009) and used by Li et al. (2006). Clearly, this simplified MLP estimate still matches the measured proton position and propagation angle in front of the object because  $t_{\text{MLP}}(u = 0) = t_{\text{in}}$  and  $\theta_{\text{MLP}}(u = 0) = t'_{\text{MLP}}(u = 0) = \theta_{\text{in}}$ . By construction of the pre-factors  $c_0$  and  $c_1$  (see section 2.4), the same holds true on the exit side, i.e.  $t_{\text{MLP}}(u = u_{\text{out}}) = t_{\text{out}}$  and  $\theta_{\text{MLP}}(u = u_{\text{out}}) = t'_{\text{MLP}}(u = u_{\text{out}}) = \theta_{\text{out}}$ .

## 2.7 Computational efficiency of MLP estimation

How computationally intensive the polynomial expression in equation 18 is to evaluate depends on the number of arithmetic operations needed to calculate the most likely transverse position  $t_{\text{MLP}}$  in every depth  $u$  required by a reconstruction algorithm. In this light, it is convenient to combine common powers in the two sums in equation 18 into a single sum. After an index shift ( $n = m + 3$ ), one obtains

$$\begin{aligned} t_{\text{MLP}} &= t_{\text{in}} + \theta_{\text{in}}u + c_0 \frac{b_0}{2} u^2 + \sum_{m=0}^{M-1} \frac{c_0 b_{m+1} + c_1 b_m}{(m+2)(m+3)} u^{m+3} + c_1 \frac{b_M}{(M+2)(M+3)} u^{M+3} \\ &= \sum_{n=0}^{N=M+3} d_n u^n = d_0 + u(d_1 + u(d_2 + \dots u(d_{N-1} + d_N u) \dots)), \end{aligned} \quad (21)$$

with

$$\begin{aligned} d_0 &= t_{\text{in}}; & d_1 &= \theta_{\text{in}}; & d_2 &= \frac{c_0 b_0}{2} \\ d_{n, (3 \leq n \leq M+2)} &= \frac{c_0 b_{n-2} + c_1 b_{n-3}}{n(n-1)}; & d_{N=M+3} &= c_1 \frac{b_M}{(M+2)(M+3)}. \end{aligned} \quad (22)$$

The polynomial coefficients  $b_m$  can be precomputed once before the reconstruction and the prefactors  $c_0$  and  $c_1$  (equation 15) need to be calculated only once per proton. Consequently, also the coefficients  $d_n$  need to be calculated only once per proton. Evaluating the expanded polynomial on the right hand side of equation 21 above amounts to  $M + 3$  multiplications and  $M + 3$  additions, i.e.  $2M + 6$  floating point operations per proton and depth  $u$ . For example, only 6 or 8 floating point operations are required per depth  $u$  if the MLP is approximated by a polynomial of order  $M = 0$  or  $M = 1$ , respectively. On the other hand, evaluating the conventional MLP expression of Schulte et al. (2008) (equation 23) for  $M = 5$  results in at least 47 floating point operations (McAllister, 2009), even if sparsity of the matrices  $R_0$  and  $R_1$  is exploited.

## 2.8 Verification with Monte Carlo simulations

We performed Monte Carlo simulations to verify the developed analytical methods presented in this paper using the Geant4/Gate toolkit (Agostinelli et al., 2003; Sarrut et al., 2014). A pin-like beam of mono energetic protons was simulated impinging onto a homogeneous box of water of  $50 \times 50 \text{ cm}^2$  lateral dimension and 20 cm depth. The maximum step size was set to 1 mm and each proton’s three dimensional position was stored at every step. The *emstandard* physics list was used to consider electromagnetic interactions only and disregard nuclear scattering as this is not included in the analytical models.

We first quantified how accurate it is in general to approximate a proton trajectory resulting from MCS by a polynomial function of depth. To this end, we fitted polynomials of different orders to each simulated proton trajectory and determined the RMSE between the so approximated transverse position and the true one.

Furthermore, we quantified how closely the analytically generated trajectories would resemble those simulated with Monte Carlo by two means: first, we compared the distribution of transverse proton positions at some given depth and the beam spread as a function of depth. Second, we fitted polynomials to all simulated trajectories (see above) and compared the distribution of the fit coefficients with the analytically generated coefficient distributions (see section 2.3).

To compare the accuracy of our alternative MLP formulation with the conventional one (Schulte et al., 2008) for different polynomial degrees (see section 2.5), an MLP was calculated for each proton history based on the entry and exit position and angle. The root mean square error (RMSE) between the MLP estimate and the true transverse position as a function of depth was calculated over an ensemble of  $10^4$  protons.

## 2.9 Tomographic reconstruction using the alternative MLP formulæ

We reconstructed proton CT images of the CTP528 high resolution module of the Catphan phantom (The Phantom Laboratory, Salem, NY) to test our alternative MLP formulæ presented in this work. Input data were simulated with Geant4/Gate using a mono energetic 200 MeV point source at 100 cm distance from iso-centre. Each proton’s position and propagation angle were recorded in two tracker planes upstream and downstream from the phantom, respectively. The downstream tracker also recorded the proton’s residual energy as would a calorimeter in a real proton CT scanner. The *QGSP-BIC* physics list was used and 3-sigma cuts applied to the exit energy and angular distribution to filter out protons which have undergone nuclear scattering (Schulte et al., 2005) to reflect how experimental data are typically processed. The source was rotated around the iso centre in 0.5 degree steps leading to 720 projections. A total of  $2.88 \times 10^8$  primary protons were simulated corresponding to  $225 \text{ protons} \cdot \text{mm}^{-2} \cdot \text{projection}^{-1}$ . Tomographic images with  $0.1 \times 0.1 \text{ mm}^2$  pixel size were reconstructed using a distance-driven binning Feldkamp-Davis-Kress algorithm (Rit et al., 2013), where the MLP intervenes during the binning step.

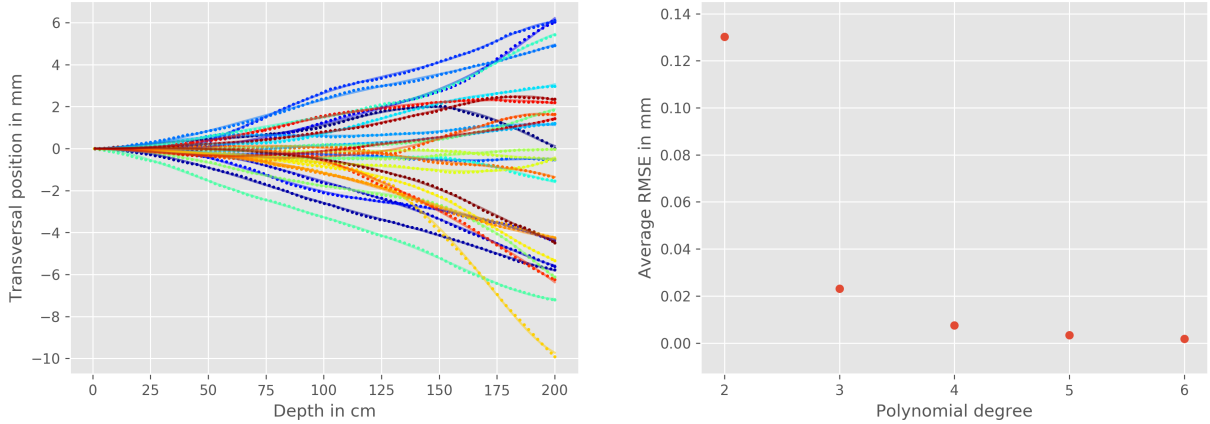


Figure 3: Left: Proton trajectories simulated with Monte Carlo (dots) and order 6 polynomial fits to each of them (solid lines). Right: RMSE averaged over depth and over the ensemble of  $10^5$  proton trajectories.

### 3 RESULTS

#### 3.1 Analytical sampling of proton trajectories

Figure 3 left shows an example of 20 proton trajectories simulated at 200 MeV in water as described in section 2.8. The dots represent the transverse positions registered by the Monte Carlo simulation. The solid lines are polynomial fits, in this case of order  $N = 6$ . The lowest two coefficients are fixed to  $a_0 = t_{\text{in}} = 0$  and  $a_1 = \theta_{\text{in}} = 0$  in accordance with the proton source geometry. The right panel shows the average RMSE between the simulated trajectories and the polynomial fits (see section 2.8). Empirically, we found that the RMSE values vary approximately according to a Poisson distribution among the ensemble of  $10^5$  trajectories so that the uncertainties to be associated with the data points have the same magnitude as the RMSE values themselves.

The left panels in figure 4 show sets of 50 example trajectories analytically generated using the procedure detailed in section 2.3 with polynomial orders  $N = 3$  and  $N = 6$ , respectively. All had the same initial position  $t_{\text{in}} = 0$  and angle  $\theta_{\text{in}} = 0$ . The colour coding represents each trajectory's likelihood. The right panel in figure 4 is essentially a cross-section through the left plot at depth  $u = 10$  cm where each green dot corresponds to one trajectory and its likelihood is represented along the y-axis. The red histogram was obtained by binning the transverse positions of the trajectories. Note that for better visibility, only  $10^4$  trajectories are shown as dots while position information from  $10^5$  trajectories were used to generate the histograms. The blue histogram was obtained based on Monte Carlo simulated proton trajectories such as the ones shown in figure 3 left. The widths of the dashed black Gaussian distributions are calculated analytically according to equation 25.

Figure 5 shows polynomial coefficients obtained from fits to Monte Carlo simulated proton trajectories (see section 2.8) and from analytical sampling (see section 2.3). A polynomial trajectory was fitted to each simulated trajectory as shown in figure 3 yielding a coefficient vector  $\vec{a}$ . Each dot in the scatter plots corresponds to a pair of components of one vector  $\vec{a}$ , i.e., of one simulated trajectory. The histograms in the upper row show marginalised distributions of one coefficient only. It is interesting to observe that these latter are appreciably broader for the Monte Carlo data than for the analytically generated polynomials, i.e.  $\langle a_n^2 \rangle$  is larger, while the spatial distributions shown in figure 4 match well. The reason is that the coefficients from Monte Carlo are also more anti-correlated ( $\langle a_n a_{n'} \rangle < 0$ ) than the analytical ones which lowers the spatial variation  $\langle t^2(u) \rangle$  (see equation 10).

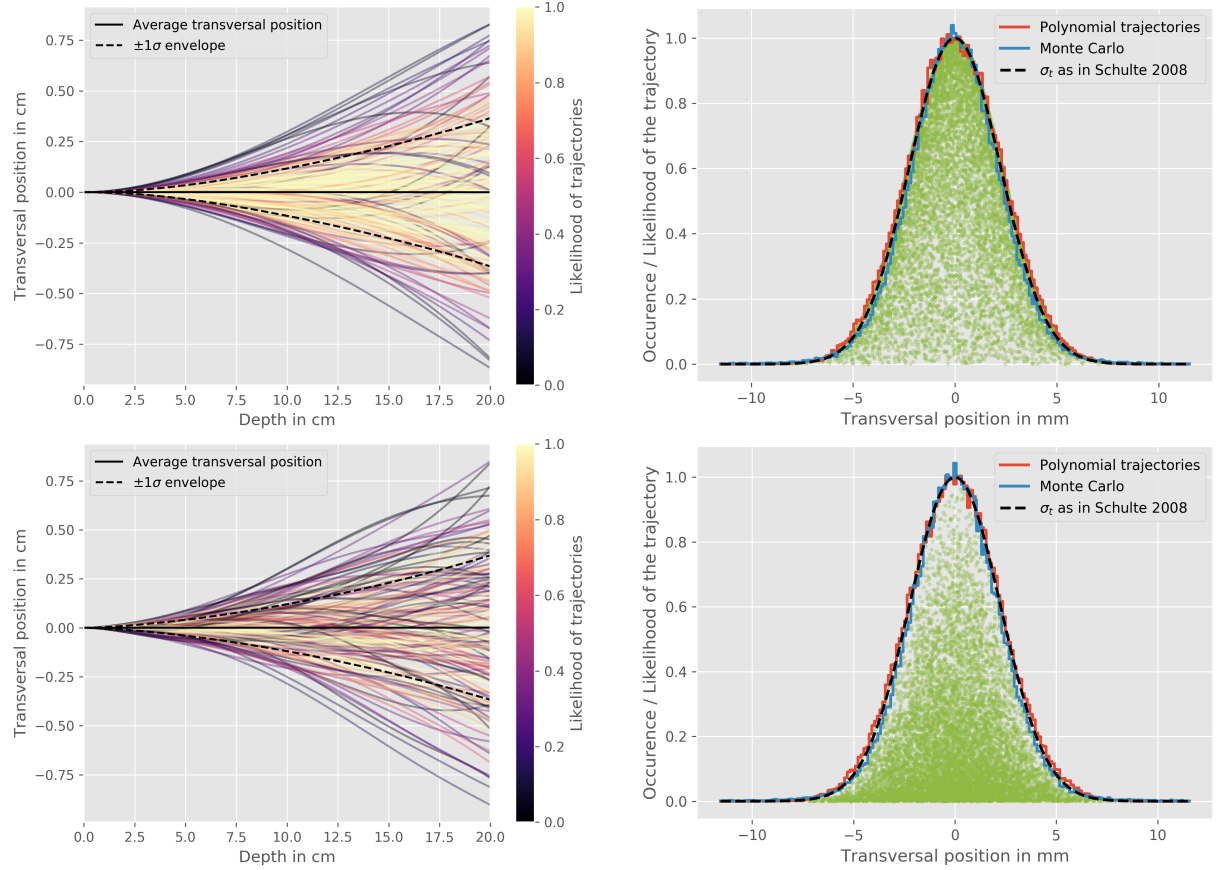


Figure 4: Left: Analytically generated proton trajectories (section 2.3) using a polynomial order  $N = 3$  (upper row) and  $N = 6$  (lower row). The entry energy was assumed to be 200 MeV. Right: distribution of transverse positions at 10 cm depth obtained from analytical trajectories and those simulated with Monte Carlo under identical conditions. Each green dot corresponds to one trajectory in the left panel.



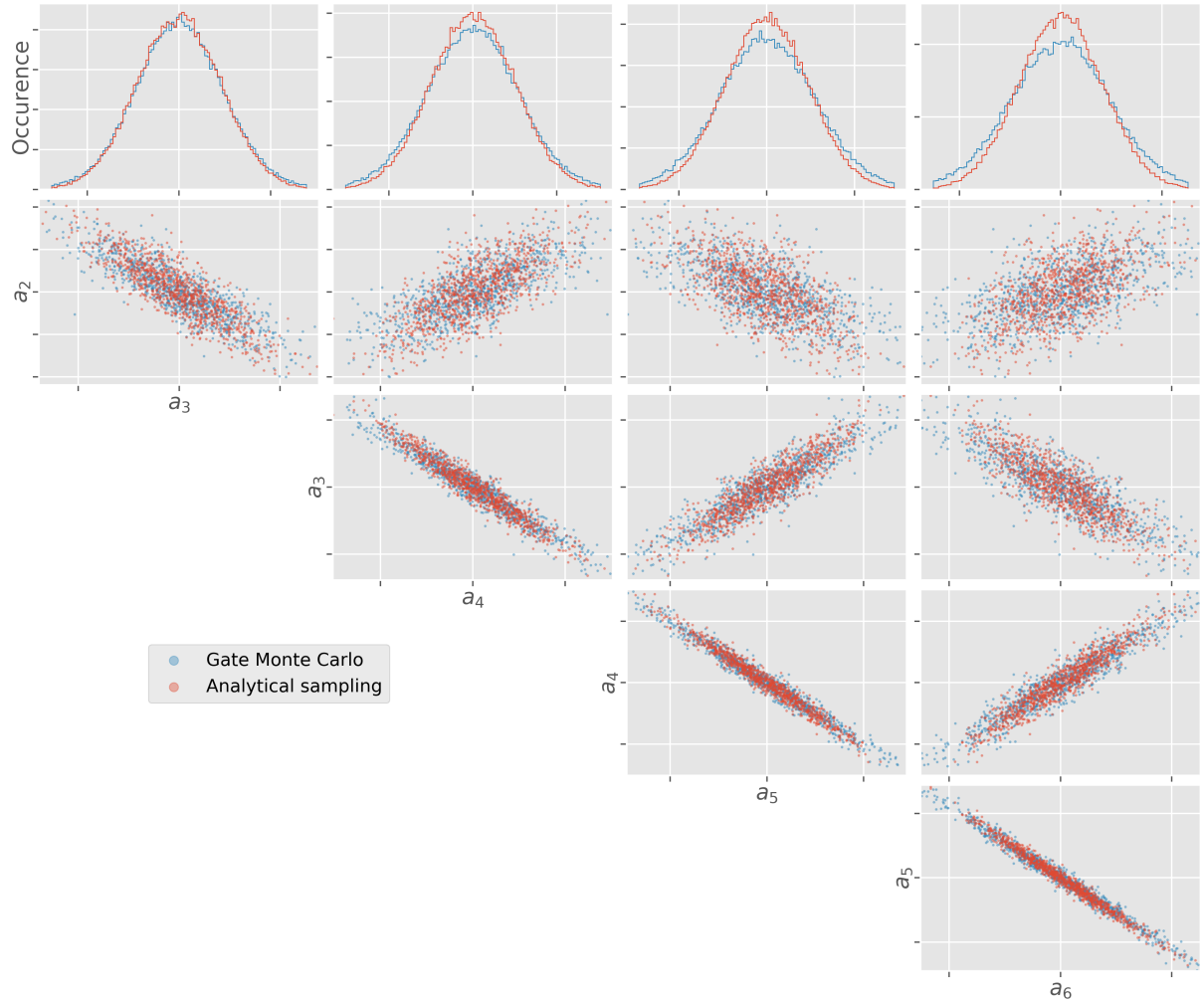


Figure 5: Pair-wise distribution of polynomial coefficients and marginalised distributions (upper row). Blue dots were obtained by fitting polynomials of order  $N = 6$  to Monte Carlo simulated trajectories (see figure 3). Red dots were generated analytically (see section 2.3) using the same polynomial degree.

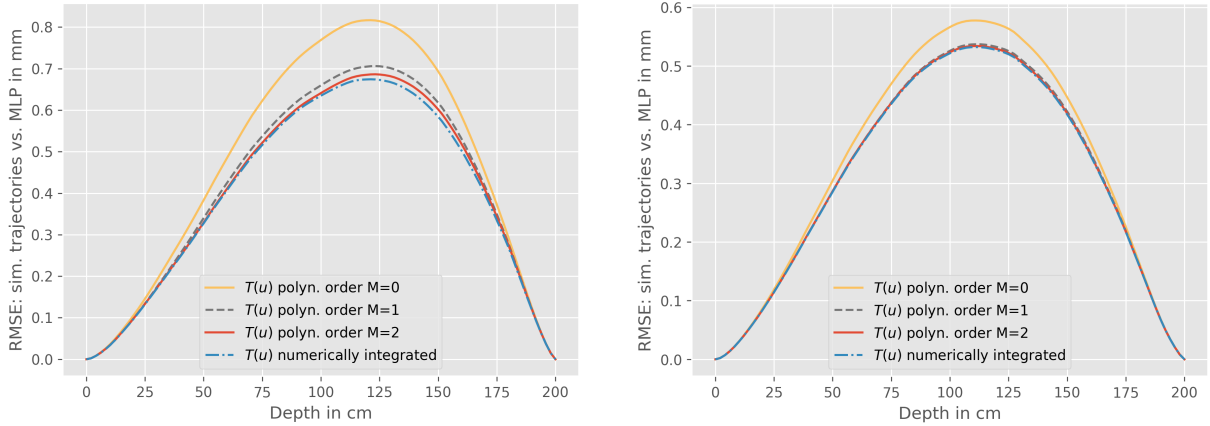


Figure 6: RMSE between simulated proton trajectories and different MLP estimates as a function of depth in 20 cm water phantom for 180 MeV (left) and 200 MeV (right) protons. For the yellow, grey, and red curves, we used our polynomial expression (equation 21) and for the blue dash-dotted one, we used the expression from Schulte et al. (2008) integrating the scattering power  $T(u)$  numerically without any polynomial approximation.

### 3.2 MLP estimation in proton CT reconstruction

Figure 6 shows the RMSE between proton trajectories simulated with Monte Carlo and several MLP estimates: on one hand using a polynomial approximation to the scattering power  $T(u)$  (equation 17) and consequently the polynomial expression for the MLP (equation 18) for different orders  $M$ ; and on the other hand using the general MLP expression (equation 14) and integrating the  $T(u)$  function numerically. The boundary conditions  $t_{\text{in}}$ ,  $t_{\text{out}}$ ,  $\theta_{\text{in}}$ , and  $\theta_{\text{out}}$ , were obtained from the Monte Carlo data. A total of  $10^5$  trajectories were simulated and for each of them the MLP was estimated. For  $M = 5$  (not shown), the RMSE is indistinguishable from the one obtained when integrating  $T(u)$  numerically. Likewise, the RMSE obtained with the conventional MLP (not shown) is identical to the results for  $M = 5$ . For 200 MeV protons (right panel), already  $M = 1$  yields a RMSE as good as for higher orders. For 180 MeV (left panel) on the other hand, the RMSE remains slightly higher even for  $M = 1$  or  $M = 2$  because the polynomial approximation of  $T(u)$  is less accurate in this case (see figure 2).

### 3.3 Reconstructed proton CT images

Figure 7 shows reconstruction results of the Catphan phantom obtained with different MLP estimates at 200 MeV and 180 MeV. On one hand, we used the expression given in equation 21 and approximated the scattering power  $T(u)$  function with polynomials of varying degrees (see figure 2). On the other hand, we estimated the MLP through the expression from Schulte et al. (2008) approximating  $T(u)$  by a fifth order polynomial. The root mean squared difference between the proton CT images using our polynomial MLP estimate and Schulte's is about 3.2 % for  $N = 5$  and 4.5 % for  $N = 0$ , compared to the noise level of the images of about 3.3 % (upper row). A faint ring-shaped structure is visible in the difference image between  $N = 5$  and  $N = 0$  (upper right) where the inserts are located, which is to be attributed to the slightly more blurry line patterns in the latter case. This is coherent with the detail images in the centre row which are a bit more blurred for  $N = 0$  than when using higher order approximations of  $T(u)$ . At 200 MeV (lower row), the line patterns appear almost identical in all cases.

Reconstruction was performed on a MacBook Pro (2017) on a single Intel i7 processor. MLP estimation took 2.7 s per projection on average (about 68000 protons per projection) using Schulte's expression. This amounts to 32 min for the 720 projections. Evaluating our polynomial expression with order 1 took 0.1 s per projection on average, or 72 s for all 720 projections. This

is essentially negligible compared to the time needed for binning data into projection images, reconstructing, and reading/writing data. The gain in computation speed is more than what would be expected by a pure comparison of floating point operations probably because the compiler is able to optimize the code more efficiently given the simpler structure of the mathematical operations (see section 2.7). We underline that similar optimisation may be possible for Schulte’s expression and the reported computation speeds should serve as indications only.

## 4 DISCUSSION

The likelihood functional introduced in this work provides a way to generate proton trajectories in medium as polynomial functions as explained in section 2.3. The medium is required to be homogeneous in the direction transverse to the beam, but may show heterogeneities along the beam, as we will discuss further down. Each polynomial proton trajectory is represented by an  $N + 1$  dimensional coefficient vector, where  $N$  is the polynomial order. The first two coefficients represent the entry positions and angle of the proton. The method is computationally very efficient because it only requires sampling the coefficient vectors according to a normal distribution followed by a multiplication with an eigenvector matrix for each coefficient vector. A limitation of the method is that nuclear interactions are neglected. In this sense, analytically generated projections correspond to Geant4 Monte Carlo, e.g., with the *emstandard* physics list, i.e., electromagnetic interactions only.

One potential application of the trajectory sampling method could be the analytical generation of proton CT data without running a full Monte Carlo simulation. Projection images (containing water equivalent thickness) could be obtained by integrating relative stopping power maps along trajectories by ray tracing methods. The technical effort necessary to implement such an analytical proton CT simulation as in Gianoli et al. (2019) was beyond the scope of this work. Another application of the analytical trajectory generation could be in Monte Carlo treatment planning systems. The most abundant process when transporting a proton through a patient geometry is MCS. Instead of performing this explicitly by successively deflecting the proton path by small angles, the overall MCS contribution to the trajectory could be sampled analytically. Discrete nuclear events would need to be simulated additionally and protons towards their end of range would potentially need to be transported explicitly. Developing such methods was beyond the scope of this work, but would deserve further investigation in our opinion.

Figure 4 confirms that the spatial distribution of analytically sampled trajectories resembles closely the one generated by Monte Carlo. Interestingly, if the polynomial degree  $N$  is low, such as in the upper right panel where  $N = 3$ , the generated trajectories are intrinsically constrained to be relatively smooth and tend to have a high likelihood especially if concentrated around the central axis of the beam. On the other hand, with  $N = 6$  (lower right panel), the polynomials have more freedom to generate small scale curvature so that their likelihood is lower on average and less correlated with the position relative to the beam centre. Furthermore, figure 5 shows that the distribution of coefficients obtained by fitting polynomials to Monte Carlo trajectories resembles closely the analytical  $(N - 1)$ -dimensional Gaussian distributions. The higher order coefficients ( $a_5, a_6$ ) obtained from Monte Carlo spread to slightly larger values than the analytically generated ones. We note in this context that the angular distribution due to MCS is actually not purely Gaussian, but contains a Rutherford-like component leading to higher tails (Scott, 1963) which are expected to generate more small scale curvature (larger high order coefficients). It is also known that the MCS models implemented in Geant4 (and other Monte Carlo codes) bear some uncertainty compared to theoretical MCS models (Makarova et al., 2017). In any case, the parametrisation of the scattering power  $T(u)$  could be refined to optimise the match between analytical and Monte Carlo trajectories. For this work, we aligned the definition of  $T(u)$  with the ones used in other contributions on MLP estimation.

We derived an expression for the MLP by maximising a likelihood functional using the Euler-

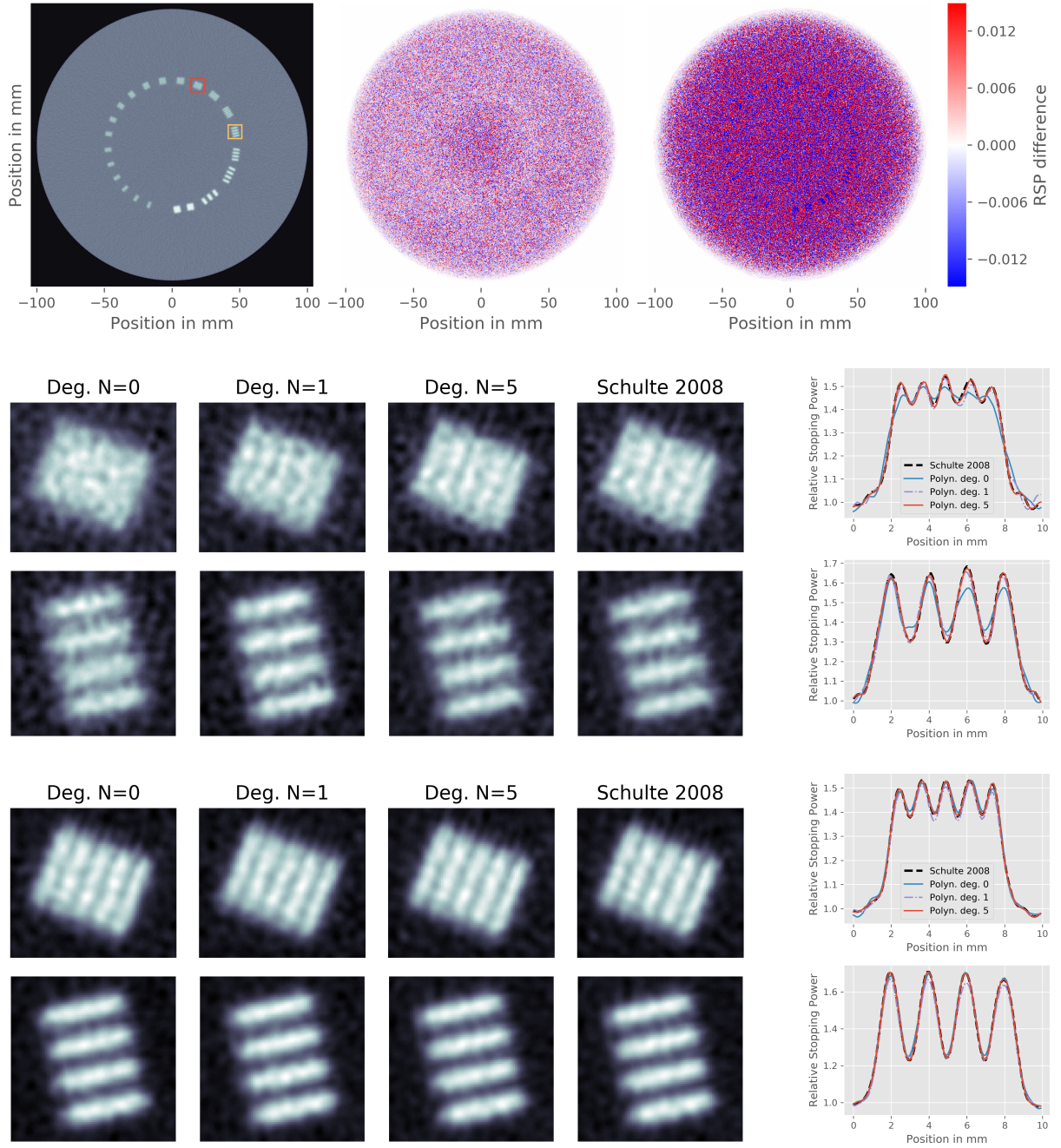


Figure 7: Upper row: Simulated proton CT image (left) of the Catphan phantom at 180 MeV energy reconstructed using equation 21 and a fifth order polynomial to fit the scattering power  $T(u)$ ; difference to the proton CT image obtained with the MLP expression from Schulte et al. (2008) (centre) and using an order  $M = 0$  polynomial to approximate  $T(u)$  (right). Center (180 MeV) and lower (200 MeV) row: Details of the phantom indicated by red and yellow boxes in the upper left panel; intensity profiles across the two line patterns averaged along the line direction.

Lagrange formalism. The resulting general expression (equation 14) is strictly identical to the conventional ones in Williams (2004) and Schulte et al. (2008) (equation 23), except for the logarithmic pre-factor in Schulte’s expression for the integrated scattering power (see section A). In fact, it would be possible to derive our MLP starting from the conventional equation 23. When the function  $T(u)$  is approximated by a polynomial, which is common practice in proton CT reconstruction, the MLP expression automatically becomes a polynomial (equation 18). The order  $M$  of the  $T(u)$  polynomial determines the order  $N = M + 3$  of the MLP polynomial. For  $M = 5$ , as used by Williams (2004) and Schulte et al. (2008), the proton CT images of the Catphan phantom reconstructed with our MLP and with the conventional expression are virtually identical (figure 7). The residual noise is mainly due to the additional pre-factors in Schulte’s MLP formula which are not present in ours (see equations 25 to 27). There are also slight numerical differences in the polynomial fits to the  $1/\beta^2 p^2(u)$  function used in Schulte’s formalism and to the scattering power  $T(u)$  used in our method, respectively.

A cubic MLP, such as used by Li et al. (2006), is equivalent to  $M = 0$  and thus implies neglecting energy loss in the medium. Inside the medium, the cubic MLP expression yields a less accurate estimate than an MLP estimate with  $M > 0$  (figure 6). Accordingly, the reconstructed Catphan images have a lower spatial resolution (visually assessed), although only slightly. This is coherent with the results reported by Li et al. (2006). Collins-Fekete et al. (2015) also estimate the MLP with a cubic polynomial, however only impose the constraints on entry and exit position,  $t_{\text{MLP}}(u = 0) = t_{\text{in}}$  and  $t_{\text{MLP}}(u = u_{\text{out}}) = t_{\text{out}}$ . The propagation angles at  $\theta_{\text{MLP}}(u = 0)$  and  $\theta_{\text{MLP}}(u = u_{\text{out}})$  are scaled through an empirical Monte Carlo based calibration.

In view of the methods reported in this work, an alternative approach would be to select the desired order  $M$  and obtain  $N$  as a consequence rather than fixing it a priori. The choice of  $M$  generally depends on the water equivalent thickness of the object compared to the proton range at the beam energy used for image acquisition (see figure 2). Our results (see figure 7) suggest that  $M = 0$  would be sufficient for a 20 cm (water equivalent) thick object if the proton range is 26 cm (200 MeV), but that  $M = 1$  (or higher) would be preferable when the range is only 21.7 cm (180 MeV). We note that ideal trackers were used in the simulation and that the measurement uncertainty of a real proton CT scanner would additionally degrade the image resolution (Krah et al., 2018).

Furthermore, we evaluated the scattering power  $T(u)$  only in water in this study so that the analytically generated trajectories as well as the MLP formulae are strictly valid only in water. The same is true in conventional MLP estimation methods used in proton CT reconstruction. We underline, however, that material heterogeneities along the beam path, i.e. which are mainly a function depth  $u$ , could be easily considered. An example could be protons passing through different slabs of soft tissue and bone. To this end, the protons’ energy loss would need to be calculated in the heterogeneous geometry to adjust the  $\beta^2 p^2$ -term in  $T(u)$  and the  $X_0(u)$  factor would need to reflect the varying radiation length of the material, similar to what has been suggested by (Collins-Fekete et al., 2017). Such an approach would provide a way to sample proton trajectories and estimate an MLP in media with the above described heterogeneities. Clearly, this would require some prior knowledge about the imaged object.

If variation of material properties in depth are considered by appropriately adjusting the  $\beta^2 p^2$ -term in  $T(u)$  and the radiation length  $X_0(u)$ , the polynomial degree  $M$  will likely depend on the degree of heterogeneity. The impact of material heterogeneities on the MLP has been studied elsewhere (Wong et al., 2009; Khellaf et al., 2019) and was not the subject of this work.

## 5 CONCLUSION

We proposed a functional to quantify the likelihood of proton trajectories under the influence of multiple Coulomb scattering and achieved two things: first, the likelihood functional provides an efficient way to analytically sample proton trajectories in homogeneous media instead of gen-

erating them through a Monte Carlo simulation. Second, maximising the functional yields an expression for the most likely path of a proton which can be used in proton computed tomography reconstruction. When the protons' energy loss is approximated by a polynomial function of depth, the most likely path is a polynomial itself. Our most likely path expression is analytically equivalent to the conventional ones available in the literature. Our results suggest that it is sufficient to describe the energy loss with a linear or at most quadratic function of depth instead of a fifth order polynomial. Treating the proton energy as constant leads to a slightly degraded spatial resolution. Evaluating our most likely path expression requires fewer floating point operations than the conventional ones.

## Acknowledgements

The work of Nils Krah was supported by funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 753370. This work was performed within the framework of the project DIC20161236452 of the "Fondation pour la Recherche Médicale" (FRM), of the SIRIC LYriCAN Grant INCa\_INSERM\_DGOS\_12563 and of the LABEX PRIMES (ANR-11-LABX-0063) of Université de Lyon, within the program "Investissements d'Avenir" (ANR-11-IDEX-0007) operated by the ANR.

## A Relation with conventional MLP expression

We summarize the steps to link our MLP expression (equation 14) with the ones reported by Williams (2004) and Schulte et al. (2008), which we state here for completeness:

$$y_{\text{MLP}} = (t_{\text{MLP}}, \theta_{\text{MLP}})^T = (\Sigma_1^{-1} + R_1^T \Sigma_2^{-1} R_1)^{-1} \cdot (\Sigma_1^{-1} R_0 y_{\text{in}} + R_1^T \Sigma_2^{-1} y_{\text{out}}), \quad (23)$$

where

$$\Sigma_1 = \begin{pmatrix} \sigma_{t_1}^2 & \sigma_{t_1 \theta_1}^2 \\ \sigma_{t_1 \theta_1}^2 & \sigma_{\theta_1}^2 \end{pmatrix}, \quad \Sigma_2 = \begin{pmatrix} \sigma_{t_2}^2 & \sigma_{t_2 \theta_2}^2 \\ \sigma_{t_2 \theta_2}^2 & \sigma_{\theta_2}^2 \end{pmatrix}, \quad R_0 = \begin{pmatrix} 1 & u - u_{\text{in}} \\ 0 & 1 \end{pmatrix}, \quad R_1 = \begin{pmatrix} 1 & u_{\text{out}} - u \\ 0 & 1 \end{pmatrix}. \quad (24)$$

The matrices  $R_0$  and  $R_1$  are essentially small angle rotation matrices and  $\Sigma_1$  and  $\Sigma_2$  quantify the spatial and angular dispersion in a certain depth due to MCS. Their components are calculated through the following integral expressions:

$$\sigma_{t_1}^2 = \left(1 + 0.038 \ln \frac{u - u_{\text{in}}}{X_0}\right)^2 \times \int_{u_{\text{in}}}^u T(v) (v - u_{\text{in}})^2 dv \quad (25)$$

$$\sigma_{\theta_1}^2 = \left(1 + 0.038 \ln \frac{u - u_{\text{in}}}{X_0}\right)^2 \times \int_{u_{\text{in}}}^u T(v) dv \quad (26)$$

$$\sigma_{\theta_1 t_1}^2 = \left(1 + 0.038 \ln \frac{u - u_{\text{in}}}{X_0}\right)^2 \times \int_{u_{\text{in}}}^u T(v) (v - u_{\text{in}}) dv. \quad (27)$$

One obtains the components of  $\Sigma_2$  by replacing  $u_{\text{in}}$  with  $u$  and  $u$  with  $u_{\text{out}}$  in the equations above. We recall that we have chosen  $u_{\text{in}} = 0$  for simplicity in our work. The quadratic pre-factor containing the logarithmic term is reported only by Schulte et al. (2008), not by Williams (2004).

To derive the relation between our MLP formulæ and the conventional one, we observe that the integral expressions in equations 13 and 14 can be linked to the conventional ones (equations 25, 26, 27; neglecting the logarithmic pre-factor) as

$$\int_0^u dv T(v) = \sigma_{\theta_1}^2 \quad (28)$$

$$\int_0^u dv T(v)v = (u - u_{\text{in}})\sigma_{\theta_1}^2 - \sigma_{\theta_1 t_1}^2 \quad (29)$$

$$\int_0^u dv \int_0^v dw T(w) = \sigma_{\theta_1 t_1}^2 \quad (30)$$

$$\int_0^u dv \int_0^v dw T(w)w = \sigma_{\theta_1 t_1}^2 (u - u_{\text{in}}) - \sigma_{t_1}^2 \quad (31)$$

and that the terms in equation 16 can be expressed as

$$A = \sigma_{\theta_1}^2 + \sigma_{\theta_2}^2 \quad (32)$$

$$B = (u - u_{\text{in}})\sigma_{\theta_1}^2 - \sigma_{\theta_1 t_1}^2 + (u_{\text{out}} - u)\sigma_{\theta_2}^2 - \sigma_{\theta_2 t_2}^2 + (u - u_{\text{in}})\sigma_{\theta_2}^2 \quad (33)$$

$$C = (u_{\text{out}} - u)\sigma_{\theta_1}^2 + \sigma_{\theta_1 t_1}^2 + \sigma_{\theta_2 t_2}^2 \quad (34)$$

$$D = (u_{\text{out}} - u)(u - u_{\text{in}})\sigma_{\theta_1}^2 - (u_{\text{out}} - u)\sigma_{\theta_1 t_1}^2 + (u - u_{\text{in}})\sigma_{\theta_1 t_1}^2 - \sigma_{t_1}^2 - \sigma_{t_2}^2 + (u_{\text{out}} - u_{\text{in}})\sigma_{\theta_2 t_2}^2 \quad (35)$$

Inserting these relations into equations 13, 14, and 15, and after some matrix manipulations, one recovers equation 23. This means that the conventional MLP, although constructed as succession of most likely positions, is actually the trajectory which minimises our likelihood functional (equation 3). The only slight difference to Schulte et al. (2008) is the logarithmic pre-factor in equations 25, 26, and 27.

## References

- Agostinelli, S. et al. (2003). “GEANT4 - A simulation toolkit”. In: *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3, pp. 250–303. ISSN: 01689002. DOI: 10.1016/S0168-9002(03)01368-8.
- Civinini, C. et al. (2013). “Recent results on the development of a proton computed tomography system”. In: *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 732, pp. 573–576. ISSN: 01689002. DOI: 10.1016/j.nima.2013.05.147.
- Collins-Fekete, C.-A. et al. (2015). “Developing a phenomenological model of the proton trajectory within a heterogeneous medium required for proton imaging”. In: *Physics in Medicine and Biology* 60.13, pp. 5071–5082. ISSN: 0031-9155. DOI: 10.1088/0031-9155/60/13/5071.
- Collins-Fekete, C.-A. et al. (2016). “A theoretical framework to predict the most likely ion path in particle imaging”. In: *Physics in Medicine & Biology* 62.October, pp. 1777–1790.
- Collins-Fekete, C.-A. et al. (2017). “Extension of the Fermi-ÅŒyges most-likely path in heterogeneous medium with prior knowledge information”. In: *Physics in Medicine & Biology* 62.24, pp. 9207–9219. ISSN: 1361-6560. DOI: 10.1088/1361-6560/aa955d.
- Erdelyi, B. (2009). “A comprehensive study of the most likely path formalism for proton-computed tomography”. In: *Phys. Med. Biol. Phys. Med. Biol* 54.54, pp. 6095–6095. ISSN: 0031-9155. DOI: 10.1088/0031-9155/54/20/005.
- Gianoli, C. et al. (2019). “Analytical simulator of proton radiography and tomography for different detector configurations”. In: *Physica Medica* 59, pp. 92–99. ISSN: 11201797. DOI: 10.1016/j.ejmp.2019.03.002.
- Gottschalk, B. et al. (1993). “Multiple Coulomb scattering of 160 MeV protons”. In: *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* 74.4, pp. 467–490. ISSN: 0168583X. DOI: 10.1016/0168-583X(93)95944-Z.



- Highland, V. L. (1975). “Some practical remarks on multiple scattering”. In: *Nuclear Instruments and Methods* 129.2, pp. 497–499.
- Johnson, R. P. (2018). “Review of medical radiography and tomography with proton beams”. In: *Reports on Progress in Physics* 81.1, p. 016701. ISSN: 0034-4885. DOI: 10.1088/1361-6633/aa8b1d.
- Khellaf, F. et al. (2019). “Effects of transverse heterogeneities on the most likely path of protons”. In: *Physics in Medicine & Biology* 64.6, p. 065003. ISSN: 1361-6560. DOI: 10.1088/1361-6560/ab02a8.
- Krah, N et al. (2018). “A comprehensive theoretical comparison of proton imaging set-ups in terms of spatial resolution”. In: *Physics in Medicine & Biology* 63.13, p. 135013. ISSN: 1361-6560. DOI: 10.1088/1361-6560/aaca1f.
- Li, T. et al. (2006). “Reconstruction for proton computed tomography by tracing proton trajectories: A Monte Carlo study”. In: *Medical Physics* 33.3, pp. 699–706. ISSN: 00942405. DOI: 10.1118/1.2171507.
- Lynch, G. R. et al. (1991). “Approximations to multiple Coulomb scattering”. In: *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* 58.1, pp. 6–10. ISSN: 0168583X. DOI: 10.1016/0168-583X(91)95671-Y.
- Makarova, A. et al. (2017). “Comparison of Geant4 multiple Coulomb scattering models with theory for radiotherapy protons”. In: *Physics in Medicine and Biology* 62.15, pp. 5959–5974. ISSN: 13616560. DOI: 10.1088/1361-6560/aa6ce3.
- McAllister, S. A. (2009). “Efficient Proton Computed Tomography Image Reconstruction using General Purpose Graphics Processing Units”. PhD thesis. California State University. DOI: 10.1.1.220.2543.
- Parodi, K. (2014). “Heavy ion radiography and tomography”. In: *Physica Medica* 30.5, pp. 539–543. ISSN: 1724191X. DOI: 10.1016/j.ejmp.2014.02.004.
- Penfold, S. N. et al. (2011). “Geometrical optimization of a particle tracking system for proton computed tomography”. In: *Radiation Measurements* 46.12, pp. 2069–2072. ISSN: 13504487. DOI: 10.1016/j.radmeas.2011.04.032.
- Penfold, S. et al. (2015). “Techniques in Iterative Proton CT Image Reconstruction”. In: *Sensing and Imaging* 16.1, p. 19. ISSN: 1557-2064. DOI: 10.1007/s11220-015-0122-3.
- Poludniowski, G et al. (2014). “Proton computed tomography reconstruction using a backprojection-then-filtering approach”. In: *Physics in Medicine & Biology* 59.24, p. 7905.
- Poludniowski, G. et al. (2015). “Proton radiography and tomography with application to proton therapy”. In: *The British Journal of Radiology* 88.1053, p. 20150134. ISSN: 0007-1285. DOI: 10.1259/bjr.20150134.
- Rit, S. et al. (2013). “Filtered backprojection proton CT reconstruction along most likely paths.” In: *Medical physics* 40.3, p. 031103. ISSN: 0094-2405. DOI: 10.1118/1.4789589.
- Sarrut, D. et al. (2014). “A review of the use and potential of the GATE Monte Carlo simulation code for radiation therapy and dosimetry applications”. In: *Medical Physics* 41.6Part1, p. 064301. ISSN: 00942405. DOI: 10.1118/1.4871617.
- Scaringella, M. et al. (2014). “A proton Computed Tomography based medical imaging system”. In: *Journal of Instrumentation* 9.12, pp. C12009–C12009. ISSN: 1748-0221. DOI: 10.1088/1748-0221/9/12/C12009.
- Schulte, R. et al. (2004). “Conceptual design of a proton computed tomography system for applications in proton radiation therapy”. In: *IEEE Transactions on Nuclear Science* 51.3, pp. 866–872. ISSN: 0018-9499. DOI: 10.1109/TNS.2004.829392.
- Schulte, R. W. et al. (2008). “A maximum likelihood proton path formalism for application in proton computed tomography”. In: *Medical Physics* 35.11, p. 4849. ISSN: 00942405. DOI: 10.1118/1.2986139.
- Schulte, R. W. et al. (2005). “Density resolution of proton computed tomography”. In: *Medical Physics* 32.4, pp. 1035–1046. ISSN: 00942405. DOI: 10.1118/1.1884906.



- Scott, W. T. (1963). “The Theory of Small-Angle Multiple Scattering of Fast Charged Particles”. In: *Reviews of Modern Physics* 35.2, pp. 231–313. ISSN: 0034-6861. DOI: 10.1103/RevModPhys.35.231.
- Taylor, J. T. et al. (2016). “An experimental demonstration of a new type of proton computed tomography using a novel silicon tracking detector”. In: *Medical Physics* 43.11, pp. 6129–6136. ISSN: 00942405. DOI: 10.1118/1.4965809.
- Williams, D. C. (2004). “The most likely path of an energetic charged particle through a uniform medium”. In: *Physics in Medicine and Biology* 49.13, pp. 2899–2911. ISSN: 0031-9155. DOI: 10.1088/0031-9155/49/13/010.
- Wong, K. et al. (2009). “The Effect of Tissue Inhomogeneities on the Accuracy of Proton Path Reconstruction for Proton Computed Tomography”. In: *AIP Conference Proceedings*. AIP, pp. 476–480. DOI: 10.1063/1.3120078.