A Deterministic Adjoint-Based Semi-Analytical Algorithm for Fast Response Change Computations in Proton Therapy

Burlacu, Tiberiu; Lathouwers, Danny; Perkó, Zoltán

DOI
10.1080/23324309.2023.2166077

Publication date
2023

Document Version
Final published version

Published in
Journal of Computational and Theoretical Transport

Journal of Computational and Theoretical Transport

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ltty21

A Deterministic Adjoint-Based Semi-Analytical Algorithm for Fast Response Change Computations in Proton Therapy

Tiberiu Burlacu, Danny Lathouwers & Zoltán Perkó

To cite this article: Tiberiu Burlacu, Danny Lathouwers & Zoltán Perkó (2023): A Deterministic Adjoint-Based Semi-Analytical Algorithm for Fast Response Change Computations in Proton Therapy, Journal of Computational and Theoretical Transport, DOI: 10.1080/23324309.2023.2166077

To link to this article: https://doi.org/10.1080/23324309.2023.2166077

Published online: 27 Jan 2023.

JOURNAL OF COMPUTATIONAL AND THEORETICAL TRANSPORT
https://doi.org/10.1080/23324309.2023.2166077

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

OPEN ACCESS Check for updates

A Deterministic Adjoint-Based Semi-Analytical Algorithm for Fast Response Change Computations in Proton Therapy

Tiberiu Burlacu[a,b], Danny Lathouwers[a,b†], and Zoltán Perkó[a,b†]

[a]Department of Radiation Science & Technology, Faculty of Applied Sciences, Delft University of Technology, Delft, Netherlands; [b]HollandPTC consortium – Erasmus Medical Center, Rotterdam, Holland Proton Therapy Centre, Delft, Leiden University Medical Center (LUMC), Leiden and Delft University of Technology, Delft, The Netherlands

ABSTRACT

In this paper we propose a solution to the need for a fast particle transport algorithm in Online Adaptive Proton Therapy capable of cheaply, but accurately computing the changes in patient dose metrics as a result of changes in the system parameters. We obtain the proton phase-space density through the product of the numerical solution to the one-dimensional Fokker-Planck equation and the analytical solution to the Fermi-Eyges equation. Moreover, a corresponding adjoint system was derived and solved for the adjoint flux. The proton phase-space density together with the adjoint flux and the metric (chosen as the energy deposited by the beam in a variable region of interest) allowed assessing the accuracy of our algorithm to different perturbation ranges in the system parameters and regions of interest. The algorithm achieved negligible errors ($1.1 \times 10^{-6}\%$ to $3.6 \times 10^{-3}\%$) for small Hounsfield unit (HU) perturbation ranges (–40 HU to 40 HU) and small to moderate errors (3% to 17%) – in line with the well-known limitation of adjoint approaches – for large perturbation ranges (–400 HU to 400 HU) in the case of most clinical interest where the region of interest surrounds the Bragg peak. Given these results coupled with the capability of further improving the timing performance it can be concluded that our algorithm presents a viable solution for the specific purpose of Online Adaptive Proton Therapy.

KEYWORDS

Charged particle transport; functional analysis; numerical methods

CONTACT Tiberiu Burlacu t.burlacu@tudelft.nl Department of Radiation Science & Technology, Faculty of Applied Sciences, Delft University of Technology, Delft, Netherlands.
†Both coauthors contributed equally

1. Introduction

1.1. Charged particle transport

The importance of studying charged particle transport is perhaps best illustrated by its applications in a wide-ranging set of fields such as radiation protection, radiotherapy, space radiation shielding, electron and ion beam microscopy, or surface analysis and lithography (Zheng-Ming and Brahme 1993). The goal of charged particle transport problems is to obtain the phase-space density of particles using modeled or empirically sourced reaction cross-sections. The phase-space density of particles provides a complete description of the particle fluence and all quantities of interest that can be derived from it. To obtain a general integro-differential equation that describes the phase-space density of particles in a scattering medium the collision-free Boltzmann equation (Duderstadt and Martin 1979) is altered to account for collisions via a scatter term. This equation is Boltzmann's general transport equation and for most realistic applications its solution is highly complex.

In practice several application-dependent approximations are applied to the Boltzmann equation in order to obtain an analytical or numerical solution, with all approaches having their individual tradeoffs. For example, while Monte Carlo (MC) methods have as advantages high precision and an ease of understanding (since" as long as the interaction processes between charged particles and atoms or the solid are known, the difficulties of the Monte-Carlo method are reduced to effectively realizing statistical sampling of the relevant interaction processes" (Zheng-Ming and Brahme 1993)), their main disadvantage is the slow computation times which deem them inapplicable in many scenarios (Zheng-Ming and Brahme 1993), especially when (near) real-time calculations are necessary (Botas et al. 2018). Diametrically opposite to MC methods from a computational expense standpoint are analytical methods such as the pencil beam approaches, which are obtained through fits and approximations. As expected, these methods tradeoff the high precision for the low computational expense. In between these two extremes lie several numerical or semi-numerical approaches, such as the moment method or the phase space time evolution method (Cordaro and Zucker 1972).

The focus of this paper is on a combination of numerical and analytical methods (the pencil beam and energy straggling methods) that are deemed promising for fulfilling the needs of our specific application, so called online adaptive proton therapy, which is currently the state-of-the-art form of radiotherapy.

1.2. Particle transport needs in online adaptive proton therapy

Within the field of radiation therapy, proton therapy (PT) has emerged as an alternative to conventional photon radiotherapy for cancer treatment

due to its promises of increased dose conformity and lowered doses achievable in healthy tissues (Paganetti 2016). These benefits are due to the presence of the Bragg peak (BP) in the depth-dose distribution, as charged particles deposit most of their energy within a small volume near the end of their range. The Bragg peak however also makes proton doses highly susceptible to uncertainties (Lomax 2008; Perkó et al. 2016). Some of the common sources of range uncertainties are related to computed tomography (CT) imaging, treatment delivery or changes in the anatomy of the patient (Paganetti 2012).

Currently, the state-of-the-art in dealing with uncertainties in clinical practice is to apply robust optimization (Rojo-Santiago et al. 2021; van der Voort et al. 2016). In robust optimization irradiation plans are optimized such that they ensure good performance of the plan under even the most extreme uncertainty scenarios (Unkelbach and Paganetti 2018). Due to the complexity of the potential scenarios however, certain scenarios – such as anatomical variations (e.g., weight loss over the course of often weeks long treatments) – are typically not accounted for (Paganetti et al. 2021). Most importantly, robust optimization essentially enlarges the high dose volume around the tumor, increasing the dose in the surrounding healthy tissues, which in turn increases the probability of detrimental side effects (van de Water et al. 2016).

The ideal solution would be to use Online Adaptive Proton Therapy (OAPT) instead. In OAPT, a daily CT scan of the patient is acquired and within 30 seconds (the time for a robotic arm to move the patient from the in-room CT scanner to the irradiation location) a new, fully re-optimized plan is created (Botas et al. 2018). Having up-to-date anatomical information allows accurately targeting the tumor (Paganetti et al. 2021) without needing robust optimization, leading to smaller irradiated volumes and fewer side effects. Unfortunately however, the computational expense of dose calculations and plan re-optimization (Men, Jia, and Jiang 2010), and the time needed for the presently mostly manual plan quality assurance (QA) (Barrett et al. 2009) is far larger than 30 seconds, making such workflows currently clinically infeasible.

Fast proton transport methods that are accurate in highly heterogeneous patient geometries are key to overcome these computational and QA related bottlenecks, and represent one of the missing enabling technology for online adaptive workflows and further improving cancer treatments. First, they are necessary for re-optimization, as plan optimization requires the dose distribution from each of the typically hundreds or even thousands of individual proton beams as input (Schwarz 2011). Second, they are crucial for replacing the current manual, measurement based plan QA with fast computational alternatives. Traditional plan QA measurements

assess the differences between planned and delivered doses in order to ensure they are within the clinically acceptable $\pm 3\%$ (Gottschalk 2004) range and that the irradiation delivery system functions as intended (Frank and Zhu 2020). Since manual measurements are clearly infeasible in OAPT, independent dose calculation methods (Li et al. 2013) have been proposed as a viable alternative, showing similar precision when using accurate MC transport methods (Meier et al. 2015). As further advantage, such automated QA procedures yield clinically more relevant metrics than measurements and could potentially even increase clinical throughput and treatment accessibility (Meijers et al. 2020). While the benefits of automated QA procedures based on independent dose calculation and machine log-files (measured the outgoing radiation from the treatment machine) are clear, MC calculations (Matter et al. 2018), even when multi-threaded (Meijers et al. 2020) are not fast enough to perform (near) real-time QA necessary in the OAPT workflow.

1.3. A semi-analytical adjoint-based deterministic algorithm for OAPT

To overcome these issues we propose a semi-analytical adjoint-based deterministic algorithm that could serve as (near) real-time plan QA, using machine log-files and the patient geometry. The semi-analytical component aims to provide a balance between the accuracy of MC algorithms and the speed of analytical dose calculation algorithms. The adjoint component aims to provide real-time quality assurance through efficient computations of the effect of perturbations in the system parameters (beam spatial and energy spread, its particle number or the patient geometry) on the desired clinical metrics (dose, or more complex responses).

The semi-analytical component has as a starting point, similarly to the MC algorithms, the Linear Boltzmann Equation (LBE). Through the continuous slowing down, energy straggling and Fokker-Planck approximations the LBE can be reduced to two partial differential equations (PDEs). One of the PDEs is the one-dimensional Fokker-Planck (FP) equation while the other one is the Fermi-Eyges (FE) equation. The advantage of this approach is threefold. First, the approach derives a system which is described by two PDEs. The presence of the PDEs (as opposed to for example a machine learning (ML) based dose engine (Pastor-Serrano and Perkó 2022)) allows the application of adjoint methods. Using functional analysis an adjoint system can be derived which can be used to avoid the expensive process of re-computing the solution to the two PDEs for each new set of system parameters. Second, the physical approximation will not suffer from the typical drawbacks of ML models such as out-of-distribution samples, i.e., will remain accurate despite the input not being already seen

by the algorithm. Third, while the one-dimensional FP equation requires a numerical solution the FE equation has a known analytical solution (Eyges 1948; Brahme 1975). This coupling will ensure the computational effectiveness as the lateral part of the proton flux is computed through a straightforward function evaluation.

1.4. Paper outline

Section 2 covers the theoretical background of reducing the LBE to two simplified PDEs. Section 3 describes the one-dimensional Fokker-Planck equation and its numerical solution while Section 4 covers the Fermi-Eyges solution. In Section 5 the application of the functional analysis framework for the derivation of the adjoint system with its associated adjoint solution is detailed, the solution methodology of the adjoint system is explained and the response change computations due to perturbations in the system parameters are given. Section 6 covers benchmarks of our own algorithm versus TOPAS and Bortfeld's algorithm and provides comparisons between the forward and adjoint computation of the response changes due to system parameter perturbations. Lastly, Section 7 contains some conclusions and future intended research directions.

2. The system model

The physical system under consideration is given by a proton beam irradiating the patient. This system can be characterized through the (steady-state) LBE, the validity of which for PT has been discussed by Börgers (Börgers 1999). The LBE describes the proton balance in an arbitrary volume. Its derivation is obtained by equating all the gain and loss mechanisms for protons at position $\mathbf{r} \in \mathbb{R}^3$ with a certain energy E in $\mathrm{d}E$ and direction given by the unit vector $\hat{\mathbf{\Omega}} = \mathbf{v}/|\mathbf{v}|$ (with $\mathbf{v}$ the velocity vector of the protons) in $\mathrm{d}\hat{\mathbf{\Omega}}$ in an arbitrary volume V with a boundary denoted by ∂V as outlined by Duderstadt & Hamilton (Duderstadt and Hamilton 1991). The equation is an integro-differential equation for the proton flux ($\varphi = vn$) with v the proton speed and $n(\mathbf{r}, E, \hat{\mathbf{\Omega}})$ the angular proton density,

$$\begin{aligned} &\hat{\mathbf{\Omega}}\cdot\nabla\varphi + \Sigma_t(\mathbf{r}, E)\varphi(\mathbf{r}, E, \hat{\mathbf{\Omega}}) \\ &\quad = \int_{4\pi} \mathrm{d}\hat{\mathbf{\Omega}}' \int_0^\infty \mathrm{d}E' \Sigma_s(E' \to E, \hat{\mathbf{\Omega}}' \to \hat{\mathbf{\Omega}})\varphi(\mathbf{r}, E', \hat{\mathbf{\Omega}}') + s(\mathbf{r}, E, \hat{\mathbf{\Omega}}) \end{aligned} \tag{1}$$

$$\mathrm{BC}: \ \varphi(\mathbf{r}_s, E, \hat{\mathbf{\Omega}}) = 0 \ \text{ if } \ \hat{\mathbf{\Omega}}\cdot\hat{\mathbf{e}}_s < 0 \ \text{ with } \ \mathbf{r}_s \in \partial V, \tag{2}$$

where BC is a boundary condition of the non-reentrant type, $\mathbf{r}_s$ denotes a vector on the boundary surface ∂V of the volume V, $\hat{\mathbf{e}}_s$ is the unit outward

pointing normal vector to the boundary ∂V at $\mathbf{r}_s$, Σ_t is the total macroscopic cross section, Σ_s is the macroscopic double differential scattering cross section and s is the source of protons.

Currently, the LBE in its form is computationally expensive to solve. A first step is to divide the total Σ_t and scatter Σ_s cross sections according to the main interactions that a proton undergoes as it propagates through the medium, namely $\Sigma_t = \Sigma_a + \Sigma_e + \Sigma_{in}$ where Σ_a is the catastrophic (absorption) scatter cross section, Σ_e is the elastic scatter cross section between the incident protons and the nuclei of tissue, Σ_{in} is the inelastic scatter cross section between the incident protons and atomic electrons. By doing so, Equation (1) can be written as

$$\begin{aligned}\hat{\mathbf{\Omega}}\cdot\nabla\varphi = &\int_{4\pi} \mathrm{d}\hat{\mathbf{\Omega}}' \int_{E}^{\infty} \mathrm{d}E' \Sigma_a(E' \to E, \hat{\mathbf{\Omega}}' \to \hat{\mathbf{\Omega}})\varphi(\mathbf{r}, E', \hat{\mathbf{\Omega}}') - \Sigma_a(\boldsymbol{r}, E)\varphi(\mathbf{r}, E, \hat{\mathbf{\Omega}}) \\ &+ \int_{4\pi} \mathrm{d}\hat{\mathbf{\Omega}}' \Sigma_e(\mathbf{r}, E, \hat{\mathbf{\Omega}}' \to \hat{\mathbf{\Omega}})\varphi(\mathbf{r}, E, \hat{\mathbf{\Omega}}') - \Sigma_e(\boldsymbol{r}, E)\varphi(\mathbf{r}, E, \hat{\mathbf{\Omega}}) \\ &+ \int_{0}^{\infty} \mathrm{d}Q \Sigma_{in}(\boldsymbol{r}, E + Q \to E, \hat{\mathbf{\Omega}})\varphi(\mathbf{r}, E + Q, \hat{\mathbf{\Omega}}) - \Sigma_{in}(\boldsymbol{r}, E)\varphi(\mathbf{r}, E, \hat{\mathbf{\Omega}}),\end{aligned} \tag{3}$$

with Q defined as the amount of energy transferred during an interaction. In this splitting it is assumed that the energy transfer in Coulomb elastic scatter interactions is negligible and that the angular deflection in Coulomb inelastic scatter interactions is negligible (Zheng-Ming and Brahme 1993). The next step is to apply approximations to each of the collision integrals in Equation (3).

The inelastic scatter integral is approximated using the Continuous Slowing Down Approximation (CSDA) and the Energy-loss Straggling (ELS) approximation (Zheng-Ming and Brahme 1993). Given the difference between the proton and electron mass, the energy loss of a proton beam in each such individual collision is small. Thus, the stopping process can effectively be approximated by a continuous energy loss process with a mean (called the stopping power) and a deviation around the mean (called the straggling coefficient).

Therafter, we orient the beam along the z-direction and apply the small angle (i.e., $\Omega_z = 1$ in Cartesian angular coodinates) Fokker-Planck approximation to the elastic scatter angular integral. In doing so, $\hat{\mathbf{\Omega}}$ is redefined as $\hat{\mathbf{\Omega}} = (\Omega_x, \Omega_y) \in \mathbb{R}^2$. Similarly to the CSDA process, due to the small angular deflection that the proton beam suffers through its Coulomb interactions with the atom the scattering process is approximated as a continuous

diffusion term in the lateral angular plane. Moreover, in the elastic scattering cross section $\Sigma_e(E, \hat{\mathbf{\Omega}} \cdot \hat{\mathbf{\Omega}}')$ the energy is replaced by the depth-dependent mean energy $E_a(z)$ (Gebäck and Asadzadeh 2012; Zheng-Ming and Brahme 1993).

The catastrophic inscatter integral is neglected completely with only the absorption catastrophic scatter cross section term remaining. Applying these approximations to the LBE reduces the integro-differential equation to the following PDE

$$\frac{\partial \varphi}{\partial z} + \Omega_x \frac{\partial \varphi}{\partial x} + \Omega_y \frac{\partial \varphi}{\partial y} - \frac{\partial S(\boldsymbol{r}, E)\varphi}{\partial E} - \frac{1}{2}\frac{\partial^2 T(\boldsymbol{r}, E)\varphi}{\partial E^2} + \Sigma_a(\boldsymbol{r}, E)\varphi - \Sigma_{tr}(z, E_a(z))\left(\frac{\partial^2 \varphi}{\partial \Omega_x^2} + \frac{\partial \varphi^2}{\partial \Omega_y^2}\right) = 0, \tag{4}$$

where $S(\boldsymbol{r}, E)$ is the stopping power (the mean energy loss per unit path of the proton), $T(\boldsymbol{r}, E)$ is the straggling coefficient (the deviation of the energy loss around its mean value), Σ_a is the absorption cross section (the removal of protons from the beam due to nuclear interactions) and Σ_{tr} is the transport cross section (the rate at which protons diffuse in the lateral angular plane). The resulting PDE is linear in the dependent variable φ which in turn depends on the six independent system variables $\mathbf{r}, \hat{\mathbf{\Omega}}, E$.

We generalize the work of Gebäck and Asadzadeh (Gebäck and Asadzadeh 2012) by considering a laterally homogeneous, in-depth heterogeneous geometry and write the flux as

$$\varphi = \varphi_{FE}(\boldsymbol{r}, \hat{\mathbf{\Omega}}) \cdot \varphi_{FP}(z, E). \tag{5}$$

To simplify notation, the dependence of the fluxes $\varphi_{FE}(\boldsymbol{r}, \hat{\mathbf{\Omega}})$ and $\varphi_{FP}(z, E)$ on their respective independent variables $(\boldsymbol{r}, \hat{\mathbf{\Omega}})$ and (z, E) will be suppressed in the rest of this paper. Substitution in Equation (4) results in

$$\Upsilon(\varphi_{FE}) \cdot \varphi_{FP} + \varphi_{FE} \cdot \mathrm{1DFP}(\varphi_{FP}) = 0, \tag{6}$$

where $\Upsilon(\varphi_{FE})$ is the Fermi-Eyges equation and $\mathrm{1DFP}(\varphi_{FP})$ is the one-dimensional Fokker-Planck equation. In order to avoid the trivial solution both of these equations are set to zero (for a proof of separability see Appendix), yielding

$$\Upsilon(\varphi_{FE}) = \frac{\partial \varphi_{FE}}{\partial z} + \Omega_x \frac{\partial \varphi_{FE}}{\partial x} + \Omega_y \frac{\partial \varphi_{FE}}{\partial y} - \Sigma_{tr}(z, E_a(z))\left(\frac{\partial^2 \varphi_{FE}}{\partial \Omega_x^2} + \frac{\partial^2 \varphi_{FE}}{\partial \Omega_y^2}\right) = 0 \tag{7}$$

and

$$1\mathrm{DFP}(\varphi_{FP}) = \frac{\partial \varphi_{FP}}{\partial z} - \frac{\partial S(z,E)\varphi_{FP}}{\partial E} - \frac{1}{2}\frac{\partial^2 T(z,E)\varphi_{FP}}{\partial E^2} + \Sigma_a(z,E)\varphi_{FP} = 0. \quad (8)$$

Searching for the solution in the split form defined by Equation (5) is a usual mathematical trick for the separation of variables, ensuring that the solutions of Equations (7) and (8) yield the exact solution of Equation (4). However, such a split of the proton flux also has strong physics foundations, even in the more general setting. Since catastrophic inscatter interactions are rare, the process mostly responsible for energy change is the Coulomb inelastic scatter. The energy loss of protons is therefore primarily determined by the stopping power and range straggling properties of the materials they traverse through. Due to the laterally homogeneous (or at least not too inhomogeneous) geometry on the scale of a highly focused beam (typically only 2-3 mm in clinical proton beams) and the strong forward scattering, the materials along the traversed through path till any given depth z are very similar for protons traveling under slightly different angles $\hat{\mathbf{\Omega}}$, resulting in strong coupling between the energy spectrum of the beam and the depth. Moreover, since the elastic Coulomb scatter mostly responsible for the angular spread of the beam causes negligible energy change, the energy spectrum of protons having slight deviations $\hat{\mathbf{\Omega}}$ from the main beam direction $\mathbf{\Omega}_z = 1$ and that of protons with the original uncollided direction $\hat{\mathbf{\Omega}} = (0,0,1)^T$ is similar. These observations provide strong reasoning for searching for the solution in the form of Equation (5) – with a depth dependent energy spectrum $\varphi_{FP}(z,E)$ that is independent from the spatially dependent angle distribution $\varphi_{FE}(\boldsymbol{r},\hat{\mathbf{\Omega}})$ – even in more general cases.

Using the solution of Equations (7) and (8), the response of the system can be defined which in this case was chosen as the energy deposited in a certain region of interest (ROI). The method is applicable to other, more general, responses (defined as functionals or operators) as long as the chosen response satisfies a weak Lipschitz condition in the system state vector and parameters (Cacuci 2003). In the case of this work, the response R is given by

$$\begin{aligned} R(\varphi) &= -\int_{ROI} \mathrm{d}V \int_{4\pi} \mathrm{d}\hat{\mathbf{\Omega}} \int_{E_{\min}}^{E_{\max}} \mathrm{d}E E \left[\frac{\partial S(z,E)\varphi}{\partial E} + \frac{1}{2}\frac{\partial^2 T(z,E)\varphi}{\partial E^2} - \Sigma_a(z,E)\varphi \right] \\ &= -\int_{ROI} \mathrm{d}V \int_{4\pi} \mathrm{d}\hat{\mathbf{\Omega}} \varphi_{FE} \int_{E_{\min}}^{E_{\max}} \mathrm{d}E E \left[\frac{\partial S(z,E)\varphi_{FP}}{\partial E} + \frac{1}{2}\frac{\partial^2 T(z,E)\varphi_{FP}}{\partial E^2} - \Sigma_a(z,E)\varphi_{FP} \right], \end{aligned} \quad (9)$$

where in the last equality Equation (5) was employed.

3. Approximating the one-dimensional Fokker-Planck equation

The one-dimensional Fokker-Planck equation is a convection-diffusion equation in energy whose character is well suited for Discontinuous Galerkin (DG) methods. Consequently, its semi-discrete form was obtained using the Symmetric Interior Penalty Galerkin (SIPG). The main advantages of the SIPG method over other finite element methods (FEM) are the relative ease with which the approximating polynomial can be changed on different mesh elements, the fact that the method allows unstructured or adaptive meshes, and that the method satisfies a local energy balance (as opposed to the global energy balance satisfied by continuous Galerkin methods) (Rivière 2008). The semi-discrete form was solved using the Crank-Nicholson (CN) method which is a second order accurate implicit finite difference method. The advantage of the CN method is that in a geometry that is piece-wise constant it relies on only one of the previous points (as opposed to schemes such as the Backward Differentiation Formula 2 that require two previous points for the same order of accuracy (Suli and Mayers 2003)).

The one-dimensional Fokker-Planck equation can be written in a more standard convection-diffusion form

$$\frac{\partial \varphi_{FP}}{\partial z} - \frac{\partial S^*(z,E)\varphi_{FP}}{\partial E} - \frac{\partial}{\partial E}\left(T^*(z,E)\frac{\partial \varphi_{FP}}{\partial E}\right) + \Sigma_a(z,E)\varphi_{FP} = 0, \tag{10}$$

where the modified stopping power $S^*(z,E) = S(z,E) + \frac{1}{2}\frac{\partial T(z,E)}{\partial E}$ and the modified straggling coefficient $T^*(z,E) = T(z,E)/2$ are introduced. To simplify notation, from here on the stars will be dropped and the explicit (z, E) dependence of the stopping power and the straggling coefficients on the depths z and energy E will only be shown if necessary. Moreover, it is Equation (10) that will from now on be referred to as the one-dimensional Fokker-Planck equation.

3.1. Domain definition and discretization

The computational domain of the equation is given as $\mathscr{D} = (0, z_{\max}) \times (E_{\min}, E_{\max}), \mathscr{D} \subset \mathbb{R}^2$. The solution of the one-dimensional Fokker-Planck equation is the Fokker-Planck flux $\varphi_{FP}(z,E) \in \mathscr{H}$ where $\mathscr{H} = \mathcal{L}_2(\mathbb{R}^2)$ is a real Hilbert space of square integrable functions with an associated inner product defined as

$$\langle f, g \rangle = \int_0^\infty \mathrm{d}z \int_0^\infty \mathrm{d}E f g.$$

To ensure a unique solution to Equation (10) boundary conditions must be imposed, namely

$$\text{BCE}: \left.\varphi_{FP}(z,E)\right|_{E=E_{\max}} = 0, \left.\frac{\partial \varphi_{FP}(z,E)}{\partial E}\right|_{E=E_{\max}} = 0. \tag{11}$$

$$\text{BCS}: \varphi_{FP}(0,E) = Ae^{-\left(\frac{E-E_0}{\sigma_E}\right)^2}. \tag{12}$$

The boundary conditions in energy (BCE) are homogeneous Dirichlet and Neumann conditions while the boundary condition in space (BCS) is given by a Gaussian function in energy. Gerbershagen (Gerbershagen et al. 2017) showed that this is a realistic energy spectrum for a proton beam that has suffered energy degradation. The amplitude of the Gaussian is given by A, the mean by E_0 and the energy spread is given by σ_E. In line with usual practice, a rigorous proof of the existence and uniqueness of the solution to Equation (10) and its associated boundary conditions is not given and these properties are assumed to be true. The energy component of the domain $\mathscr{D}$ is discretized into a number NG of groups with each group having the same width. The minimum and maximum energy of the domain are chosen to encompass the standard clinical proton energies range of 1 MeV to 200 MeV. In a given group g the high energy boundary is denoted by $E_{g-1/2}$, the low energy one by $E_{g+1/2}$ and the center value by E_g. Thus, $E_{\max} = E_{1/2}$ and $E_{\min} = E_{NG+1/2}$. An illustration of this discretization can be seen in Figure 1.

The spatial part of the domain $\mathscr{D}$ is discretized into a number of steps N_s with the interval length Δz allowed to vary on a per step basis and the start and end points of the spatial domain are given by $z_0 = 0$ and $z_{Ns} = z_{max}$.

Equation (10) can also be written in a short-hand form as

$$L(\boldsymbol{\alpha})\varphi_{FP} = 0$$

where the vector of system parameters $\boldsymbol{\alpha}$ and the differential operator $L(\boldsymbol{\alpha})$ acting on the flux are introduced as

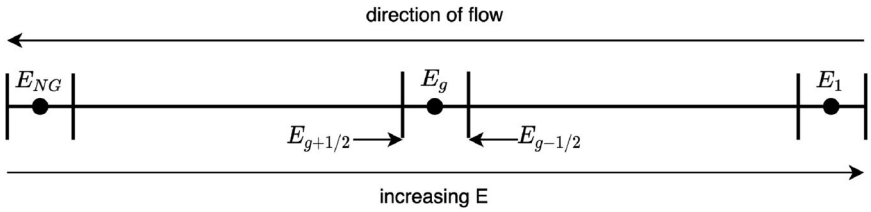

Figure 1. Energy domain discretization.

$$L(\boldsymbol{\alpha})(\cdot) = \frac{\partial}{\partial z}(\cdot) - \frac{\partial S^*(\cdot)}{\partial E} - \frac{\partial}{\partial E}\left[T^* \frac{\partial(\cdot)}{\partial E}\right] + \Sigma_a(\cdot),$$
$$\text{and } \boldsymbol{\alpha} = (S^*(z,E), T^*(z,E), \Sigma_a(z,E)).$$

The stopping power, energy straggling and absorption cross sections are all approximated as continuous, piece-wise linear functions in the NG energy groups. Since these properties are unique for each nuclide, we typically need as many material datasets as many voxels (defined as a cubic element in the CT scan) the beam traverses, as the CT HU units are mapped to different material compositions. Denoting the space of univariate polynomials with real coefficients and degree at most k as $\mathbb{P}^k$, and the mesh in the energy variable as $\mathcal{T} = \{EI_i\}_{i\in 1,\ldots,NG}$, with $EI_i = [E_{i+1}, E_i]$ being the continuous energy interval in the i^{th} group, all our material data can represented by the vector space of continuous, piecewise linear functions defined as $P^1_{\mathcal{T}} = \left\{\nu_{\mathcal{T}} \in \mathcal{C}^0(R^+) | \forall i \in \{1,\ldots,NG\},\ \nu_{\mathcal{T}}|_{I_i} \in \mathbb{P}^1\right\}$. Thus, the space in which $\boldsymbol{\alpha}$ resides is the tensor product constructed from the individual spaces $P^1_{\mathcal{T}}$ to which the stopping powers, straggling coefficients and absorption cross sections of all material domains belong.

3.2. Semi-discrete variational formulation

The first step to obtain an approximation to the solution of Equation (10) and its associated boundary conditions 11, 12 is to obtain the semi-discrete variational formulation. To do so, several quantities must be defined. First, the jump and the average of the flux at the edges of an energy group are defined as

$$[\varphi] = \varphi(E_j^-) - \varphi(E_j^+),$$
$$\text{and } \{\varphi\} = \frac{1}{2}\left(\varphi(E_j^-) + \varphi(E_j^+)\right),$$

where $j = \frac{1}{2}, \ldots, NG + \frac{1}{2}$ and with $E_j^- = \lim_{\epsilon \downarrow 0}(E_j - \epsilon)$ and $E_j^+ = \lim_{\epsilon \downarrow 0}(E_j + \epsilon)$. Special cases are defined at the boundary of the energy domain where

$$\left[\nu(E_{NG+1/2})\right] = -\nu(E^+_{NG+1/2}),\ \left\{\nu(E_{NG+1/2})\right\} = \nu(E^+_{NG+1/2}),\ \text{and}$$
$$\left[\nu(E_{1/2})\right] = \nu(E^-_{1/2}),\ \left\{\nu(E_{1/2})\right\} = \nu(E^-_{1/2}).$$

Second, the penalty term is defined as

$$J_0(\nu, w) = \sum_{j=1/2}^{NG+1/2} \frac{\sigma^0}{h_{j-1,j}} \left[\nu(E_j)\right]\left[w(E_j)\right]$$

where $h_{j-1,j} = \max(\Delta E_{j-1}, \Delta E_j)$ and σ^0 is a real and nonnegative number bounded from below. The role of this term is to penalize the jumps in the solution.

By multiplying Equation (10) with a test function v, integrating over one group, thereafter summing over all energy groups and making use of the definitions of the jump and the average, the semi- discrete variational formulation is found to be

$$\int_{E_{\min}}^{E_{\max}} \mathrm{d}E \frac{\partial \varphi_{FP}}{\partial z} v + a_{SIPG}(\varphi_{FP}, v) - \int_{E_{\min}}^{E_{\max}} \mathrm{d}E \frac{\partial S^* \varphi_{FP}}{\partial E} v + \int_{E_{\min}}^{E_{\max}} \mathrm{d}E \Sigma_a \varphi_{FP} v = 0, \quad (13)$$

where the SIPG bilinear a_{SIPG} is (Rivière 2008)

$$a_{SIPG}(\varphi_{FP}, v) = \int_{E_{\min}}^{E_{\max}} T \frac{\partial \varphi_{FP}}{\partial E} \frac{\mathrm{d}v}{\mathrm{d}E} \mathrm{d}E + \sum_{\Gamma_i} -\left\{ T \frac{\partial \varphi_{FP}}{\partial E} \right\} \cdot [v] - [\varphi_{FP}] \cdot \left\{ T \frac{\mathrm{d}v}{\mathrm{d}E} \right\} + \frac{\sigma^0}{h_{j-1,j}} [\varphi_{FP}][v], \quad (14)$$

where Γ_i denotes the interior points of the energy domain. Following Hillewaert's work (Hillewaert 2013), the penalty parameter was set as a function of the maximum polynomial degree $\max(\deg(p_g^i))$ of the basis functions, namely

$$\sigma^0 = \frac{(\max(\deg(p_g^i)) + 1)^2}{2}. \quad (15)$$

Both a coercivity analysis and the proof of equivalence between the semi-discrete variational formulation from Equation (13) and the model problem 10 with its associated boundary conditions 11 and 12 are beyond the scope of this paper and can be found in the work of Hillewaert and Riviere respectively (Hillewaert 2013; Rivière 2008).

3.3. Basis functions

The first three group-centered Legendre polynomials[1]

$$p_g^i(E) \equiv P_i\left(\frac{2}{\Delta E_g} (E - E_g) \right), i = 0, 1, 2 \quad (16)$$

[1]Initially, the algorithm used first order basis functions. However, the resulting fluxes for coarse energy and spatial grids resulted in unphysical negative values. The addition of the third quadratic basis function reduced the behaviour to negligible values.

were used as the basis functions for the expansion of the flux in the computational domain as

$$\varphi_{FP}(z,E) = \sum_{g=1}^{NG} \sum_{i=0}^{2} \varphi_g^i(z) p_g^i(E). \tag{17}$$

Introducing the expansion from Equation (17) into the semi-discrete variational formulation from Equation (13) and sequentially replacing the function v with the chosen basis functions $p_g^i(E)$ yields a system of linear equations. This system can be written as

$$M\frac{\mathrm{d}\mathbf{\Phi}}{\mathrm{d}z} + G\mathbf{\Phi} = 0, \tag{18}$$

where $\mathbf{\Phi}$ is a vector with dimension $(1 + \max(\deg(p_g^i))) \times NG$ and its elements are given by the unknown coefficients $\varphi_g^i(z)$ from Equation (17), the mass matrix M is a diagonal matrix that in a given group g has elements $\int \mathrm{d}E p_g^i(E) p_g^i(E)$ with $i = 0,\ 1,\ 2$ along the diagonal and G is the system matrix which receives contributions from the stopping power, straggling coefficient and absorption cross section discretization.

This resulting system is discretized in space using the Crank-Nicholson method. Depending on the chosen number of groups the size of the resulting system is on the order of 10^3. This relatively small size of the system of equations implies that direct solution methods are comparable in computational time to iterative ones. To this end, the banded system solver DGBSV from the LAPACK library (Anderson et al. 1999) was used.

4. The Fermi-Eyges equation

This section describes the analytical solution to the Fermi-Eyges equation and the steps taken to implement it. This solution is based on refinements brought to Fermi's original theory on the distribution of charged particles undergoing multiple elastic scattering in their passing through matter. Authors such as Eyges, Brahme and Asadzadeh (Eyges 1948; Brahme 1975; Gebäck and Asadzadeh 2012) have brought the theory into its form presented here. A full derivation from basic principles is beyond the scope of this document and can be found in the previously mentioned publications.

The Fermi-Eyges equation

$$\Upsilon[\varphi_{FE}] = \frac{\partial \varphi_{FE}}{\partial z} + \Omega_x \frac{\partial \varphi_{FE}}{\partial x} + \Omega_y \frac{\partial \varphi_{FE}}{\partial y} - \Sigma_{tr}(z)\left(\frac{\partial^2 \varphi_{FE}}{\partial \Omega_x^2} + \frac{\partial^2 \varphi_{FE}}{\partial \Omega_y^2}\right) = 0 \tag{19}$$

can be solved by separating the x and y directions, namely $\varphi_{FE}(\boldsymbol{r}, \Omega_x, \Omega_y) = H(z,x,\Omega_x) \cdot H(z,y,\Omega_y)$. This results in two separate PDEs for each

direction

$$\frac{\partial H(z,\xi,\omega)}{\partial z}+\omega\frac{\partial H(z,\xi,\omega)}{\partial \xi}-\Sigma_{tr}(z)\frac{\partial^2 H(z,\xi,\omega)}{\partial \omega^2}=0, \tag{20}$$

where ξ stands for one of x, y and ω stands for one of Ω_x, Ω_y. The resulting PDEs have the same boundary condition imposed, namely

$$H(0,\xi,\omega)=C\exp\left(-\left(a_1\xi^2+a_2\xi\omega+a_3\omega^2\right)\right), \tag{21}$$

with $a_i\in\mathbb{R}, \forall i=1,2,3$ and $C>0$. The solution of Equation (20) is found by artificially extending the usual domain of Ω_x and Ω_y from $\{(\Omega_x,\Omega_y): |\hat{\mathbf{\Omega}}|=1\}$ to $(\Omega_x,\Omega_y)\in[-\infty,\infty]^2$, applying two-dimensional Fourier transforms in ξ and ω and accounting for the Gaussian initial condition as detailed by Eyges and Brahme (Brahme 1975; Eyges 1948). In doing so the solution to the Fermi-Eyges Equation (19) is found to be

$$\varphi_{FE}(z,\boldsymbol{\rho},\hat{\mathbf{\Omega}})=\frac{A^2}{4\pi^2}\frac{\exp\left(-\frac{|\boldsymbol{\rho}|^2}{2\overline{\xi^2}(z)}\right)}{\overline{\xi^2}(z)}\frac{\exp\left(-\frac{1}{2B(z)}\left|\hat{\mathbf{\Omega}}-\frac{\overline{\theta\xi}(z)}{\overline{\xi^2}(z)}\boldsymbol{\rho}\right|^2\right)}{B(z)}, \tag{22}$$

where $\boldsymbol{\rho}=(x,y)$, $\hat{\mathbf{\Omega}}=(\Omega_x,\Omega_y)$ and the remaining quantities are defined as

$$B(z)=\overline{\theta^2}(z)-\frac{\left(\overline{\theta\xi}(z)\right)^2}{\overline{\xi^2}(z)},\quad A=\frac{2\pi C}{D}, D=4a_1a_3-a_2^2.$$

Jette (Jette 1988 Mar-Apr) showed that if there is any scattering at all, then $B\geq 0$ must hold. This was used as a check that the obtained coefficient values were not spurious. The coefficients $\overline{\theta^2}, \overline{\theta\xi}, \overline{\xi^2}$ present in the Fermi-Eyges solution are the moments of the Σ_{tr} transport cross section and are found from the following equations

$$\overline{\theta^2}(z)=\overline{\theta^2}(0)+\int_0^z\Sigma_{tr}(z')\mathrm{d}z',\ \text{with}\ \overline{\theta^2}(0)=\frac{2a_3}{D}, \tag{23a}$$

$$\overline{\theta\xi}(z)=\overline{\theta\xi}(0)+\overline{\theta^2}(0)z+\int_0^z(z-z')\Sigma_{tr}(z')\mathrm{d}z',\ \text{with}\ \overline{\theta\xi}(0)=\frac{a_2}{D}, \tag{23b}$$

$$\overline{\xi^2}(z)=\overline{\xi^2}(0)+2\overline{\theta\xi}(0)z+\overline{\theta^2}(0)z^2+\int_0^z(z-z')^2\Sigma_{tr}(z')\mathrm{d}z',\ \text{with}\ \overline{\xi^2}(0)=\frac{2a_1}{D}, \tag{23c}$$

where

$$\Sigma_{tr}(z) = \int_{-1}^{1} \mathrm{d}\mu \Sigma_s(E_a(z), \mu, z)(1-\mu), \text{ with } \mu = \cos(\hat{\mathbf{\Omega}} \cdot \hat{\mathbf{\Omega}}'), \tag{24}$$

and Σ_s is the macroscopic elastic scatter cross section. Gottschalk (Gottschalk 2012) showed that the FE coefficients $\overline{\theta^2}(z), \overline{\xi^2}(z), \overline{\theta\xi}(z)$ can be intepreted as the variances of the angular direction, the lateral position and the covariance of the lateral position and the angular direction respectively.

Next to its analytical nature an important feature of the Fermi-Eyges solution from Equation (22) is that it is a Gaussian function in both the spatial and angular directions with coefficients that are determined by the average depth-dependent beam energy and the elastic scatter cross section from Equation (24) corresponding to that energy. A disadvantage of this solution is that only the average depth dependent energy instead of the full beam energy spectrum is used to calculate the coefficients. As lower energy protons tend to scatter more it is expected that only using the average beam energy will result in an underestimation of the amount of scatter that the proton beam suffers.

4.1. Solution method

The coefficients of the boundary condition are chosen in such a way that Equation (21) represents the two-dimensional normal distribution. By setting the average values in ξ and ω to zero the coefficients $a_i, i \in [1,2,3]$ from Equation (21) are easily identified to be equal to

$$a_1 = \frac{1}{2(1-\varrho^2)\sigma_\xi^2}, \quad a_2 = -\frac{\varrho}{(1-\varrho^2)\sigma_\xi \sigma_\omega}, \quad a_3 = \frac{1}{2(1-\varrho^2)\sigma_\omega^2}, \tag{25}$$

where ϱ is the correlation coefficient between the spatial dimension ξ and the angular dimension ω, σ_ξ standard deviation in ξ and σ_ω standard deviation in ω. The a_i coefficients are thereafter used to initialize the values of the FE coefficients.

To compute the FE coefficients at a given depth the average beam energy at that depth must be known. This quantity was defined as

$$E_a(z) = \frac{\int_0^\infty \mathrm{d}E \varphi_{FP}(z,E) E}{\int_0^\infty \mathrm{d}E \varphi_{FP}(z,E)}. \tag{26}$$

The average energy is thereafter introduced into the elastic scatter cross section via the classical relationship between speed and energy $v_p = \sqrt{2E_a(z)/m_p}$. The elastic scatter cross section is in turn used to compute

the transport cross section from Equation (24). To compute the angle integral the QAGE routine from the QUADPACK library was used (Piessens et al. 1983).

Once $\Sigma_{tr}(z)$ is known for all the points of the z dimension, the coefficients given in Equation (23) can be calculated. As z increases in the integrals from Equation (23) so do the integrands and the computational expense of these integrals. We chose to approximate Σ_{tr} in a given step as the average of its values at the start and endpoint of the step thereby employing the trapezoidal integration rule. In doing so, the integrals could be re-written to depend only on the previous value. Given the fact that the segments over which Σ_{tr} is integrated are small (≤ 0.01 cm) the trapezoidal integration scheme is adequately accurate for our purposes.

4.2. The planar integral approximation

In the response computation the angle integrated FE flux is needed. In the derivation of the Fermi-Eyges solution the domain of Ω_x and Ω_y is extended from its normal range to the $(-\infty, \infty)$ range in order to apply the Fourier transforms. Since the Fermi-Eyges solution is obtained through this extension, the same extension should be consistently applied throughout the calculations that involve this solution. Thus, the angular integral can be approximated to

$$\begin{aligned}
&\int_{4\pi} \exp\left(-\frac{1}{2B(z)}\left|\hat{\boldsymbol{\Omega}}-\frac{\overline{\theta\xi}(z)}{\overline{\xi^2}(z)}\boldsymbol{\rho}\right|^2\right)\sin\theta \mathrm{d}\theta\mathrm{d}\phi \approx \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2B(z)}\left|\hat{\boldsymbol{\Omega}}-\frac{\overline{\theta\xi}(z)}{\overline{\xi^2}(z)}\boldsymbol{\rho}\right|^2\right)\mathrm{d}\Omega_x\mathrm{d}\Omega_y \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2B(z)}\left[(\Omega_x - c_x)^2 + (\Omega_y - c_y)^2\right]\right)\mathrm{d}\Omega_x\mathrm{d}\Omega_y \\
&= \sqrt{2\pi B(z)}\cdot\sqrt{2\pi B(z)} = 2\pi B(z).
\end{aligned} \tag{27}$$

where $c_i = [\overline{\theta\xi}(z)/\overline{\xi^2}(z)]i$, with i being x or y. Thus, the angularly integrated FE flux is

$$\Psi_{FE}(x,y,z) = \int_{4\pi} \varphi_{FE}(x,y,z,\Omega_x,\Omega_y)\mathrm{d}\hat{\boldsymbol{\Omega}} = \frac{A^2}{2\pi}\frac{1}{\overline{\xi^2}(z)}\exp\left(-\frac{|\boldsymbol{\rho}|^2}{2\overline{\xi^2}(z)}\right). \tag{28}$$

4.3. Data sources

In order to obtain the solution to the two PDEs and the response, the stopping power, straggling coefficient, absorption cross section and elastic scatter cross section must be known as a function of energy and tissue

composition. The CT scan HU values were converted to density and fractional compositions according to Schneider's method (Schneider, Bortfeld, and Schlegel 2000). The density and fractional composition were used to interpolate nuclide specific tables of the stopping power versus energy. The tables were extracted from TOPAS (Perl et al. 2012) using an adapted extension distributed on the TOPAS forum. The stopping power for protons in water versus energy can be seen in Figure 2.

The straggling coefficient represents the statistical variation around the mean of the energy loss of a proton in a material. The consequence of energy straggling is the spreading of the energy spectrum of an initially mono-energetic beam (Noshad and Bahador 2012). The equation that was used for the straggling coefficient is (Williams and Bragg 1932)

$$T(E, N_{\mathbb{A}}(z)) = \sum_{i\in\mathbb{A}} \frac{1}{(4\pi\epsilon_0)^2} N_i(z) 4\pi e^4 Z_i(z) \left(1 + \frac{4I_i(z)}{3m_e v_p^2} \ln \frac{2m_e v_p^2}{I_i(z)}\right), \tag{29}$$

where $N_{\mathbb{A}}$ is the set of atomic densities corresponding to the set of atoms atoms $\mathbb{A}$ that were considered to constitute human tissue, namely $\mathbb{A} = \{\text{H, C, N, O, Na, Mg, P, S, Cl, Ar, K, Ca}\}$. Moreover, Z_i is the atomic number of the target atom i with $i \in \mathbb{A}$, ϵ_0 is the vacuum permitivity constant, e is the elementary charge, m_e is the electron mass, v_p is the proton speed, I_i is the mean atomic excitation energy of atom i. The straggling coefficient for protons versus energy in water can be seen in Figure 3.

The elastic scatter cross section can be found by considering the deflection that a proton suffers due to the Coulomb field of the nucleus. A derivation of this can be found in the work of Goldstein (Goldstein, Poole, and Safko 2002) who gives the microscopic elastic scatter cross section for protons incident on a target nucles $t, t \in \mathbb{A}$ with atomic number Z_t and atomic mass numbers A_t as

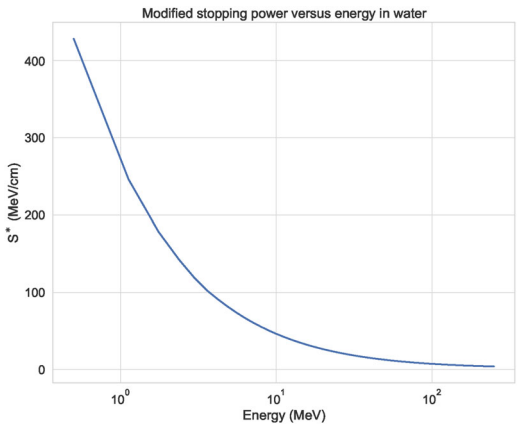

Figure 2. Water stopping power versus energy.

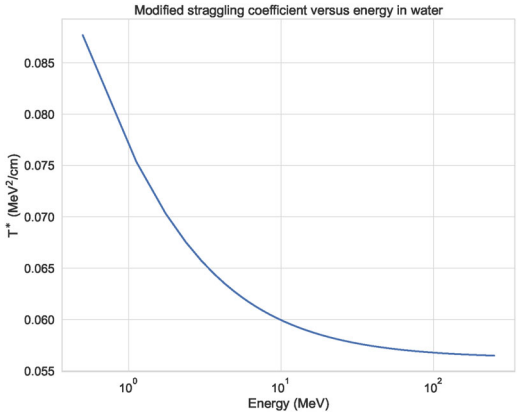

Figure 3. Water straggling coefficient versus energy.

$$\sigma_{s,t}(E,\mu,z) = \frac{\left(1+\frac{2\mu}{A_t(z)}+\frac{1}{A_t(z)^2}\right)^{3/2}}{1+\frac{\mu}{A_t(z)}}\left(\frac{Z_t(z)e^2}{4\pi\epsilon_0 m_0 v_p^2}\right)^2 \frac{1}{(1-\mu+2\eta(z))^2}, \tag{30}$$

where m_0 is the reduced mass which is defined by

$$\frac{1}{m_0} = \frac{1}{m_p} + \frac{1}{m_t(z)}$$

with m_p the mass of the proton and m_t the mass of the target nucleus, v_p is the incident speed of the proton, ϵ_0 is the vacuum permittivity, e is the elementary charge and

$$\eta(z) = \Theta^2_{\min}(z) = \left(\frac{Z_t^{1/3}(z)\alpha m_e c}{p}\right)^2$$

with m_e the electron mass, α the fine structure constant, c the speed of light and p the momentum of the incident proton. Equation (30) is used to define the macroscopic scatter cross section as

$$\Sigma_s(\mu, E, N_{\mathbb{A}}(z)) = \sum_{i\in\mathbb{A}} N_i(z)\sigma_{s,i}(E,\mu,z) \tag{31}$$

with $N_i, i \in \mathbb{A}$ the individual atomic density in the material under consideration. This cross section can be seen in Figure 4 for a range of energies.

5. Response change

Using the Fokker-Planck flux φ_{FP}, the FE coefficients from Equation (23) and a given set of system parameters $\boldsymbol{\alpha}$ the deposited energy in an arbitrary ROI is computed via

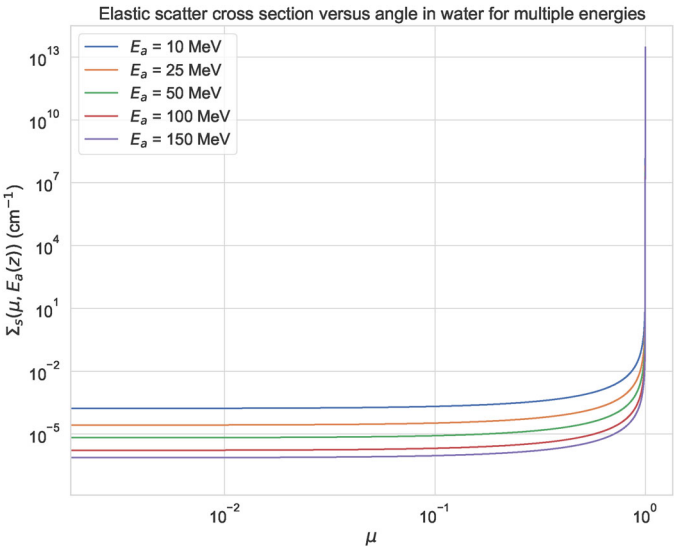

Figure 4. Water elastic scatter cross section for protons versus angle for multiple energies.

$$R(\boldsymbol{\alpha}, \varphi_{FP}) = -\int_{ROI} \mathrm{d}V \int_{4\pi} \mathrm{d}\hat{\boldsymbol{\Omega}} \int_{E_{\min}}^{E_{\max}} \mathrm{d}E \left[E\frac{\partial S\varphi}{\partial E} + E\frac{\partial}{\partial E}\left(T\frac{\partial \varphi}{\partial E}\right) - E\Sigma_a\varphi \right],$$

where the z and E dependency of S, T and Σ_a has been supressed for notational conveniency. It is of interest to assess how the response changes depending on changes in the system parameters $\boldsymbol{\alpha}$. This change in the response can be described as a direct and an indirect change. The direct change is the one that results from the change in the system parameters being directly used to compute the response. The indirect change comes through the FP flux and FE coefficients which are perturbed when changes in the system parameters are present. Thus, for each new vector of system parameters $\boldsymbol{\alpha}$ a new solution to the Fokker-Planck and Fermi-Eyges systems must be obtained.

This section describes the functional relationship between the change in the response δR and the changes in the system parameters $\delta\boldsymbol{\alpha}$ and the FP flux $\delta\varphi_{FP}$. Moreover, it describes a methodology that allows cheaply evaluating the desired change in the response without re-computing the Fokker-Planck flux and the Fermi-Eyges coefficients.

5.1. The change in the response

The response can be written using Equation (5) as

$$R(\boldsymbol{\alpha}, \varphi_{FP}) = -\int_{ROI} \mathrm{d}V \int_{4\pi} \mathrm{d}\hat{\boldsymbol{\Omega}}\varphi_{FE} \int_{E_{\min}}^{E_{\max}} \mathrm{d}E \left[E\frac{\partial S\varphi_{FP}}{\partial E} + E\frac{\partial}{\partial E}\left(T\frac{\partial \varphi_{FP}}{\partial E}\right) - E\Sigma_a\varphi_{FP} \right]. \quad (32)$$

Making use of the definition from Equation (28) for the 4π integrated Fermi-Eyges flux allows writing the response from Equation (32) in a shortened form

$$R(\boldsymbol{\alpha}, \varphi_{FP}) = \int_{ROI} \mathrm{d}V \Psi_{FE}(\mathbf{r}) D_{FP}(z, \varphi_{FP}, \boldsymbol{\alpha}), \tag{33}$$

where

$$D_{FP}(z, \varphi_{FP}, \boldsymbol{\alpha}) = \left[ES\varphi_{FP}\Big|_{E_{\min}} + \int_{E_{\min}}^{E_{\max}} \mathrm{d}ES\varphi_{FP} + \int_{E_{\min}}^{E_{\max}} \mathrm{d}ET \frac{\partial \varphi_{FP}}{\partial E} + \sum_{\Gamma_i} -[\varphi]T + \int_{E_{\min}}^{E_{\max}} \mathrm{d}EE\Sigma_a \varphi_{FP} \right].$$

The change in the response due to changes in the system parameters can be found by computing the Gateaux-differential. Let $\boldsymbol{e^0} = (\boldsymbol{\alpha}, \varphi_{FP})$ and $\mathbf{h} = (\delta\boldsymbol{\alpha}, \delta\varphi_{FP})$. Then,

$$\delta R(\boldsymbol{e^0}, \boldsymbol{h}) = \frac{\mathrm{d}}{\mathrm{d}t} R(\boldsymbol{e^0} + t\boldsymbol{h})\Big|_{t=0} = \int_{ROI} \mathrm{d}V[\delta\Psi_{FE}(\mathbf{r}) D_{FP}(z) + \Psi_{FE}(\mathbf{r}) \delta D_{FP}(z)]. \tag{34}$$

The Gateaux-differential of D_{FP} is computed to be

$$\begin{aligned}
\delta D_{FP} = &\left\{ E\delta S\varphi_{FP}\Big|_{E_{\min}} + \int_{E_{\min}}^{E_{\max}} \mathrm{d}E\delta S\varphi_{FP} + \int_{E_{\min}}^{E_{\max}} \mathrm{d}E\delta T \frac{\partial \varphi_{FP}}{\partial E} + \sum_{\Gamma_i} -[\varphi]\delta T + \int_{E_{\min}}^{E_{\max}} \mathrm{d}EE\delta\Sigma_a \varphi_{FP} \right\} \\
&+ \left\{ ES\delta\varphi_{FP}\Big|_{E_{\min}} + \int_{E_{\min}}^{E_{\max}} \mathrm{d}ES\delta\varphi_{FP} + \int_{E_{\min}}^{E_{\max}} \mathrm{d}ET \frac{\partial \delta\varphi_{FP}}{\partial E} + \sum_{\Gamma_i} -[\delta\varphi]T + \int_{E_{\min}}^{E_{\max}} \mathrm{d}EE\Sigma_a \delta\varphi_{FP} \right\} \\
&= \delta D_{FP,\,dir}(\delta\boldsymbol{\alpha}, \varphi_{FP}) + \delta D_{FP,\,indir}(\boldsymbol{\alpha}, \delta\varphi_{FP})
\end{aligned}$$

where $\delta D_{FP,\,dir}(\delta\boldsymbol{\alpha}, \varphi_{FP})$ is the direct change in the response due to the change in the system parameter vector and $\delta D_{FP,\,indir}(\boldsymbol{\alpha}, \delta\varphi_{FP})$ is the indirect change in the response due to the perturbation in the dependent system variable.

The Gateaux-differential of $\Psi_{FE}(\boldsymbol{r})$ is found to be

$$\begin{aligned}\delta\Psi_{FE}(\boldsymbol{r}) &= \frac{\mathrm{d}}{\mathrm{d}t}\Psi_{FE}(A + t\delta A, \overline{\xi^2}(z) + t\delta\overline{\xi^2}(z))\Big|_{t=0} \\ &= \Psi_{FE}\left[\frac{2\delta A}{A} - \frac{\delta\overline{\xi^2}(z)}{\overline{\xi^2}(z)} + \frac{x^2+y^2}{2}\frac{\delta\overline{\xi^2}(z)}{\overline{\xi^2}(z)^2}\right]. \end{aligned} \tag{35}$$

The quantity $\delta\Psi_{FE}(\boldsymbol{r})$ is thereafter laterally integrated over the X and Y extents of the ROI, namely

$$\begin{aligned}\iint\limits_{ROI_{XY}} \mathrm{d}x\mathrm{d}y\delta\Psi_{FE} &= \iint\limits_{ROI_{XY}} \mathrm{d}x\mathrm{d}y\Psi_{FE}\left[\frac{2\delta A}{A} - \frac{\delta\overline{\xi^2}(z)}{\overline{\xi^2}(z)} + \frac{x^2+y^2}{2}\frac{\delta\overline{\xi^2}(z)}{\overline{\xi^2}(z)^2}\right] \\ &= \left[\frac{2\delta A}{A} - \frac{\delta\overline{\xi^2}(z)}{\overline{\xi^2}(z)}\right]\underbrace{\iint\limits_{ROI_{XY}} \mathrm{d}x\mathrm{d}y\Psi_{FE}}_{I_{f1}} + \frac{\delta\overline{\xi^2}(z)}{\overline{\xi^2}(z)^2}\underbrace{\iint\limits_{ROI_{XY}} \mathrm{d}x\mathrm{d}y\frac{x^2+y^2}{2}\Psi_{FE}}_{I_{f2}} \\ &= \left[\frac{2\delta A}{A} - \frac{\delta\overline{\xi^2}(z)}{\overline{\xi^2}(z)}\right]I_{f,1}(z) + \frac{\delta\overline{\xi^2}(z)}{\overline{\xi^2}(z)^2}I_{f,2}(z). \end{aligned} \tag{36}$$

At this point the term δA is set to zero. This term is only non-zero when the Fermi-Eyges initial condition is perturbed. For the purpose of this work, no such perturbation was included. Thus, the laterally integrated $\delta\Psi_{FE}$ is given as

$$\iint\limits_{ROI_{XY}} \mathrm{d}x\mathrm{d}y\delta\Psi_{FE} = \delta\overline{\xi^2}(z)\left[-\frac{I_{f,1}(z)}{\overline{\xi^2}(z)} + \frac{I_{f,2}(z)}{\overline{\xi^2}(z)^2}\right]. \tag{37}$$

The term $\delta\overline{\xi^2}(z)$ contains the unknown $\delta\varphi_{FP}$ via the implicit E_a computation in the elastic scatter cross section. Continuing the Gateaux-differential process results in

$$\delta\overline{\xi^2}(z) = \int\limits_0^z (z-z')^2\delta\Sigma_{tr}(z')\mathrm{d}z' \text{ where } \delta\Sigma_{tr}(z) = \int\limits_{-1}^{1} \delta\Sigma_s(E_a(z),\mu)(1-\mu)\mathrm{d}\mu. \tag{38}$$

As the system parameters change, so does the Fokker-Planck flux φ_{FP}. This in turn results in a change in the average depth-dependent energy $E_a(z)$ of the beam which in turn results ultimately in changes in the FE coefficients. To compute the effect of a change in the FP flux on the FE coefficients, the elastic scatter cross section can be re-written to illustrate the energy dependence by using the classical kinetic energy relationship between speed and energy as

$$\Sigma_s(E_a(z),\mu,N_{\mathbb{A}}) = \sum_{i\in\mathbb{A}} N_i F_1(\mu,A_i) F_2(Z_i,m_{0i}) \frac{1}{E_a^2} \frac{1}{\left(1-\mu+\frac{2c_{\eta,i}}{E_a}\right)^2},$$

$$\text{where } F_1(\mu,A_i) = \frac{\left(1+\frac{2\mu}{A_i}+\frac{1}{A_i^2}\right)^{3/2}}{1+\frac{\mu}{A_i}},\ F_2(Z_i,m_{0i}) = \left(\frac{Z_i e^2 m_p}{8\pi\epsilon_0 m_{0i}}\right)^2 \text{ and } c_{\eta,i} = E_a\eta_i.$$

The Gateaux-differential of the elastic scatter cross section is

$$\begin{aligned}
\delta\Sigma_s(E_a(z),\mu,N_{\mathbb{A}}) &= \frac{\mathrm{d}}{\mathrm{d}t}\sum_{i\in\mathbb{A}}(N_i+t\delta N_i)F_1(\mu,A_i)F_2(Z_i,m_{0i})\frac{1}{(E_a(z)+t\delta E_a(z))^2}\frac{1}{\left(1-\mu+\frac{2c_{\eta,i}}{E_a(z)+t\delta E_a(z)}\right)^2}\Bigg|_{t=0} \\
&= \sum_{i\in\mathbb{A}} \delta N_i F_1(\mu,A_i)F_2(Z_i,m_{0i})\frac{1}{E_a(z)^2}\frac{1}{\left(1-\mu+\frac{2c_{\eta,i}}{E_a(z)}\right)^2} \\
&+ \sum_{i\in\mathbb{A}} N_i F_1(\mu,A_i)F_2(Z_i,m_{0i})\frac{\delta E_a(z)}{E_a(z)^2}\frac{1}{\left(1-\mu+\frac{2c_{\eta,i}}{E_a(z)}\right)^2}\left[\frac{-2}{E_a(z)}+\frac{4c_{\eta,i}}{E_a^2(z)}\frac{1}{1-\mu+\frac{2c_{\eta,i}}{E_a(z)}}\right] \\
&= \delta\Sigma_{s1}(\delta N_{\mathbb{A}}) + \delta\Sigma_{s2}(\delta E_a(z)),
\end{aligned} \tag{39}$$

where $\delta N_{\mathbb{A}}$ is the set of perturbations in the atomic density set previously defined as $N_{\mathbb{A}}$. It can be seen that here as well the change in the elastic scatter cross section can be described as a direct change due to the atomic composition in the tissue $\delta N_{\mathbb{A}}$ and the change due to the change in the beam energy spectrum. The last component is to obtain the relationship between the change in the energy spectrum $\delta E_a(z)$ and the change in the flux itself $\delta\varphi_{FP}$. The Gateaux-differential of the average depth-dependent energy is

$$\delta E_a(z) = \frac{1}{N_p(z)}\int_0^\infty \mathrm{d}E E\delta\varphi_{FP} - \frac{E_b(z)}{N_p(z)^2}\int_0^\infty \mathrm{d}E\delta\varphi_{FP}, \tag{40}$$

where the number of particles and the beam energy at a given depth are defined as

$$N_p(z) = \int_0^\infty \mathrm{d}E\varphi_{FP}(z,E) \text{ and} \tag{41}$$

$$E_b(z) = \int_0^\infty \mathrm{d}E E\varphi_{FP}(z,E). \tag{42}$$

At this point the functional relationship between δR and the changes in $\delta\boldsymbol{\alpha}$ and $\delta\varphi_{FP}$ can be obtained. The first step is to introduce the result from Equation (39) into the Gateaux-differential of $\overline{\xi^2}$ given in Equation (38), yielding

$$\delta\overline{\xi^2}(z) = \int_0^z \mathrm{d}z'(z-z')^2\delta\Sigma_{tr}(z') = \int_0^z \mathrm{d}z'(z-z')^2 \int_{-1}^{1} \mathrm{d}\mu[\delta\Sigma_{s1}(\delta N_{\mathbb{A}}) + \delta\Sigma_{s2}(\delta E_a(z))](1-\mu)$$
$$= \delta\overline{\xi^2}_1(\delta N_{\mathbb{A}}) + \delta\overline{\xi^2}_2(\delta E_a(z)). \tag{43}$$

The term $\delta\overline{\xi^2}_1$ does not contain any dependencies on the unknown $\delta\varphi_{FP}$ and contributes to the direct effect. The term $\delta\overline{\xi^2}_2$ does on the other hand contain a dependency on $\delta\varphi_{FP}$. Its $\delta\varphi_{FP}$ dependency is obtained by using Equation (40), namely

$$\delta\overline{\xi^2}_2 = \int_0^z \mathrm{d}z'(z-z')^2 \underbrace{\int_{-1}^{1} \mathrm{d}\mu(1-\mu)\sum_{i\in\mathbb{A}} N_i F_1(\mu, A_i)F_2(Z_i, m_{0i})\frac{\delta E_a(z')}{E_a(z')^2}\frac{1}{\left(1-\mu+\frac{2c_{\eta,i}}{E_a(z')}\right)^2}\left[\frac{-2}{E_a(z')} + \frac{4c_{\eta,i}}{E_a^2(z')}\frac{1}{1-\mu+\frac{2c_{\eta i}}{E_a(z')}}\right]}_{\delta E_a(z')\cdot I_\mu(z')}$$
$$= \int_0^z \mathrm{d}z'(z-z')^2\delta E_a(z')I_\mu(z') = \int_0^z \mathrm{d}z'(z-z')^2 I_\mu(z')\int_0^\infty \mathrm{d}E\left(\frac{E}{N_p(z')} - \frac{E_b(z')}{N_p(z')^2}\right)\delta\varphi_{FP}(z', E).$$

The next step is to introduce the expression from Equation (43) into the lateral ROI integration of $\delta\Psi_{FE}$ from Equation (37). For simplicity of notation, let

$$\Psi_{FE}^{ROI_{xy}}(z) = \int_{ROI_{xy}(z)} \mathrm{d}x\mathrm{d}y\Psi_{FE}.$$

Using this, δR becomes

$$\delta R = \int_{ROI_z} \mathrm{d}z D_{FP}(z) \int_{ROI_{xy}} \mathrm{d}x\mathrm{d}y\delta\Psi_{FE} + \int_{ROI_z} \mathrm{d}z\Psi_{FE}^{ROI_{xy}}(z)\delta D_{FP}(z). \tag{44}$$

Introducing in this expression the direct and indirect contributions results in

$$\delta R = \int_{ROI_z} \mathrm{d}z\left\{D_{FP}(z)\delta\overline{\xi^2}(z)\left[-\frac{I_{f,1}(z)}{\overline{\xi^2}(z)} + \frac{I_{f,2}(z)}{\overline{\xi^2}(z)^2}\right]\right\}$$
$$+ \int_{ROI_z} \mathrm{d}z\Psi_{FE}^{ROI_{xy}}(z)[\delta D_{FP,dir}(\delta\boldsymbol{\alpha}) + \delta D_{FP,indir}(\delta\varphi_{FP})] \tag{45}$$

$$
\begin{aligned}
&= \int_{ROI_z} \mathrm{d}z D_{FP}(z)\delta\overline{\xi^2}_1(\delta N_{\mathbb{A}})\left[-\frac{I_{f,1}(z)}{\overline{\xi^2}(z)}+\frac{I_{f,2}(z)}{\overline{\xi^2}(z)^2}\right] \rightarrow \text{ direct change due to } \delta\boldsymbol{\alpha} \\
&+ \int_{ROI_z} \mathrm{d}z D_{FP}(z)\delta\overline{\xi^2}_2(\delta\varphi_{FP})\left[-\frac{I_{f,1}(z)}{\overline{\xi^2}(z)}+\frac{I_{f,2}(z)}{\overline{\xi^2}(z)^2}\right] \rightarrow \text{ indirect change due to } \delta\varphi_{FP} \\
&+ \int_{ROI} \mathrm{d}z \Psi_{FE}^{ROI_{xy}}(z)\delta D_{FP,dir}(\delta\boldsymbol{\alpha},\varphi_{FP}) \rightarrow \text{ direct change due to } \delta\boldsymbol{\alpha} \\
&+ \int_{ROI} \mathrm{d}z \Psi_{FE}^{ROI}(z)\delta D_{FP,indir}(\boldsymbol{\alpha},\delta\varphi_{FP}) \rightarrow \text{indirect change due to } \delta\varphi_{FP}.
\end{aligned}
\tag{46}
$$

5.2. Relating $\delta\varphi$ to $\delta\alpha$

The expression in Equation (46) shows that the response change depends on the changes in the system parameter vector $\delta\boldsymbol{\alpha}$ and the corresponding change in the Fokker-Planck flux $\delta\varphi_{FP}$. Since the Fokker-Planck flux and the vector of system parameters are related through Equation (10) and its associated boundary conditions it must also be that the perturbations in both of these quantities are related. A first-order relationship between $\delta\varphi_{FP}$ and $\delta\boldsymbol{\alpha}$ can be obtained by taking the Gateaux-differential of the Fokker-Planck equation and it's boundary conditions (Cacuci 2003). In this process a new PDE is obtained for the unknown $\delta\varphi_{FP}$ as a function of the initial operator $L(\boldsymbol{\alpha})$ and the perturbations in the system parameters $\delta\boldsymbol{\alpha}$. Given that this PDE has to be solved for each new vector of system parameter perturbations and the number of such vectors is large, this process is computationally too expensive to implement in practice.

5.3. Adjoint sensitivity analysis procedure

An alternative to the procedure from subsection 5.2 is the Adjoint Sensitivity Analysis Procedure (ASAP). ASAP aims to eliminate the unknown value of $\delta\varphi_{FP}$ from the response change Equation (46). This is done by constructing a new system called the adjoint system that is independent of $\delta\varphi_{FP}$ with the property that the solution of this system (denoted by $\varphi^{\dagger}$) can be used to eliminate the unknown $\delta\varphi_{FP}$ from Equation (46). In the process of constructing the adjoint system the boundary conditions that ensure its unique solution will also have to be imposed. These must be chosen such that (Cacuci 2003):

- they are independent of $\delta\varphi_{FP}$, $\delta\boldsymbol{\alpha}$ and Gateaux-derivatives with respect to $\boldsymbol{\alpha}$, and
- the evaluation of boundary terms does not contain unknown values of $\delta\varphi_{FP}$.

5.4. Adjoint system derivation

The starting point of the adjoint system derivation is the inner product between the adjoint flux and the operator acting on the perturbation $\delta\varphi_{FP}$, namely

$$\langle\varphi^{\dagger}, L(\boldsymbol{\alpha})\delta\varphi_{FP}\rangle = \int_0^{\infty} \mathrm{d}z \int_0^{\infty} \mathrm{d}E\varphi^{\dagger}\left[\frac{\partial\delta\varphi_{FP}}{\partial z} - \frac{\partial S\delta\varphi_{FP}}{\partial E} - \frac{\partial}{\partial E}\left(T\frac{\partial\delta\varphi_{FP}}{\partial E}\right) + \Sigma_a\delta\varphi_{FP}\right] \tag{47}$$

At this point we extend φ_{FP} and consequently $\delta\varphi_{FP}$ to the whole $\mathbb{R}^2$ plane with the condition that these quantities are zero everywhere outside of the computational domain $\mathscr{D}$. Through partial integration along the z-direction for the first term, and along the E-direction for the stopping power and range straggling terms, Equation (47) is found to be equal to

$$\begin{aligned}\langle\varphi^{\dagger}, L(\boldsymbol{\alpha})\delta\varphi_{FP}\rangle &= \int_0^{\infty} \mathrm{d}E\varphi^{\dagger}(0,E)\delta\varphi_{FP}(0,E) + \left\langle -\frac{\partial\varphi^{\dagger}}{\partial z} + S\frac{\partial\varphi^{\dagger}}{\partial E} - \frac{\partial}{\partial E}T\frac{\partial\varphi^{\dagger}}{\partial E} + \Sigma_a\varphi^{\dagger}, \delta\varphi_{FP}\right\rangle \\ &= \int_0^{\infty} \mathrm{d}E\varphi^{\dagger}(0,E)\delta\varphi_{FP}(0,E) + \langle L^{\dagger}(\alpha^0)\varphi^{\dagger}, \delta\varphi_{FP}\rangle.\end{aligned} \tag{48}$$

In the process of deriving Equation (48) the adjoint operator $L^{\dagger}$ together with its associated boundary conditions were found to be

$$L^{\dagger}\varphi^{\dagger} = -\frac{\partial\varphi^{\dagger}}{\partial z} + S\frac{\partial\varphi^{\dagger}}{\partial E} - \frac{\partial}{\partial E}\left(T\frac{\partial\varphi^{\dagger}}{\partial E}\right) + \Sigma_a\varphi^{\dagger} \tag{49}$$

$$\mathrm{BCE}: \ \varphi^{\dagger}(z, E_{\min}) = 0, \left.\frac{\partial\varphi^{\dagger}}{\partial E}\right|_{E=E_{\min}} = 0, \tag{50}$$

$$\mathrm{BCS}: \ \varphi^{\dagger}(z_{\max}, E) = 0. \tag{51}$$

To achieve the desired $\delta\varphi_{FP}$ elimination from Equation (46) we note that δR is linear in both $\delta\boldsymbol{\alpha}$ and $\delta\varphi_{FP}$. This allows writing the Gateaux-differential of the response as (Cacuci 2003)

$$\delta R(\boldsymbol{e}^0; \boldsymbol{h}) = R'_{\varphi}(\boldsymbol{e}^0)\delta\varphi_{FP} + R'_{\alpha}(\boldsymbol{e}^0)\delta\boldsymbol{\alpha}. \tag{52}$$

The quantity denoted as $R'_{\varphi}(\boldsymbol{e}^0)\delta\varphi_{FP}$, identified as

$$R'_{\varphi}(\boldsymbol{e}^0)\delta\varphi_{FP} = \int\limits_{ROI_z} \mathrm{d}z D_{FP}(z)\delta\overline{\xi^2}_2(\delta\varphi_{FP})\left[-\frac{I_{f,1}(z)}{\overline{\xi^2}(z)} + \frac{I_{f,2}(z)}{\overline{\xi^2}(z)^2}\right] + \int\limits_{ROI} \mathrm{d}z \Psi_{FE}^{ROI_{xy}}(z)\delta D_{FP,\,indir}(\delta\varphi_{FP}),$$

is itself also linear in $\delta\varphi_{FP}$. Coupling this with the self-duality of Hilbert spaces certifies the application of Riesz's representation theorem. Using this theorem the quantity $R'_{\varphi}(\boldsymbol{e}^0)\delta\varphi_{FP}$ can be written as an inner product between a quantity $r^{\dagger} \in \mathscr{H}$ and $\delta\varphi_{FP}$ (Cacuci 2003), namely,

$$R'_{\varphi}(\boldsymbol{e}^0)\delta\varphi_{FP} = \langle r^{\dagger}, \delta\varphi_{FP}\rangle. \tag{53}$$

Identifying $r^{\dagger}$ as the right-hand side of the adjoint system allows writing the Gâteaux-differential of the response as

$$\delta R(\boldsymbol{e}^0;\boldsymbol{h}) = \langle r^{\dagger}, \delta\varphi_{FP}\rangle + R'_{\alpha}(\boldsymbol{e}^0)\delta\boldsymbol{\alpha} = \langle L^{\dagger}(\boldsymbol{\alpha})\varphi^{\dagger}, \delta\varphi_{FP}\rangle + R'_{\alpha}(\boldsymbol{e}^0)\delta\boldsymbol{\alpha}. \tag{54}$$

The inner product in the second equality has already been computed in Equation (48). Thus,

$$\delta R(\boldsymbol{e}^0;\boldsymbol{h}) = \langle \varphi^{\dagger}, L(\boldsymbol{\alpha})\delta\varphi_{FP}\rangle - \int\limits_0^{\infty} \mathrm{d}E\varphi^{\dagger}(0,E)\delta\varphi_{FP}(0,E) + R'_{\alpha}(\boldsymbol{e}^0)\delta\boldsymbol{\alpha}. \tag{55}$$

The quantity $L(\boldsymbol{\alpha})\delta\varphi_{FP}$ can be derived by taking the Gateaux-differential of the Fokker-Planck equation, and is found to be

$$L(\boldsymbol{\alpha})\delta\varphi_{FP} = -\left[L'_{\alpha}(\boldsymbol{\alpha})\varphi_{FP}\right]\delta\boldsymbol{\alpha} = \frac{\partial\delta S\varphi_{FP}}{\partial E} + \frac{\partial}{\partial E}\left(\delta T\frac{\partial\varphi_{FP}}{\partial E}\right) - \delta\Sigma_a\varphi_{FP}. \tag{56}$$

Making use of this transforms $\delta R(\boldsymbol{e}^0;\boldsymbol{h})$ to

$$\delta R(\boldsymbol{e}^0;\boldsymbol{h}) = \langle \varphi^{\dagger}, -\left[L'_{\alpha}(\alpha)\varphi_{FP}\right]\delta\boldsymbol{\alpha}\rangle - \int\limits_0^{\infty} \mathrm{d}E\varphi^{\dagger}(0,E)\delta\varphi(0,E) + R'_{\alpha}(\boldsymbol{e}^0)\delta\boldsymbol{\alpha}, \tag{57}$$

where the first inner product is equal to

$$\langle \varphi^{\dagger}, -\left[L'_{\alpha}(\alpha)\varphi_{FP}\right]\delta\boldsymbol{\alpha}\rangle = \int_0^{\infty}\mathrm{d}z\int_0^{\infty}\mathrm{d}E\varphi^{\dagger}\left[\frac{\partial}{\partial E}\delta S\varphi_{FP} + \frac{\partial}{\partial E}\delta T\frac{\partial\varphi_{FP}}{\partial E} - \delta\Sigma_a\varphi_{FP}\right]. \tag{58}$$

For simplicity, the quantity $\delta\varphi(0,E)$ was set to zero in the case of this work and $R'_{\alpha}(\boldsymbol{e}^0)\delta\boldsymbol{\alpha}$ is the direct change that has been previously computed

in Equation (46) as

$$R'_{\alpha}(\boldsymbol{e}^0)\delta\boldsymbol{\alpha} = \int_{ROI_z} \mathrm{d}z D_{FP}(z)\delta\overline{\xi^2}_1(\delta N_i)\left[-\frac{I_{f,1}(z)}{\overline{\xi^2}(z)} + \frac{I_{f,2}(z)}{\overline{\xi^2}(z)^2}\right] + \int_{ROI_z} \mathrm{d}z \Psi_{FE}^{ROI}(z)\delta D_{FP,dir}(\delta\boldsymbol{\alpha}). \tag{59}$$

As can be seen from Equations (57) to (59) the goal of the ASAP has been reached. The indirect change in the response due to the perturbation of the Fokker-Planck flux $\delta\varphi_{FP}$ has been replaced in Equation (46) by the inner product from Equation (58). Thus, for a given number N of different vectors of system parameters $\boldsymbol{\alpha}$ the computational expense has been decreased from the initial N necessary solutions of the 1DFP system to just two solutions, namely those of the adjoint and Fokker-Planck systems, with a similar computational expense for both systems.

6. Results and discussion

This section details the computational set-up of the algorithm in Subsection 6.1, the comparison between the dose calculation of our engine and those of TOPAS and Bortfeld's method in Subsection 6.2 and the accuracy of the previously illustrated ASAP methodology for response change computations in Subsection 6.3.

6.1. Computational set-up

The domain of the CT scan was discretized using an arbitrarily chosen number of 51 bins in the X and Y direction and 100 bins in the Z direction. The spatial extent of the CT scan was set to -2 to 2 cm in the X and Y directions and 0 to 10 cm in the Z direction. Within this geometry a slab was placed along the depth of the tank with its depth and precise position being variable. The slab had variable HU values set while the rest of the tank was set to either 0 HU (water) or the arbitrarily chosen value of 550 HU. The ROI was defined to be a box with variable extents in all three directions. An illustration of this set-up can be seen in Figure 5

In all test cases the beam started at the point of $\boldsymbol{r}_{start} = (0, 0, 0)$ and ended at $\boldsymbol{r}_{end} = (0, 0, 10)$. The tracking of the beam within the geometry was performed using an in-house ray-tracing procedure based on work of

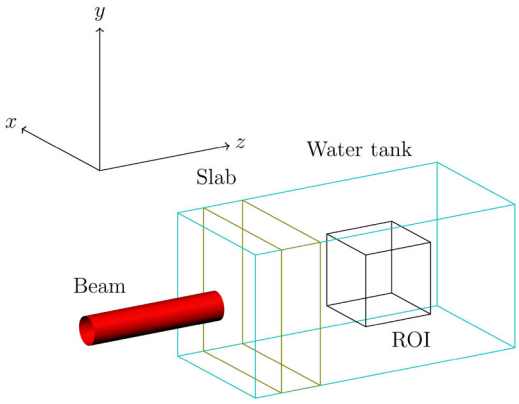

Figure 5. Illustration of the CT scan, the slab with perturbed HU values, the ROI and the incident beam.

de Sutter et al. (Sundermann et al. 1998) with a maximal step size set to 0.01 cm.[2]

The initial spread of the Gaussian σ_ξ from Equation (25) in X and Y was set to 0.3 cm and the space angle correlation ϱ from Equation (25) was set to zero. Due to the singularity in the angular spread variable of the coefficients from Equation (25), σ_ω could not be set to zero. However, it was found that values below 10^{-4} did not affect the resulting energy deposition distributions and thus the angular spread was set to the dimensionless value of 10^{-8}. The energy domain was fixed between $E_{\min} = 1\,\text{MeV}, E_{\max} = 200\,\text{MeV}$ with a number of groups of $NG = 300$. The coefficients of the energy initial condition from Equation (11) were set to correspond to a normal distribution and were matched such that the number of particles was either 1 or 2×10^7. Moreover, the initial beam average energy E_0 from Equation (12) was set to 100 MeV while the energy spread σ_E from the same equation was set to 0.757 504 MeV. The energy spread value was chosen to match the standard value that TOPAS initializes for a proton pencil beam.

6.2. Forward results

To gauge the accuracy of the reponse computation engine, it was benchmarked against the TOPAS MC algorithm (Perl et al. 2012). In a homogeneous water tank the energy deposition can readily be converted to dose deposition. Laterally integrating this three dimensional quantity results in the dose-depth curve named integrated depth dose (IDD),

[2]While a von Neumann stability analysis was not performed, through empirical observations it was found that in water the maximal step size for accurate, artifact free fluxes outputted by the CN scheme is 0.01 cm. This was used in a wrapper function for the ray-tracing procedure to divide each segment into a corresponding number of sub-segments.

$$IDD(z) = \int\int_{X,Y} \mathrm{d}x\mathrm{d}yD(x,y,z). \tag{60}$$

The comparison against TOPAS and Bortfeld's popular pencil beam algorithm (Bortfeld 1997) can be seen in Figure 6 showing that our algorithm is capable of accurately predicting the dose in the Bragg peak, the region of most clinical interest. Our algorithm slightly over-estimates the dose in the entrance region, due to the assumption that 100% of the energy released in nuclear interactions is deposited locally. To confirm this, a comparison was performed where nuclear interactions were deactivated in TOPAS. This result can be seen in Figure 7 which shows near overlap between the two normalized IDDs. The slight difference in range is attributed to the slighly different stopping powers and straggling coefficients used. A refinement of the treatment of nuclear interactions can be envisioned using convolutional methods or through more simplistic fits to empirical data. However, this is not the purpose of this paper, whose main focus is the sensitivity of the algorithm.

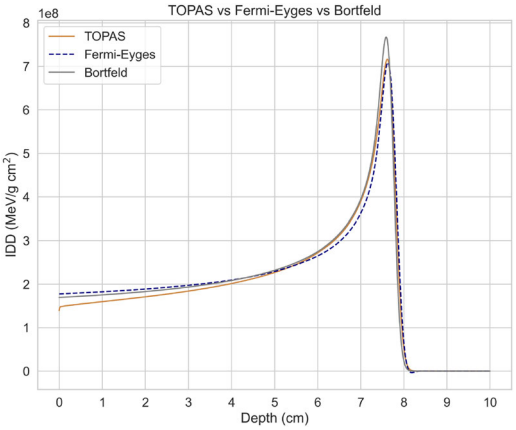

Figure 6. IDD of the in-house algorithm, TOPAS and Bortfeld's model.

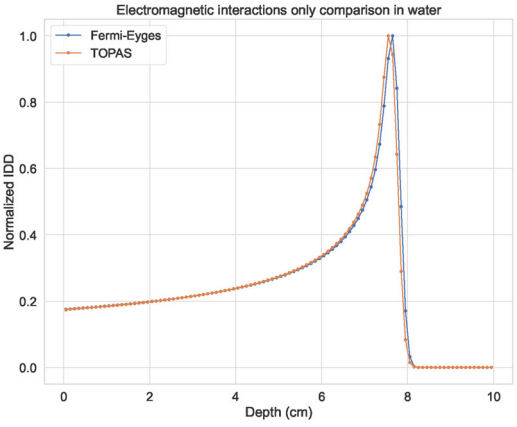

Figure 7. IDD of the in-house algorithm and TOPAS without nuclear interactions.

6.3. Adjoint results

A variety of test cases for the response change computation was performed, an overview of which can be seen in Table 1. For all test cases Σ_a was set to zero due to a lack of data for all other materials other than 0 HU.

Table 1. Overview of case numbers and the corresponding slab location and perturbation values, ROI values and maximal percentage errors between the re-computation and the adjoint result.

Case number	Slab		ROI extent			Maximal percentage error
	Location (cm)	Perturbation (HU)	X (cm)	Y (cm)	Z (cm)	
I	[2, 3]	[-40, 40]	[-2, 2]	[-2, 2]	[0, 2]	0
II	[2, 3]	[-40, 40]	[-2, 2]	[-2, 2]	[2, 6.5]	1.1×10^{-6}
III	[2, 3]	[-40, 40]	[-2, 2]	[-2, 2]	[5, 6.5]	3.6×10^{-3}
III	[2, 3]	[-40, 40]	[-0.3, 0.3]	[-0.3, 0.3]	[5, 6.5]	3.6×10^{-3}
IV	[2, 3]	[-40, 40]	[-0.3, 0]	[-0.3, 0.3]	[5, 6.5]	3.6×10^{-3}
V	[2, 3]	[-400, 400]	[-2, 2]	[-2, 2]	[7, 9]	3.0
V	[2, 3]	[-400, 400]	[-0.3, 0.3]	[-0.3, 0.3]	[7, 9]	3.0
VI	[4, 6]	[-40, 40]	[-2, 2]	[-2, 2]	[5, 6.5]	3.6×10^{-3}
VI	[4, 6]	[-40, 40]	[-0.3, 0.3]	[-0.3, 0.3]	[5, 6.5]	3.6×10^{-3}
VI	[4, 6]	[-400, 400]	[-2, 2]	[-2, 2]	[7, 9]	17.7
VI	[4, 6]	[-400, 400]	[-0.3, 0.3]	[-0.3, 0.3]	[7, 9]	17.7

6.3.1. Case I

First, a small range of [-40, 40] HU perturbations around the nominal slab value of 550 HU was used. The slab was placed in the plateau region of the energy deposition versus depth curve between 2 and 3 cm in depth. Given this set-up if the ROI has dimensions $X, Y \in [-2, 2], Z \in [0, 2]$ and under the assumption of a forward propagating (not backscattered) flux then the expectation is that the response change is identically zero. This corectness test can be seen in Figure 8.

6.3.2. Case II

Changing the ROI to be the box with $X, Y \in [-2, 2], Z \in [2, 6.5]$ results in Figure 9. There it can be seen that adjoint theory provides a first order approximation to the forward response. In this case the maximal percent error occurs for the 510 HU slab perturbation and it is equal to $\approx 1.1 \times 10^{-6}\%$.

6.3.3. Case III

In the case of proton therapy, it is likely that a ROI of practical interest is the Bragg Peak region. Thus, the slab is maintained in its previous position and the ROI is set to the box with $X, Y \in [-2, 2], Z \in [5, 6.5]$. This result can be seen in Figure 10. Another scenario of interest could be the one in which a tumor is surrounded by organs at risk. In this case, the ROI is restricted to only part of the lateral extent. In Figure 11 the lateral extent

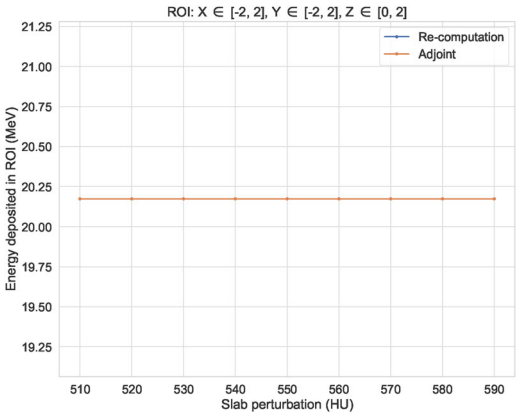

Figure 8. Adjoint versus re-computation for a small HU range with a full lateral ROI in X and Y and $Z \in [0, 2]$.

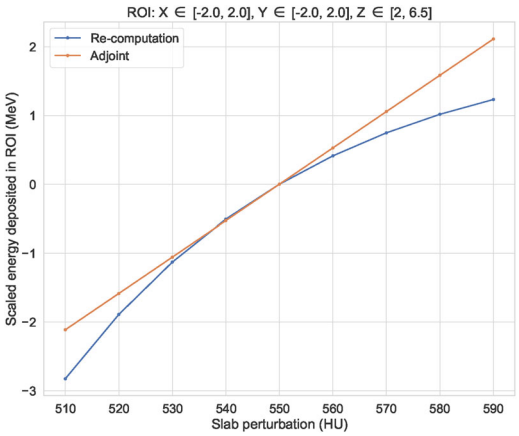

Figure 9. Adjoint versus re-computation for a small HU range with a full lateral ROI in X and Y, $Z \in [2, 6.5]$.

was constrained to $X, Y \in [-0.3, 0.3]$ while the depth was kept to $Z \in [5, 6.5]$. Both Figures 10 and 11 show that the adjoint method is capable of accurately computing the response changes due to the slab in the ROI down to a percentage error of $3.6 \times 10^{-3}\%$ that is likely clinically insignificant. This can be seen from the fact that a fraction that is typically delivered is on the order of 2 Gy which is equivalent to $\approx 6.3 \times 10^{9}\,\mathrm{MeV\,g^{-1}}$.

6.3.4. Case IV

Tests with asymmetric ROIs have also been performed. Restricting the ROI such that $X \in [-0.3, 0]$, $Y \in [-0.3, 0.3]$ and keeping $Z \in [5, 6.5]$ resulted in a similar error (depicted as the same in Table 1 due to round-off) to the previous test cases. The two curves for this case can be seen in Figure 12.

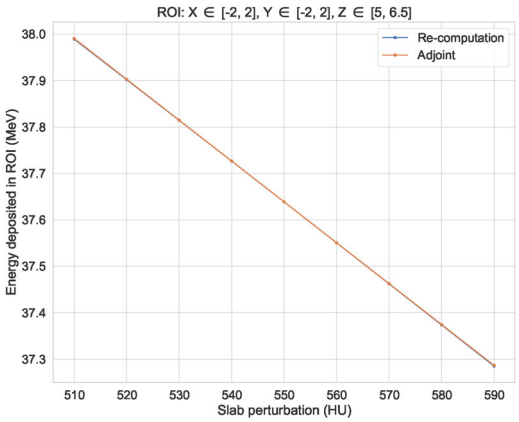

Figure 10. Adjoint versus re-computation for a small HU range with a full lateral ROI in X and Y and Z $\in$ [5, 6.5].

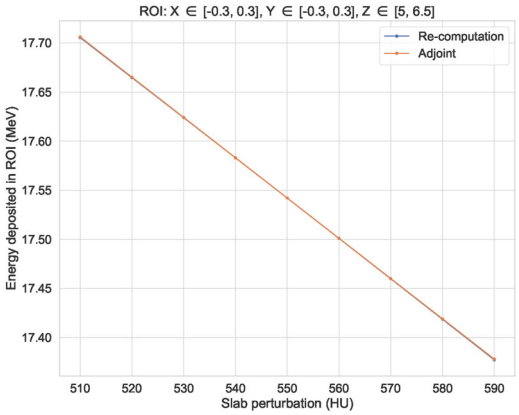

Figure 11. Adjoint versus re-computation for a small HU range with a reduced lateral ROI in X and Y and Z $\in$ [5, 6.5].

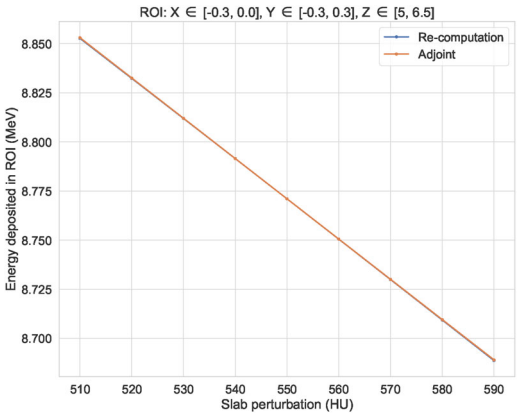

Figure 12. Adjoint versus re-computation for a small HU range with a laterally asymmetric ROI in X and Y and Z $\in$ [5, 6.5].

6.3.5. Case V

Next to the small HU range, a large [-400, 400] HU perturbation range around the nominal value of 0 HU was tested. The nominal value in this case corresponded to a homogeneous water tank. This set-up simulates more clinically relevant test-cases as the -400 HU value roughly corresponds to a tissue similar to lung while a value of 400 HU correponds to bone. As in the small perturbation range cases, ROIs with full and reduced lateral X and Y extents were tested. Figure 13 illustrates the case when $X \in [-2, 2]$, $Y \in [-2, 2], Z \in [7, 9]$ while Figure 14 illustrates the case when $X \in [-0.3, 0.3]$, $Y \in [-0.3, 0.3], Z \in [7, 9]$. It should be noted that as opposed to the straight lines shown before, these figures do not contain straight lines. This is due to the regions of discontinuity that appear in the Scheinder's conversion (Schneider, Bortfeld, and Schlegel 2000) from HU values to density

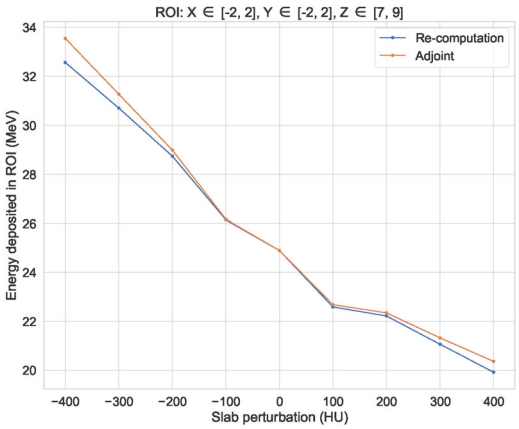

Figure 13. Adjoint versus re-computation for a large HU range with a laterally symmetric ROI and Z ∈ [7, 9].

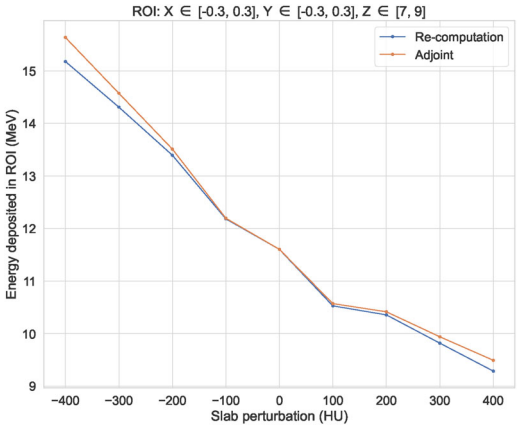

Figure 14. Adjoint versus re-computation for a large HU range with a laterally symmetric reduced ROI and Z ∈ [7, 9].

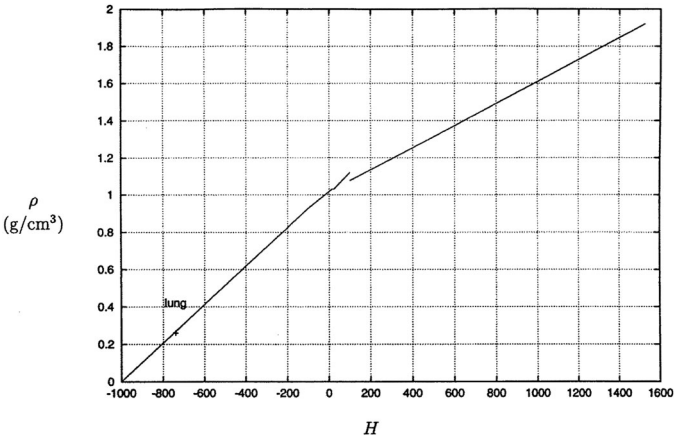

Figure 15. Illustration of the discontinuity in Schneider's HU to density conversion curve. Picture taken from (Schneider, Bortfeld, and Schlegel 2000). [The image is reproduced with permission from IOP Publishing, Ltd.]

and atomic composition. The discontinuity for the density to HU converison can be seen in Figure 15 and it is also why the range of 510–590 was chosen initially, as in this range the afore mentioned conversion is continuous.

6.3.6. Case VI

Tests were also performed with a 2 cm slab placed between 4 and 6 cm deep, in the vicinity of the BP. The same large and small variations in the perturbation range together with ROI contractions were investigated. In the case of a small perturbation range, good agreement was found for both full and symmetrically reduced lateral X and Y extents with maximal percentage errors being $3.6 \times 10^{-3}\%$. In the case of the large perturbation range, the algorithm resulted in a moderate percentage error of 17%. This result is not unexpected as adjoint theory provides a first order approximation to the response change and it is expected that the approximation worsens as a function of increasing perturbations (Figures 16–19).

7. Conclusion

In this paper we have developed a methodology for the approximate solution of the Linear Boltzmann Equation that takes heterogeneity in the depth direction into account. This method requires the solution to two PDEs, namely the one-dimensional FP equation and the FE equation. The one-dimensional FP equation was numerically solved through a combination of the SIPG method using quadratic energy basis functions in energy and the CN method in space. Using the 1DFP flux φ_{FP} the average depth-dependent energy $E_a(z)$ of the beam is computed. This quantity is thereafter used in the computation of the depth-dependent FE coefficients which

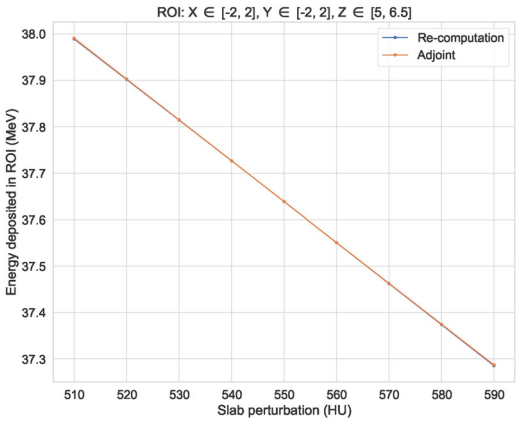

Figure 16. Adjoint versus re-computation for a small HU range with a laterally symmetric ROI and Z ∈ [5, 6.5].

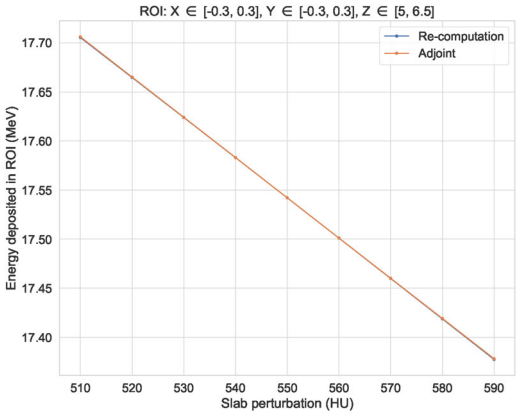

Figure 17. Adjoint versus re-computation for a large HU range with a laterally symmetric reduced ROI and Z ∈ [5, 6.5].

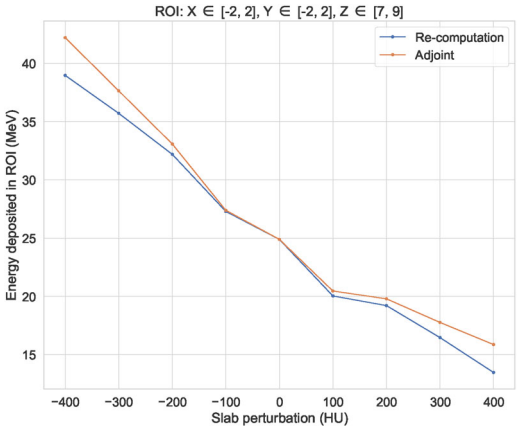

Figure 18. Adjoint versus re-computation for a large HU range with a laterally symmetric ROI and Z ∈ [7, 9].

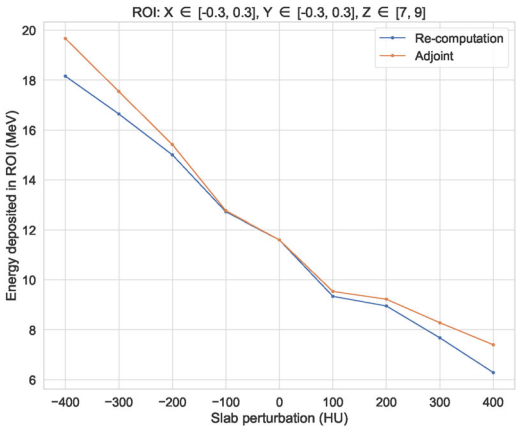

Figure 19. Adjoint versus re-computation for a large HU range with a laterally symmetric reduced ROI and Z ∈ [7, 9].

define the FE flux φ_{FE}. Using the product of these two fluxes complete knowledge of the phase-space density of protons is obtained and hereby our specific problem of charged particle transport is solved.

Using the phase-space density of protons, the response (which was defined by the deposited energy in an arbitrary ROI) was computed. Good agreement, especially in the clinically significant Bragg peak region, was obtained in a homogeneous water tank when our algorithm was benchmarked against TOPAS (taken as the reference algorithm) and Bortfeld's popular pencil beam algorithm.

Using functional analysis the adjoint system was derived and solved. The changes in the response due to slabs placed along the depth of the tank with different HU values were computed. These changes were compared against the re-computation ones with relatively good results. Adjoint theory provided (as expected) a first order approximation to the response change curve. In the case of small slab perturbations the relative difference was clinically insignificant. Even in cases of large HU perturbations, adjoint theory resulted in relatively small to moderate errors of 3% to 17%.

Future work should focus both on improving the energy/dose deposition component (forward) and on speeding up the response change computation (backward). To improve the forward component a better model is needed for the inelastic nuclear interactions between the primary protons and the irradiated tissue. Possible methods for approaching this would be a convolution-based approach or the Monte-Carlo fit method outlined by Soukup et al. (Soukup, Fippel, and Alber 2005). Moreover, it is clear that our algorithm cannot in its current state account for lateral heterogeneities and thus, a pencil beam splitting scheme is needed. A starting point for this would be the well-performing scheme proposed by Yang et. al (Yang et al. 2020). Another metric to improve on is speed. This is an area where

our algorithm performed relatively well, with an average execution time of 0.1 s for one pencil beam. Ultimately, we aim to reduce the execution time down to the ms range. This can be achieved by implementing an adaptive energy grid such that no empty energy groups are solved for. Moreover, the process of tracking many beams through the CT scan is highly parallelizable due to their independent nature.

The main drawback of the adjoint component is the long time presently needed to compute the direct contribution to the response change due to $\delta N_{\mathbb{A}}$ versus the relatively small increase in accuracy it yields. Tabulating the integrals involved should, similar to the forward component, yield significant time reductions. Moreover, the necessary data for the absorption cross section should be obtained. Another point of improvement is the inclusion of perturbations in the initial values of the FE coefficients or of the FP boundary condition coefficients. Such perturbations can be derived from machine log-files (for example the difference in Monitor Unit values for a point is related to a difference in the number of input protons and the difference in the spot positions is related to a different entry position and angle of the beam) and are the way in which our algorithm can be used for the purpose of patient-specific quality assurance.

Conflicts of interest and acknowledgements

The authors wish to acknowledge that the manuscript is partly funded by Varian, a Siemens Healthineers Company.

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. Moreover, no data was used for the research described in the article.

Credit statement

Tiberiu Burlacu: Conceptualization, methodology, software, validation, formal analysis, data curation, investigation, writing - original draft, writing - review & editing, visualization.

Danny Lathouwers: Conceptualization, methodology, software, validation, resources, writing - review & editing, supervision.

Zoltán Perkó: Conceptualization, methodology, validation, resources, writing - review & editing, supervision, project administration, funding acquisition.

Appendix: Proof of separability

Unlike in the homogeneous geometry considered by Gebäck and Asadzadeh (2012), in our more general, in-depth heterogeneous case it is not immediately straightforward that introducing the split from Equations (5) into (4) yields the Fermi-Eyges Equation (19) and the 1 D Fokker-Planck Equation (10). The difficulty is that the transport cross section $\Sigma_{tr}(z, E_a(z))$ depends on the depth-dependent mean energy $E_a(z)$. In the work of Gebäck and Asadzadeh (2012), when also accounting for energy straggling this relation is defined

implicitly via a line integral of the stopping power as $E_a(z) = E_0 - \int_0^z S(E_a(z'))dz'$. We define the average energy via Equation (26), which means that $E_a(z)$ depends on the Fokker-Planck flux $\varphi_{FP}(z,E)$, therefore the latter is present in the Υ term, namely

$$\underbrace{\frac{1}{\varphi_{FE}(\mathbf{r},\hat{\mathbf{\Omega}})}\Upsilon\Big(\varphi_{FE}(\mathbf{r},\hat{\mathbf{\Omega}}),\varphi_{FP}(z,E)\Big)}_{\lambda} + \underbrace{\frac{1}{\varphi_{FP}(z,E)}1DFP(\varphi_{FP}(z,E))}_{-\lambda} = 0. \tag{A1}$$

However, since Equation (26) defines the average energy as only depending on the energy integrated Fokker-Planck flux, the only independent variable $E_a(z)$ depends on is z. Consequently, even though $\varphi_{FP}(z,E)$ is present in the definition of $E_a(z)$, Υ does not have E dependence. The general solution of Equation A1 is therefore to equate each component to a constant $\pm\lambda \in \mathbb{R}$, as shown in Equation A1. Thus, the 1DFP component is written out as

$$\frac{\partial\varphi_{FP}}{\partial z} - \frac{\partial S(z,E)\varphi_{FP}}{\partial E} - \frac{1}{2}\frac{\partial^2 T(z,E)\varphi_{FP}}{\partial E^2} + (\Sigma_a(z,E) + \lambda)\varphi_{FP} = 0, \tag{A2}$$

while the $\Upsilon(\varphi_{FE}(\mathbf{r},\hat{\mathbf{\Omega}}),\varphi_{FP}(z,E))$ part reads

$$\frac{\partial\varphi_{FE}}{\partial z} + \Omega_x\frac{\partial\varphi_{FE}}{\partial x} + \Omega_y\frac{\partial\varphi_{FE}}{\partial y} - \Sigma_{tr}(z,E_a(z))\left(\frac{\partial^2\varphi_{FE}}{\partial\Omega_x^2} + \frac{\partial^2\varphi_{FE}}{\partial\Omega_y^2}\right) - \lambda\varphi_{FE} = 0. \tag{A3}$$

The solution of Equation A3 is found by Gebäck and Asadzadeh (Gebäck and Asadzadeh 2012) to be the usual FE solution $\varphi_{FE}(\mathbf{r},\hat{\mathbf{\Omega}})$ given in Equation (22) with a multiplicative factor, namely

$$\varphi_{FE}(\mathbf{r},\hat{\mathbf{\Omega}},\lambda \neq 0) = \varphi_{FE}(\mathbf{r},\hat{\mathbf{\Omega}})e^{\lambda z}. \tag{A4}$$

This solution also clearly follows from the structure of Equation A3, containing the $\frac{\partial\varphi_{FE}}{\partial z}$ partial derivative and $-\lambda\varphi_{FE}$ terms, in addition to other linear partial non-z derivatives. Due to the same structure, the solution of Equation A2 can similarly be written as

$$\varphi_{FP}(z,E,\lambda \neq 0) = \varphi_{FP}(z,E)e^{-\lambda z}, \tag{A5}$$

where $\varphi_{FP}(z,E)$ is the solution to usual FP Equation (10) without the additional λ term.

When multiplying $\varphi_{FE}(\mathbf{r},\hat{\mathbf{\Omega}},\lambda \neq 0)$ and $\varphi_{FP}(z,E,\lambda \neq 0)$ it becomes clear that the value of λ does not affect the solution to Equation (4) and thus simply setting $\lambda = 0$ is a valid choice, yielding the usual FE Equation (19) and FP Equation (10), even in the in-depth heterogeneous setting with exact average energy calculation. Since the boundary conditions imply the existence of a unique solution, it is also guaranteed that the splitting proposed by Gebäck and Asadzadeh (Gebäck and Asadzadeh 2012) in Equation (5) is exact in our more general, in-depth heterogeneous case too.

Funding

This work was supported by Varian Medical Systems.

ORCID

Tiberiu Burlacu http://orcid.org/0000-0001-5542-6971
Danny Lathouwers http://orcid.org/0000-0003-3810-1926
Zoltán Perkó http://orcid.org/0000-0002-0975-4226

References

Anderson, E., Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, et al. 1999. *LAPACK users' guide*. 3rd ed. Philadelphia, PA: Society for Industrial and Applied Mathematics.

Barrett, A., S. Morris, J. Dobbs, and T. Roques. 2009. *Practical radiotherapy planning*. London: CRC Press.

Börgers, C. 1999. The radiation therapy planning problem. In *Computational radiology and imaging*, edited by Willard Miller, Christoph Börgers, and Frank Natterer, Vol. 110, 1–16. New York: Springer New York.

Bortfeld, T. 1997. An analytical approximation of the Bragg curve for therapeutic proton beams. *Med. Phys.* 24 (12):2024–33. doi:10.1118/1.598116

Botas, P., J. Kim, B. Winey, and H. Paganetti. 2018. Online adaption approaches for intensity modulated proton therapy for head and neck patients based on cone beam CTs and Monte Carlo simulations. *Phys. Med. Biol.* 64 (1):015004. doi:10.1088/1361-6560/aaf30b

Brahme, A. 1975. *Simple Relations for the Penetration of High Energy Electron Beams in Matter*. Technical Report. Stockholm: Statens Straalskyddsinstitut.

Cacuci, D. 2003. *Sensitivity and uncertainty analysis, volume I: Theory*. Vol. 1. New York: Chapman and Hall/CRC.

Cordaro, M. C., and M. S. Zucker. 1972. The application of the phase space time evolution method to electron shielding. In NASA, Washington Proceedings of the National Symposium on Natural and Manmade Radiation in Space, January

Duderstadt, J. J., and L. J. Hamilton. 1991. *Nuclear reactor analysis*. New York: Wiley.

Duderstadt, J. J., and W. R. Martin. 1979. *Transport theory*. New York: Wiley.

Eyges, L. 1948. Multiple scattering with energy loss. *Phys. Rev.* 74 (10):1534–5. doi:10.1103/PhysRev.74.1534

Frank, S. J., and X. R. Zhu. 2020. *Proton therapy indications, techniques, and outcomes*. Philadelphia: Elsevier.

Gebäck, T., and M. Asadzadeh. 2012. Analytical solutions for the pencil-beam equation with energy loss and straggling. *Transp. Theory Stat. Phys.* 41 (5–6):325–36. doi:10.1080/00411450.2012.671207

Gerbershagen, A., A. Adelmann, R. Dölling, D. Meer, V. Rizzoglio, and J. Schippers. 2017. Simulations and measurements of proton beam energy spectrum after energy degradation. *J. Phys: Conf. Ser.* 874:012108. doi:10.1088/1742-6596/874/1/012108

Goldstein, H., C. P. Jr. Poole, and J. Safko. 2002. *Classical mechanics*. 3rd ed. Reading, MA: Addison-Wesley.

Gottschalk, B. 2004. *Passive beam spreading in proton radiation therapy*. Draft ed. Harvard High Energy Physics Laboratory.

Gottschalk, B. 2012. Techniques of proton radiotherapy: Transport theory. *arXiv preprint arXiv:1204.4470*.

Hillewaert, K. 2013. Development of the discontinuous galerkin method for high-resolution, large scale CFD and acoustics in industrial geometries. PhD diss., UCL - Université Catholique de Louvain.

Jette, D. 1988. Electron dose calculation using multiple-scattering theory. A. Gaussian multiple-scattering theory. *Med. Phys.* 15 (2):123–37, March–April. doi:10.1118/1.596166

Li, H., N. Sahoo, F. Poenisch, K. Suzuki, Y. Li, X. Li, X. Zhang, A. K. Lee, M. T. Gillin, and X. R. Zhu. 2013. Use of treatment log files in spot scanning proton therapy as part of patient-specific quality assurance. *Med. Phys.* 40 (2):021703. doi:10.1118/1.4773312

Lomax, A. J. 2008. Intensity modulated proton therapy and its sensitivity to treatment uncertainties 1: The potential effects of calculational uncertainties. *Phys. Med. Biol.* 53 (4):1027–42. doi:10.1088/0031-9155/53/4/014

Matter, M., L. Nenoff, G. Meier, D. C. Weber, A. J. Lomax, and F. Albertini. 2018. Alternatives to patient specific verification measurements in proton therapy: A comparative experimental study with intentional errors. *Phys. Med. Biol.* 63 (20):205014. doi:10.1088/1361-6560/aae2f4

Meier, G., R. Besson, A. Nanz, S. Safai, and A. J. Lomax. 2015. Independent dose calculations for commissioning, quality assurance and dose reconstruction of PBS proton therapy. *Phys. Med. Biol.* 60 (7):2819–36. doi:10.1088/0031-9155/60/7/2819

Meijers, A., G. Guterres Marmitt, K. Ng Wei Siang, A. van der Schaaf, A. C. Knopf, J. A. Langendijk, and S. Both. 2020. Feasibility of patient specific quality assurance for proton therapy based on independent dose calculation and predicted outcomes. *Radiother. Oncol.* 150:136–41. doi:10.1016/j.radonc.2020.06.027

Men, C., X. Jia, and S. B. Jiang. 2010. GPU-based ultra-fast direct aperture optimization for online adaptive radiation therapy. *Phys. Med. Biol.* 55 (15):4309–19. doi:10.1088/0031-9155/55/15/008

Noshad, H., and S. S. Bahador. 2012. Investigation on energy straggling of protons via Fokker–Planck equation. *Nucl. Instrum. Methods Phys. Res. Sect. B.* 288:89–93. doi:10.1016/j.nimb.2012.07.030

Paganetti, H. 2012. Range uncertainties in proton therapy and the role of Monte Carlo simulations. *Phys. Med. Biol.* 57 (11):R99–117. doi:10.1088/0031-9155/57/11/R99

Paganetti, H., ed. 2016. *Proton therapy physics*. Boca Raton, FL: CRC Press.

Paganetti, H., P. Botas, G. C. Sharp, and B. Winey. 2021. Adaptive proton therapy. *Phys. Med. Biol.* 66 (22):22TR01. doi:10.1088/1361-6560/ac344f

Pastor-Serrano, O., and Z. Perkó. 2022. Millisecond speed deep learning based proton dose calculation with Monte Carlo accuracy. *Phys. Med. Biol.* 67 (10):105006.

Perkó, Z., S. R. van der Voort, S. van de Water, C. M. H. Hartman, M. Hoogeman, and D. Lathouwers. 2016. Fast and accurate sensitivity analysis of IMPT treatment plans using polynomial chaos expansion. *Phys. Med. Biol.* 61 (12):4646–64. doi:10.1088/0031-9155/61/12/4646

Perl, J., J. Shin, J. Schumann, B. Faddegon, and H. Paganetti. 2012. TOPAS: An innovative proton Monte Carlo platform for research and clinical applications. *Med. Phys.* 39 (11): 6818–37. doi:10.1118/1.4758060

Piessens, R., E. deDoncker-Kapenga, C. Ueberhuber, and D. Kahaner. 1983. *QUADPACK: A subroutine package for automatic integration*. Berlin, Heidelberg: Springer.

Rivière, B. 2008. *Discontinuous Galerkin methods for solving elliptic and parabolic equations: Theory and implementation*. Philadelphia: Society for Industrial and Applied Mathematics.

Rojo-Santiago, J., S. J. M. Habraken, D. Lathouwers, A. Méndez Romero, Z. Perkó, and M. S. Hoogeman. 2021. Accurate assessment of a Dutch practical robustness evaluation protocol in clinical PT with pencil beam scanning for neurological tumors. *Radiother. Oncol.* 163:121–7. doi:10.1016/j.radonc.2021.07.028

Schneider, W., T. Bortfeld, and W. Schlegel. 2000. Correlation between CT numbers and tissue parameters needed for Monte Carlo simulations of clinical dose distributions. *Phys. Med. Biol.* 45 (2):459–78. doi:10.1088/0031-9155/45/2/314

Schwarz, M. 2011. Treatment planning in proton therapy. *Eur. Phys. J. Plus* 126 (7):1–10. doi:10.1140/epjp/i2011-11067-y

Soukup, M., M. Fippel, and M. Alber. 2005. A pencil beam algorithm for intensity modulated proton therapy derived from Monte Carlo simulations. *Phys. Med. Biol.* 50 (21): 5089–104. doi:10.1088/0031-9155/50/21/010

Suli, E., and D. F. Mayers. 2003. *An introduction to numerical analysis*. Cambridge: Cambridge University Press.

Sundermann, E., F. Jacobs, M. Christiaens, B. De Sutter, and I. Lemahieu. 1998. A fast algorithm to calculate the exact radiological path through a pixel or voxel space. *J. Comput. Inf. Technol.* 6.

Unkelbach, J., and H. Paganetti. 2018. Robust proton treatment planning: physical and biological optimization. *Semin. Radiat. Oncol.* 28 (2):88–96. doi:10.1016/j.semradonc.2017.11.005

van de Water, S., I. van Dam, D. R. Schaart, A. Al-Mamgani, B. J. M. Heijmen, and M. S. Hoogeman. 2016. The price of robustness; impact of worst-case optimization on organ-at-risk dose and complication probability in intensity-modulated proton therapy for oropharyngeal cancer patients. *Radiother. Oncol.: J. Eur. Soc. Ther. Radiol. Oncol.* 120 (1): 56–62. doi:10.1016/j.radonc.2016.04.038

van der Voort, S., S. van de Water, Z. Perkó, B. Heijmen, D. Lathouwers, and M. Hoogeman. 2016. Robustness recipes for minimax robust optimization in intensity modulated proton therapy for oropharyngeal cancer patients. *Int. J. Radiat. Oncol. Biol. Phys.* 95 (1):163–70. doi:10.1016/j.ijrobp.2016.02.035

Williams, E. J., and W. L. N. Bragg. 1932. The passage of α- and β- particles through matter and Born's theory of collisions. *Proc. R. Soc. London. Ser. A, Containing Pap. Math. Phys. Charact.* 135 (826):108–31.

Yang, J., P. He, H. Wang, G. Sun, H. Zheng, and J. Jia. 2020. An improved beam splitting method for intensity modulated proton therapy. *Phys. Med. Biol.* 65 (18):185015.

Zheng-Ming, L., and A. Brahme. 1993. An overview of the transport theory of charged particles. *Radiat. Phys. Chem.* 41 (4–5):673–703. doi:10.1016/0969-806X(93)90318-O