Physics in Medicine & Biology

IPEM Institute of Physics and Engineering in Medicine

PAPER • **OPEN ACCESS**

Implementation and 3D extension of dose reconstruction strategies from prompt-gamma emissions in proton therapy

To cite this article: B Foglia *et al* 2025 *Phys. Med. Biol.* **70** 235026

IOP Publishing Phys. Med. Biol. 70 (2025) 235026 https://doi.org/10.1088/1361-6560/ae1fcb

Physics in Medicine & Biology

IPEM Institute of Physics and Engineering in Medicine

OPEN ACCESS

RECEIVED
21 August 2025

REVISED
23 October 2025

ACCEPTED FOR PUBLICATION
14 November 2025

PUBLISHED
28 November 2025

PAPER

Implementation and 3D extension of dose reconstruction strategies from prompt-gamma emissions in proton therapy

B Foglia[1], **C Gianoli**[1], **E De Bernardi**[2], **T Masuda**[3], **T Du**[1], **T Bortfeld**[4], **K Parodi**[1,*] and **M Pinto**[1]

[1] Ludwig-Maximilians-Universität LMU Munich, Experimental Physics—Medical Physics, Garching b. München, DE, Germany
[2] School of Medicine and Surgery, University of Milano-Bicocca, Monza, IT, Italy
[3] Department of Accelerator and Medical Physics, National Institutes for Quantum Science and Technology (QST), Chiba, JP, Japan
[4] Massachusetts General Hospital MGH Boston, Radiation Oncology, Boston, MA, United States of America
[*] Author to whom any correspondence should be addressed.

E-mail: katia.parodi@physik.uni-muenchen.de and beatrice.foglia@physik.uni-muenchen.de

Keywords: prompt gamma, dose reconstruction, proton therapy, real-time monitoring

Supplementary material for this article is available online

Abstract

Objective. Range and dose monitoring with secondary radiation can help minimizing the issues of range uncertainties in proton cancer therapy. Prompt gammas (PGs) have been widely investigated as a promising secondary radiation for *in vivo* verification. Since it can be argued that for a proper delivery of the intended treatment the most desirable quantity to assess is the dose distribution *in vivo*, this work aims at the reconstruction of the delivered proton dose from distributions of PG radiation. *Approach.* Some techniques have already been proposed in the literature to reconstruct the dose from a distribution of detected secondary radiation, mostly positron emitters. Among them, very promising methods are the analytical deconvolution approach, the evolutionary algorithm and the maximum-likelihood expectation-maximization (MLEM) algorithm. Herein, the feasibility of the application of these approaches to PG distributions at emission stage is assessed with simulated mono- and polyenergetic proton beams, irradiating homogeneous and inhomogeneous phantoms, and a realistic case of a head and neck (H&N) tumor patient. *Main results.* The accuracy of the reconstructed dose is evaluated via comparison with the corresponding simulated ground truth dose distributions using different metrics. For the case of 1D reconstruction on phantoms, the ΔR_{80}, ΔR_{50} and ΔR_{10}, with $\Delta R_{\%}$ being the difference of the positions at the % of the dose maximum in the distal fall-off region between the simulated and reconstructed curves, are always less or of the order of 1 mm in absolute value. For 3D reconstruction on phantoms and on the H&N case, the γ(1%/1 mm) passing rate is always above or equal to 97%. *Significance.* This study demonstrates the applicability of the analytical deconvolution, the evolutionary and the MLEM algorithms to dose reconstruction from PG emissions, providing a step forward toward the final goal of real-time verification of the dose delivery for real-time adaptive particle therapy.

1. Introduction

Protons are widely used in cancer treatment because of their advantageous interaction properties compared to photons in conventional radiotherapy (Wilson 1946, Newhauser and Zhang 2015). Nevertheless, the exploitation of the full potential offered by protons in clinical practice is limited by several sources of treatment uncertainties (Knopf and Lomax 2013, Parodi 2016, Parodi and Polf 2018, Lomax 2020). Range verification and dose monitoring in proton therapy are thus of interest to fully exploit the favorable properties of proton beams.

Since protons stop inside the patient, it is not possible to detect them directly for range verification purposes. Among different proposed approaches, one promising possibility for monitoring is through secondary prompt gammas (PGs) (Pinto 2024). Range verification with PG was first proposed by Stichelbaut and Jongen (2003), then experimentally verified by Min *et al* (2006), and subsequently investigated by many groups using instrumentation able to detect different types of signatures, such as spatial distribution just as examples, Testa *et al* 2008, Richter *et al* 2016, Xie *et al* 2017, time (Golnik *et al* 2014) and energy (Hueso-Gonzalez *et al* 2018). In this work, the spatial distribution is considered: the emission of PG is correlated with the penetration path of protons in tissue, and the measurement of PG can be used to draw conclusions on the proton range.

PG emission happens in a timescale below the nanosecond. This makes PG suitable for the final purpose of real-time adaptive particle therapy applications, towards which the radiotherapy community interest is increasing. Building upon the benefits seen from online adaptive proton therapy (Albertini *et al* 2020, Paganetti *et al* 2021), real-time adaptive particle therapy aims at reducing proton therapy uncertainties not only among treatment fractions, but during a single fraction itself, on a pencil-beam delivery scale. This work is thought to represent a step forward this direction.

For range verification purposes, the expected PG signals related to the dose distributions calculated by the treatment planning program are needed, in order to compare the measurements results with the expected ones and check for mismatches. This approach can be seen as an indirect method for treatment verification. The modeling of the expected PG distributions could be performed by a full Monte-Carlo (MC) simulation. However, this way of proceeding is subject to approximation issues inherent in MC models (Verburg *et al* 2012). In fact, Verburg *et al* (2012) showed for example that the nuclear reaction models built in different MC tools predict different PG emissions. Moreover, a full MC calculation of highest accuracy is time consuming, so not suitable for the final goal of real-time verification. To overcome MC limitations, deep learning approaches were proposed (Hu *et al* 2020, Liu and Huang 2020, Zhang *et al* 2024), which nevertheless are still in an early stage of research and suffer nowadays from patient specificity and lack of available data for training. An alternative way of proceeding, which could also aid in the training of deep learning approaches, is by means of a filtering approach. The formalism was first introduced by Parodi and Bortfeld (2006) to relate the positron emission tomography (PET) activity profiles to the depth-dose profile delivered during proton irradiation. The main idea behind the filtering is the possibility to determine a PET profile from a dose distribution by performing a 1D convolution with a so-called filter function f. Validation of this approach and extension to PG came with successive works (Attanasi *et al* 2011, Frey *et al* 2014, Hofmann *et al* 2019a, Pinto *et al* 2020).

PG distributions can then give a direct information about the range of the primary protons. For this reason, the filtering can be mathematically referred to as direct problem. However, other than range, one of the main quantities of interest for cancer therapy is the deposited dose. Getting a direct information on the dose deposition from the PG emission is not as straightforward: the underlying mechanisms of these two physical phenomena are different, as they are governed by nuclear and electronic interactions, respectively. In principle, it would be possible to exploit the introduced filtering formalism inversely, setting up in this way a mathematical inverse problem, using an inverse filter function f^{-1} to obtain the actual delivered dose. However, this operation is a deconvolution problem, which is ill-posed: there are infinite possible solutions, sensitive to noise in particular.

To face this challenge, several deconvolution techniques have been proposed for different secondary emission signals. In particular, the following three approaches were identified in this work as most promising for PG-based dose reconstruction (i.e. reconstruction of dose distribution): the analytical deconvolution method (Remmele *et al* 2011), the evolutionary algorithm (Schumann *et al* 2016, Hofmann *et al* 2019b, Yao *et al* 2023) and the maximum-likelihood expectation-maximization (MLEM) algorithm (Masuda *et al* 2019, 2020), which were applied in literature to reconstruct the dose delivered by ions from secondary radiation data, mostly PET. They can all be built around the filtering formalism, which is then crucial for their implementation, and they can also potentially be integrated in a treatment planning system, as it already happened for the filtering (Pinto *et al* 2020). The analytical deconvolution and the MLEM algorithm were applied to reconstruct the dose delivered by protons from PET signals, generated through simulations and from measurements in homogeneous and heterogeneous phantom (Remmele *et al* 2011, Masuda *et al* 2019), from head and neck (H&N) phantom simulations for the case of MLEM (Masuda *et al* 2020), and from measurements in the case of a computed tomography (CT) data of a H&N patient for the deconvolution approach (Remmele *et al* 2011). The evolutionary algorithm was first introduced to predict the depth-dose profile from simulated PG distributions in water phantoms (Schumann *et al* 2016), then the same approach was also explored to reconstruct the dose profile from PET measurements generated after interactions of carbon ions (Hofmann *et al* 2019b).

More recently, a double evolutionary algorithm was used to realize a PG-based reconstruction of simulated dose in a thoracic CT considering a non-ideal Compton camera (Yao *et al* 2023).

In the present work, all the abovementioned algorithms are first implemented and applied to simulated dose deposited in homogeneous and inhomogeneous phantoms, for pristine Bragg peaks and spread-out Bragg peaks (SOBPs). Both laterally integrated 1D dose-depth profiles and 3D dose distributions were reconstructed. To investigate the 3D applicability of the dose reconstruction methods, *in silico* data were also generated on the CT data of a H&N patient. The performance of the techniques is evaluated and compared in terms of different metrics. Before presenting the obtained results in section 3, a brief recall of the established filtering approach is given in section 2, followed by the description of how the different dose reconstruction approaches considered herein work. Section 4 compares the methods and provides comments about their potential, giving an outlook on future steps.

2. Material and methods

2.1. Direct problem: filtering

This section focuses on the forward generation of depth-PG distributions from depth-dose profiles, similar to the work by Schumann *et al* (2016) and Pinto *et al* (2020). Here, the main ideas of the forward filtering framework are described, as it is an important part of the dose reconstruction workflows described in the following sections. For a more detailed description of the mathematical basics of the approach, please refer to Parodi and Bortfeld (2006) and successive works (Attanasi *et al* 2011, Frey *et al* 2014, Hofmann *et al* 2019a, Pinto *et al* 2020).

First, a proton beam irradiating a homogeneous target is taken into account. The 1D PG signal $P(z)$ along the beam propagation depth z can be obtained from the corresponding 1D dose distribution $D(z)$ by convolving (i.e. filtering) the dose profile itself with a filter function $f(z)$:

$$P(z) = D(z) * f(z) \ . \tag{1}$$

To estimate this filter function, in principle it would be possible to start from simulated data and then proceed with deconvolution methods. However, this is an ill-posed problem. The alternative is to move on analytically. The key elements of this analytical formalism are the so-called $\tilde{Q}_\nu$ functions, defined as the convolution of Gaussian $G(x)$ and powerlaw functions $P_\nu(x)$:

$$\tilde{Q}_\nu(x) = G(x) * P_\nu(x) \ . \tag{2}$$

The parameter ν is called shape and defines the peak-to-plateau ratio of the profile. Shifted and scaled versions of these $\tilde{Q}_\nu$ functions are first used to fit the simulated depth-dose and depth-PG profiles. A $\tilde{Q}_\nu$ function can then be described with four parameters: an overall normalization factor or weight w, the shape ν, the shift a as measure of the horizontal shift of the depth profiles, and the scale σ, which defines the width of the distal peak.

Then, a useful mathematical property of the $\tilde{Q}_\nu$ functions is exploited: in fact, the convolution of two $\tilde{Q}_\nu$ functions is still a $\tilde{Q}_\nu$ function, so

$$\tilde{Q}_{\nu_1}\left(\frac{x-a_1}{\sigma_1}\right) * \tilde{Q}_{\nu_2}\left(\frac{x-a_2}{\sigma_2}\right) = \frac{\left(\sigma_1^2+\sigma_2^2\right)^{\frac{\nu_1+\nu_2-1}{2}}}{\sigma_1^{\nu_1-1}\sigma_2^{\nu_2-1}} \tilde{Q}_{\nu_1+\nu_2}\left(\frac{x-a_1-a_2}{\sqrt{\sigma_1^2+\sigma_2^2}}\right) \ . \tag{3}$$

That said, if $P(z)$ and $D(z)$ are expressed as $\tilde{Q}_\nu$ functions, then $f(z)$ will be a $\tilde{Q}_\nu$ function as well. The $\tilde{Q}_\nu$ function's parameters of the filter can then be found exploiting the following relationships: if P indicates quantities related to the PG distribution, D to the dose profile and f to the filter,

$$\begin{cases} \nu_f = \nu_P - \nu_D \\ \sigma_f = \sqrt{\sigma_P^2 - \sigma_D^2} \\ a_f = a_P - a_D \\ w_f = \frac{w_P}{w_D} \cdot \frac{\sigma_D^{\nu_D-1}\sigma_f^{\nu_f-1}}{\sigma_P^{\nu_P-1}} \ . \end{cases} \tag{4}$$

These equations imply the following conditions:

$$\begin{cases} \nu_P > \nu_D \\ \sigma_P > \sigma_D \ . \end{cases} \tag{5}$$

One filter must be considered for each kind of nucleus in the target which emits PG after a nuclear interaction induced by a proton. The filters are also dependent on the energy of the primary protons. This has been previously showed by Schumann *et al* (2016), and mainly the parameters w and a are concerned. To tackle this filter energy dependence, the filter functions were here optimized for energies spanning the treatment energy window, from 110 MeV to 230 MeV. This was sufficient for the purposes of the current work, as showed in section 3.1.

The extension of the filtering to different homogeneous targets than the one considered to compute the filters and to inhomogeneous cases is straightforward (Parodi and Bortfeld 2006), and it can be performed still exploiting the filter functions computed for one homogeneous material, with additional steps. First, a water-equivalent range-conversion formalism is necessary. Protons irradiating different tissues with the same initial energy stop at different positions along the beam direction. Knowing the path length of protons with a defined initial energy in a certain material, it is possible to obtain the corresponding path length l traversed by protons in a second material taken as reference (usually water) as

$$l = \int_0^z \frac{\mathrm{SP}}{\mathrm{SP}_{\mathrm{ref}}}(z')\,\mathrm{d}z' \ , \tag{6}$$

where $\frac{\mathrm{SP}}{\mathrm{SP}_{\mathrm{ref}}}(z)$ is the relative stopping power between the traversed material and the reference one at the depth z' of the traversed medium. The filtering can then be applied in the homogeneous, reference space. Dose needs to be rescaled when moving from the target space to the reference one, exploiting the relative mass stopping power. To analytically calculate PG in actual space, one eventually has to consider the weight for each nucleus relative to the tissue composition of the original, reference homogeneous material (for which the filters were calculated). This weight can be introduced by a local factor $g_i(z)$:

$$g_i(z) = \left[\frac{w_i(z)}{w_{i,\mathrm{ref}}}\frac{\rho(z)}{\rho_{\mathrm{ref}}}\right] \ , \tag{7}$$

$w_i(z)/w_{i,\mathrm{ref}}$ is the weight fraction of the target nucleus involved in the reaction channel in the actual material at depth z relative to the reference homogeneous material, while $\rho(z)/\rho_{\mathrm{ref}}$ is the mass density of the material at depth z relative to to the reference one.

To take into account the presence of heterogeneities in all directions, and not just along the beam direction, a raytracing operation is performed. This also makes possible to analytically calculate 3D PG distributions from 3D dose distributions. By raytracing, the proton beam is divided into N narrow sub-beams, or rays, each representing a straight line in the target. The range conversion and the g_i factors are applied to each ray, which are then traced from the actual to the reference, homogeneous space. Summing up all the sub-beam doses in reference space, the laterally integrated depth-dose profile in reference space can be obtained. The laterally integrated depth-PG profiles for each nucleus of the reference target are then calculated via filtering. Eventually, it is possible to obtain the analytical 3D PG distribution from the laterally integrated depth-PG profile analytically calculated tracing every sub-beam back in the actual space by inversely applying the range conversion and g_i factors to each of them.

2.2. Inverse problem: dose reconstruction

With the described filtering approach, the PG signal at emission can be estimated from the dose distribution and indirect range verification operations can be carried out by comparing the said PG curve with a possible measured one, corrected for the detector response. This comparison can then give information on the difference between the measured range of protons and the expected one, but no knowledge would be available regarding the dose itself. In order to reconstruct the dose from a PG distribution via the filtering formalism, several alternative methods have been proposed. This section focuses on the description of the three methods, considered herein as most promising because of their relationship with the described filtering formalism. They exploit the application of a suitable deconvolution algorithm, which is an inverse problem and, after its application, noise or other artifacts in the data may be amplified, leading to a degraded reconstruction. This is why a regularization term is introduced, which uses a prior information on the dose distribution. Before describing the individual reconstruction methods, an introduction of regularization as common to all the considered algorithms is provided.

2.2.1. Regularization

A regularization term is included in all reconstruction algorithms, which implies the consideration of a prior knowledge of the signal, in order to stabilize the problem by adding some constraints to restrict the amount of the possible solutions.

When it is reasonable to assume a smooth solution, as for the case of dose profiles, the Tikhonov–Miller (TM) (Press *et al* 2002) and the total-variation (TV) regularizations (Vogel 2002) are suitable for noise amplification suppression in deconvolution problems. Both TM and TV use the gradient ∇D of the reconstructed dose signal D to find the *a priori* knowledge. The TM regularization is expressed as a L^2 norm:

$$R_{\mathrm{TM}}[D(z)] = \int |\nabla D(z)|^2 \mathrm{d}z \ , \tag{8}$$

while the TV regularization is a L^1 norm:

$$R_{\mathrm{TV}}[D(z)] = \int |\nabla D(z)| \mathrm{d}z \ . \tag{9}$$

The TM regularization has the effect of smoothing edges. The TV regularization was proposed to overcome this problem, but once applied, a staircasing effect occurs in the case of a continuous increase of signal values. That said, the choice of the regularization term depends on the features of the given signals. The TV regularization should be used in the case of sharp edges, while the TM regularization is recommended in the case of a continuous increase of signal values. For these reasons, and since its use was well motivated in the previous literature for the deconvolution approach (Remmele *et al* 2011) and the evolutionary algorithm (Hofmann *et al* 2019b) described in the following sections, the TM regularization was chosen for the implementation of the dose reconstruction algorithms explored and described in the following. TV regularization was also described for a matter of completeness.

2.2.2. Analytical deconvolution approach

The analytical deconvolution approach, initially developed by Remmele *et al* (2011) in the context of PET verification, consists of an iterative maximum *a posteriori* approach based on the minimization of the functional J, defined as:

$$J[D(z)] = S[D(z)] + \lambda R[D(z)] \ . \tag{10}$$

R is the TM regularization term, which is balanced against S by a parameter λ. A good value for λ can be found via a 'brute-force' search. S is called similarity (or fitting) term, and realizes the comparison between the expected PG signal, coming from the filtering of the dose to be reconstructed, and the simulated (potentially measured) one. The definition of S depends on the underlying noise distribution. Since the PG detection is a photon detection technique involving a low number of counts, measured signals are influenced by Poisson noise, and Gaussian noise may be associated with the detector's performance. If the Poisson noise outweighs the Gaussian one, as it can be for acquired signals, then the similarity term is better expressed as a Richardson–Lucy term:

$$S_{\mathrm{RL}}[D(z)] = \int [D(z) * f(z) - P(z) \cdot \log(D(z) * f(z))] \mathrm{d}z \ . \tag{11}$$

Within this work, high-statistics simulated data were considered, where the Gaussian noise overcomes the Poisson one. A more suitable definition of S in these conditions is accomplished with a least-squares (LSs) term:

$$S_{LS}[D(z)] = \int |D(z) * f(z) - P(z)|^2 \mathrm{d}z \ . \tag{12}$$

Remmele *et al* (2011) pointed out that the noise amplitude in the reconstructed dose reflects the noise in PET signals, and it is not possible to completely eliminate it. The same concept remains valid for PG signals. Higher noises can be smoothed with a higher regularization, which means a higher value of λ, but a perfect solution is not achievable. In this work, an attempt in trying to face this issue was made exploiting the $\tilde{Q}_\nu$ functions formalism. The dose distribution was first described using one $\tilde{Q}_\nu$ function. Then, instead of optimizing the similarity metrics bin by bin in the dose distribution as in Remmele *et al* (2011), the optimization was carried out in terms of $\tilde{Q}_\nu$ function's parameters. This process is expected to be faster, because the optimization involves only the four $\tilde{Q}_\nu$ function's parameters instead of all bins in depth. Another advantage is that noise is not reflected in the reconstruction. However, what has to be kept in mind is that a $\tilde{Q}_\nu$ function is already an approximation itself, so even in this case a perfect match is not achievable. Results related to this new approach and a comparison with the more standard, established one from Remmele *et al* (2011) are showed in section 3.2.

The analytical deconvolution approach relies on an optimization process. To implement it, the *minimize* function within the *scipy.optimize* Python package was used. The Nelder–Mead optimization algorithm was chosen. The standard deconvolution algorithm from Remmele *et al* (2011) was compared with the new strategy of deconvolution performed only in terms of $\tilde{Q}_\nu$ function parameters. For this latter algorithm, there is no need of regularization. The Nelder–Mead optimization algorithm was used also in this case. To define convergence of these two deconvolution algorithms, the tolerance value, which sets the acceptable error on both the optimized variable (i.e. $D(z)$) and the minimized cost function (i.e. $J[D(z)]$), was set at 10^{-6}.

2.2.3. Evolutionary algorithm

The evolutionary algorithm can be considered as a gradient-free optimization approach as well, as it relies on an iterative process. The algorithm begins with the generation of an initial population of N individuals, each of them representing one depth-dose profile $D_j(z)$. These initial profiles are first randomly mutated. Mutations can be of any nature. In this work, as in Hofmann *et al* (2019b), only one of the following mutation at a time was applied: a random horizontal shift in depth of integer bins in a range from -2 to 2; a random vertical scale of a factor between 0.8 and 1.2; or a local variation, which consists in adding a Gaussian distribution to a randomly chosen point, with random height (up to $\pm 10\%$ of the dose at the chosen bin) and σ within 10 sampled bins. The mutated dose individuals of the population are then convolved with the filter functions, in order to obtain a first population of predicted depth-PG profiles $P_j(z)$. Each of these jth depth-PG predictions is compared to the ground-truth one, potentially measured. The agreement of each comparison j at every iteration k is assessed by a quantity called fitness value ${f_j}^k$, defined as

$$f_j^k = \frac{1}{1 + M\left[D_j^k(z)\right] + \lambda R\left[D_j^k(z)\right]} \ . \tag{13}$$

M can be any metric used to compare two curves. In this work, the normalized root mean squared error (NRMSE) was used. R and λ are the regularization term and its weighting factor, respectively, analogous to what was introduced in section 2.2.2. The fitness value is a criterion which quantifies the quality of the reconstructed dose for that individual by comparing the depth-PG profile obtained after the filtering of that reconstructed dose and the ground-truth depth-PG profile. The higher the fitness of an individual of the population is, the more similar is the ground-truth depth-PG profile to the one calculated from filtering of that same individual. To then select which individuals will be the parents for the following iteration, a fitness selection is applied. These new parents individuals are then randomly mutated again, in order to vary their profile and reach a potential improvement in fitness. These steps are repeated until a certain termination condition is reached, which could be related to the fitness value or to the number of iterations. Eventually, the individual with the higher fitness in the population is designated as the reconstructed dose.

For the evolutionary algorithm, a population of $N = 300$ individuals was considered. Increasing the number of individuals in a population improves the fitness, but also increases the computing time linearly. As in Hofmann *et al* (2019b), the probability for the occuring mutations was set at 0.3, 0.3 and 0.4 for the horizontal shift, the scaling and the local Gaussian variation, respectively. The algorithm was stopped after 1500 iterations and, due to its stochastic nature, it was repeated 10 times, with different initial conditions introduced in the first random mutation. A single result among the 10 was eventually chosen, using the criterion of the lowest NRMSE between the laterally integrated profiles of the simulated PG and the PG calculated after filtering of the dose at the last iteration. The initial parameters were at first chosen from the literature (Hofmann *et al* 2019b) and then varied empirically (between 100–500 individuals and 500–3000 iterations) until a good compromise between computational effort and quality of the result was achieved.

2.2.4. MLEM algorithm

The MLEM algorithm is better known for its clinical use in image reconstruction. Masuda *et al* (2019) gave it a different purpose, exploiting it for dose reconstruction. The inherent filtering is here implemented with a different approach. The convolution in its common mathematical expression is not performed anymore. Instead, the equivalent matrix-vector product of the dose profile with a system matrix C is considered. If we define δ_j the dose in a certain voxel j and p_{ij} its corresponding PG yield, then p_{ij} can be calculated as

$$p_{ij} = C_{ij}\delta_j \ . \tag{14}$$

If P_i is the actual measured PG signal in the voxel i, the expected PG signal $\overline{P}_i$ coming from filtering can be computed as

$$\overline{P}_i = \sum_j C_{ij}\delta_j \ . \tag{15}$$

The construction of C is based on the generated filter function. When more tissues and inhomogeneous targets are taken into account, the range conversion and local factor application described in section 2.1 are performed in the system matrix itself (Masuda *et al* 2020).

The dose reconstruction algorithm is based on the maximization of the likelihood, defined as the probability that a certain PG distribution $\overline{P}_i$ obtained from applying the filtering to a dose distribution corresponds to the related measured signal P_i. The maximization of this quantity brings to the MLEM formula:

$$\delta_j^{k+1} = \frac{\delta_j^k}{\sum_i C_{ij}} \sum_i \frac{P_i C_{ij}}{\sum_j' C_{ij'}\delta_{j'}^k} \ . \tag{16}$$

Since this is still an inverse reconstruction problem, there is the need to introduce a regularization term, as in section 2.2.1. Masuda *et al* (2020) applied a TV superiorization to the dose signal itself at each iteration, dividing the MLEM algorithm in two steps and implementing what they referred to as EM-TV algorithm. In this work, the regularization was implemented in a different way. First, a TM regularization was applied as in the other methods, in order to make a meaningful comparison among them. For the same reason, instead of working on the signal itself, the regularization was applied to the objective function, i.e. the likelihood, by implementing a Bayesian reconstruction. To carry this out, the one-step-late approximation, introduced first by Green (1990), was used. This requires the following change in the MLEM formula:

$$\delta_j^{k+1} = \frac{\delta_j^k}{\left(1+\lambda\left[\frac{DR}{D\delta^k}\right]_j\right)\sum_i C_{ij}} \sum_i \frac{P_i C_{ij}}{\sum_j' C_{ij'}\delta_{j'}^k} \ , \tag{17}$$

where R is the regularization term, and it is derived with respect to the dose at the previous iteration k.

Before applying the MLEM, the generation of system matrices was necessary. In fact, the MLEM algorithm application foresees that each filter is incorporated in a respective system matrix, whose rows in this case were related to the PG space, and the columns to the dose space. The generation of the system matrices for 1D dose reconstruction is described in appendix A of the supplementary material. Once the system matrices were available, the MLEM algorithm ran for a total of 500 iterations, also determined empirically starting from literature as for the parameters of the evolutionary algorithm, and the selected result was the one at the iteration with lower NRMSE between the simulated PG profile and the PG calculated after filtering of the dose.

2.2.5. Consideration of lateral heterogeneities and 3D reconstruction

As for the forward filtering, each dose reconstruction algorithm can be also applied in 3D, and also when heterogeneities are present in all directions of the target, and not just along the beam direction. The complete workflow to reconstruct 3D dose distributions from 3D PG distributions works as follows:

(i) An initial guess of the dose (e.g. the treatment plan dose) is considered, and a raytracing operation is performed, to divide the proton beam into N sub-beams (or rays);
(ii) With the application of range conversion and g_i factors, the initial dose of each single ray is converted from actual to reference space;
(iii) All ray doses in reference space are summed up, and the laterally integrated depth-dose profile in reference space is obtained;
(iv) Filtering is then applied, which means that the laterally integrated depth-PG profiles for each nucleus of the reference target are analytically calculated;
(v) The total laterally integrated depth-PG profile in real space is calculated analytically after tracing every sub-beam back by inversely applying the range conversion and g_i factors to each ray;
(vi) With one of the described algorithms, the laterally integrated 1D depth-dose profiles in reference space are then reconstructed;
(vii) After the 1D reconstruction of the dose in reference space is completed, the inverse application of range conversion and tissue composition information allows to obtain the analytical 3D dose distribution in real space.

Table 1. REFMAT properties. Density = 1.54 g cm^{-3}.

Element	Fraction by mass (%)
^{1}H	1.76
^{16}C	30.8
^{14}N	25.9
^{16}O	37.7
^{31}P	1.34
^{40}Ca	2.5

It is worth noticing that, while the dose is reconstructed in reference space (as described in point vi) and converted in real space only at the end of the reconstruction (point vii), the iterative comparison of the laterally integrated depth-PG signals inherent in each algorithm happens in real space (point v). The need of moving from one space to another at each iteration for comparing the PG signals is because it is not possible, after potential detection, to distinguish which nucleus a single PG comes from. Moreover, the true relative stopping power and g_i factors, which would be needed to convert the PG signals from real to reference space, are not known. With the MLEM algorithm, the laterally integrated 1D depth-dose profiles are reconstructed directly in the actual space, since the information on the range conversion and tissue composition are inherent in the system matrix. When reconstructing with MLEM in 3D, one system matrix is considered for each ray involved in the raytracing process. The quantities necessary to perform raytracing are interpolated on a common array.

A scheme of the 3D reconstruction workflow is shown in figure B1 provided in appendix B of the supplementary material.

2.3. Voxelized phantoms simulations

The simulated dose and PG distributions in phantoms were obtained with the general purpose MC toolkit Geant4 (Agostinelli *et al* 2003), version 10.07.p03, using the physics list QGSP_BIC.

First, a rectangular phantom of $20 \times 20 \times 25$ cm^3 in x, y and z, respectively, made of an homogeneous reference material (referred to as REFMAT in this paper) was created. The properties of REFMAT are listed in table 1. The choice of this particular REFMAT is the same as in Hofmann *et al* (2019a) and Pinto *et al* (2020) and it takes into consideration the six most abundant elements in tissues, which produce a consistent amount of PG.

Then, two inhomogeneous phantoms were generated. Phantom1 has only longitudinal inhomogeneities. It is made of six slabs made of soft tissue, bone, lung tissue, bone and soft tissue. Materials were chosen from the Geant4 libraries, in particular 'G4_TISSUE_SOFT_ICRP', 'G4_BONE_CORTICAL_ICRP' and 'G4_LUNG_ICRP'. Phantom2 was inspired by Parodi *et al* (2005). It is composed of slabs disposed in a way to generate lateral heterogeneities. It is made of polyethylene (PE, 'G4_POLYETHYLENE'), polymethylmethacrylate (PMMA), muscle tissue ('G4_MUSCLE_SKELETAL_ICRP'), bone tissue and lung tissue. The geometry of Phantom1 and Phantom2 can be visualized with the results in section 3.1.

The simulations consider a proton pencil beam shot onto the target xy front face, along the z-axis. The total volume of all three targets was divided in 1 mm^3 cubic sub-volumes, building a voxelized structure. The energy deposition, consequently the dose, and the information for the emitted photons, like the coordinates at emission or the nucleus involved in the reaction, were scored for each voxel. Only photons created after hadronic interactions of protons with target nuclei were considered in the analysis. Depth-dose and PG profiles were created by means of lateral integration. Emitted PG were scored within a certain energy window only, between 1 MeV and 10 MeV, in order to consider for most common PG monitoring techniques responses (Pinto *et al* 2020 and references therein). This energy window can be modified as necessary by exploiting the look-up table formalism developed by Pinto *et al* (2020). This allows for a swift application of the approaches presented herein to the different PG cameras and energy windows considered in PG monitoring by different authors (just as examples, Richter *et al* 2016, Xie *et al* 2017, Hueso-Gonzalez *et al* 2018, Missaglia *et al* 2023).

2.4. Filter creation and scenarios for dose reconstruction in the voxelized phantoms

The filter functions, one for each target nucleus present in REFMAT (except hydrogen), were created by irradiating the REFMAT with 10^9 protons of incident energy equal to 150 MeV. The filters creation process is mentioned in section 3.1 and explained more in detail in appendix A of the supplementary

Table 2. Properties of the analyzed spots coming from the H&N plan.

Spot Number	Energy (MeV)	Level of inhomogeneities along beam path
1	106.82	High
2	112.25	Low
3	122.57	Low
4	125.06	High
5	131.11	High
6	132.30	Low

material. To tackle the energy dependence of the filters, they were optimized considering four energies in the therapeutic energy window. So, protons incident on REFMAT with energies equal to 110 MeV, 190 MeV and 230 MeV other than 150 MeV were also simulated, and the related depth-PG and dose profiles were used for energy optimization of the filters. This optimization was finally tested for 5×10^7 protons impinging on a homogeneous phantom of REFMAT with an energy of 200 MeV.

All the dose reconstruction approaches were first applied to REFMAT and to Phantom1. A first scenario foresees the reconstruction methods applied to a single pristine Bragg peak, made of a beam comprising 5×10^7 particles incident perpendicularly to the *xy* surface of the phantoms with an energy of 200 MeV and a 2D Gaussian shape with a σ of 4 mm. To check for the sensitivity of the methods to small range variations, range shifts were approximated by a change in the initial energy of the proton beams, to represent a global change. Protons incident on REFMAT with energies from 198.5 MeV to 201.5 MeV with a step of 0.5 MeV were then simulated, corresponding to expected range shifts in water of about 1 mm. The range shifts were evaluated as ΔR_{80} and ΔR_{10}, with $\Delta R_{\%}$ being the difference of the positions at the % of the dose maximum in the distal fall-off region between the expected curve at 200 MeV and the ones reconstructed from the PG distributions at the actual energies. A positive value indicates a higher range for the expected curve than the reconstructed one. The real range shifts relative to the planned energy were also obtained from simulated data and were used for comparison.

A second scenario considers the reconstruction of a SOBP, simulated following the approach developed first by Bortfeld (1997), then by Jette and Chen (2011). For the SOBP used in this work, 11 different energies (from 159.5 MeV to 180.9 MeV for the generation of the SOBP in REFMAT, from 180.7 MeV to 205 MeV for the one in the Phantom1) were used as contributions, forming a plateau as wide as the 20% of the maximum proton range and with a maximum of 2 Gy, to resemble the maximum fraction dose in a realistic clinical treatment plan. Dose was reconstructed separately for each contribution, to consider a realistic spot-by-spot delivery.

Phantom2 was exploited to extend the reconstruction methods from 1D to 3D, also tackling the presence of lateral heterogeneities, created by how the single slabs are positioned within the phantom. Additional lateral heterogeneities were also introduced simulating the phantom irradiation with different incident angles. Beams with σ from 2 to 10 mm incident on Phantom2 with energies from 110 to 200 MeV and angles of 0, 15, 30 and 45° were simulated to thoroughly test the 3D reconstruction algorithm before the extension to patient data.

2.5. In silico patient data

To extend the 3D dose reconstruction workflow to *in silico* patient data, a H&N patient treated at the Heidelberg Ion Beam Therapy Center with scanned proton beams (Pinto *et al* 2020) was considered. All the spots included in the plan were simulated with Geant4 in order to obtain simulated dose and PG 3D distributions. For the purpose of this work, the considered PG distributions only include PGs generated after interaction with nuclei of carbon, nitrogen, oxygen, phosphorus and calcium, since the filters were generated for these nuclei only. Two spots were chosen for a first testing of the algorithms, one in the layer with the highest energy (132.30 MeV) and one in the layer with the lower energy (106.82 MeV). These two spots where chosen because they cross regions with different compositions: while the first spot encounters a more homogeneous path, the second one interacts with a more heterogeneous region of the patient. Four additional spots were also chosen, two of them crossing more homogeneous paths and the remaining two irradiating more heterogeneous regions. The level of inhomogeneities was assessed from a qualitative point of view. Table 2 gathers the characteristics of the six chosen spots.

2.6. Precision evaluation

The six spots extracted from the H&N plan were also considered to assess the precision of the methods in presence of different levels of noise. Laterally integrated depth-PG profiles were scaled down from a very high statistics simulation and Poisson noise introduced. Said noise was randomly introduced and repeated 30 times for each scaled down amount of primaries (from 10^3 to 10^9). The dose reconstruction algorithms were then applied. Average $|\Delta R_{80}|$ and $\gamma(1\%/1\,\text{mm})$ index passing rate (PR) were calculated for each noise level, with the related standard deviation.

2.7. Choice of regularization factor λ

While the results for the deconvolution approach and the MLEM algorithm will always be the same given the same initial conditions, this is not the case for the evolutionary algorithm, since it has an inherent stochastic behavior introduced by random mutations of the depth-dose profiles at every iteration. This was considered when selecting the regularization factor value λ for the reconstructions. The deconvolution approaches and the MLEM algorithm ran for several values of λ, ranging from 0 to 10 in different steps, and the final regularization factor was chosen as the one giving the minimum NRMSE between the simulated and reconstructed laterally integrated 1D depth-dose profiles. The evolutionary algorithm was repeated not only considering several values of λ, but also considering different noise levels, as for the precision evaluation described in section 2.6. Eventually, the λ was chosen for the evolutionary algorithm as a suitable compromise for all the noise level cases, equal to 0.01. As mentioned, the deconvolution approach and the MLEM algorithm were also studied in different noise conditions, but there is not one single value of λ suitable for all the noise levels. Instead, a different value of λ was selected depending on the introduced noise.

2.8. Data analysis

The methods performances were compared via evaluation of different metrics to assess the match between the simulated and the reconstructed dose distributions. For the 1D reconstruction scenarios, metrics considered were the NRMSE and ΔR_{80}, ΔR_{50} and ΔR_{10}, with $\Delta R_{\%}$ as defined in section 2.3. To assess the robustness of the dose reconstruction, the dose was reconstructed 100 times with each algorithm for the case study of 1D reconstruction of a pristine Bragg peak in REFMAT, changing the initial dose guess. The mean value and the standard deviation for the $R_{\%}$ metrics were then calculated. For the SOBP case study, the plateau width percentage difference $\Delta pw(\%)$ between the reconstructed curve (rec) and the simulated one (ground truth, GT) was calculated as

$$\Delta pw(\%) = 100 \cdot \frac{pw_{\text{rec}} - pw_{\text{GT}}}{pw_{\text{GT}}}\ , \tag{18}$$

where pw is defined as the distance between the proximal and the distal positions at 80% of the dose maximum. Moreover, the weight of each reconstructed pencil beam that form the SOBP was compared with the weight initially assigned to build the SOBP itself. The percentage differences $\Delta w(\%)$ between the weights of the reconstructed and ground truth contributions were calculated in a way similar to equation (18).

For the case of 3D reconstruction, ΔR_{80}, ΔR_{50} and ΔR_{10} between the simulated and reconstructed laterally integrated depth-dose profiles were calculated, in analogy with the 1D study. Moreover, a γ-index analysis (Low *et al* 1998) was performed to compare analytical calculated distributions D_{rec} with simulated ones D_{sim}. γ PRs at different criteria (3%/3 mm, 2%/2 mm and 1%/1 mm) were obtained. The 3D relative error RE between the 3D reconstructed D_{rec} and simulated D_{sim} dose distributions was also calculated as

$$\text{RE} = 100 \cdot \frac{(D_{\text{rec}} - D_{\text{sim}})}{\max(D_{\text{sim}})}\ . \tag{19}$$

A positive RE value means that the reconstructed dose value in a certain voxel is higher than the simulated one in the same voxel.

The algorithms optimization in terms of time is out of the scope of this work. However, in order to have a first preliminary comparison in this sense, each algorithm was launched on the same machine equipped with a Intel(R) Core(TM) i9-10 940X CPU @ 3.30GHz 3.31 GHz processor and 128 GB RAM.

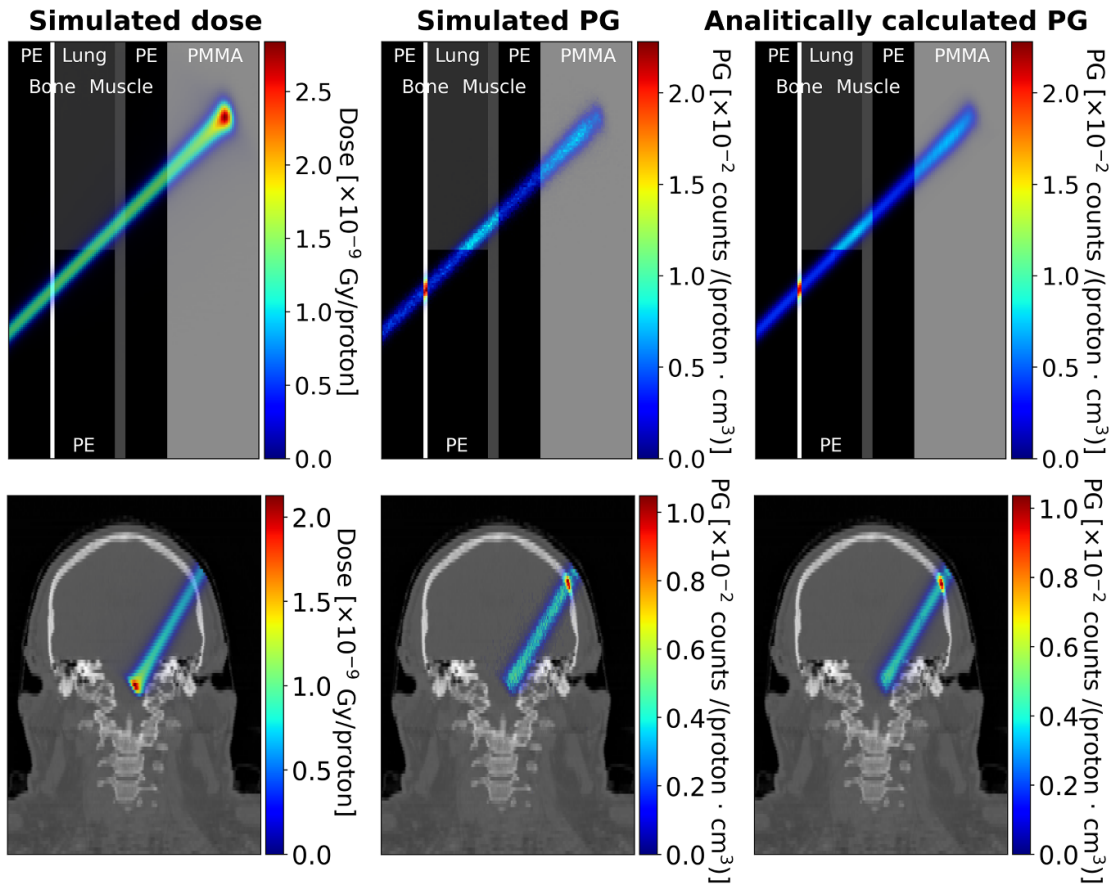

Figure 1. Upper: 3D forward filtering for Phantom2. The case with an angle of 45° from the horizontal line, an initial energy of 150 MeV and a beam sigma of 4 mm is reported. Lower: 3D forward filtering for the spot in the H&N case with an incident energy equal to 132.3 MeV. Left: simulated 3D dose distribution. Center: simulated 3D PG distribution. Right: 3D PG distribution calculated analytically with the filter functions developed in this work.

3. Results

3.1. Direct problem: filtering

The filter functions have a crucial role in the dose reconstruction methods examined, so particular care was given to their calculation. One filter function was generated for protons incident with an initial energy of 150 MeV for each element in REFMAT, excluded hydrogen. In appendix A of the supplementary material, the main results for the generation of the filters and for the application of the forward filtering approach to 1D profiles are described.

The raytracing formalism described in section 2.1 was applied to obtain predicted, analytical 3D PG distributions starting from 3D dose distributions. Results are shown on the upper part of figure 1 for Phantom2. Different angles of incidence, beam sigmas and initial energies were considered to test the feasibility of the algorithm. The same procedure was carried out for the two considered spots of the H&N patient. The 3D forward filtering for the first spot in the layer corresponding to an energy of 132.3 MeV is shown on the bottom of figure 1.

3.2. Inverse problem: dose reconstruction

3.2.1. 1D reconstruction, REFMAT

Figure 2 shows the reconstructed dose for a pristine beam of protons incident with initial energy of 200 MeV into REFMAT. Table 3 collects the different metrics used to compare the methods, described in section 2.8. As for the following tables, the number of iterations necessary to complete the reconstruction, the final regularization factor λ and an indication of the duration of the reconstruction are also reported. It is worth noticing that the R_{80}, R_{50} and R_{10} were reconstructed with an accuracy better than 1 mm in absolute value. The error comes from the 1D reconstruction robustness assessment described in section 2.8.

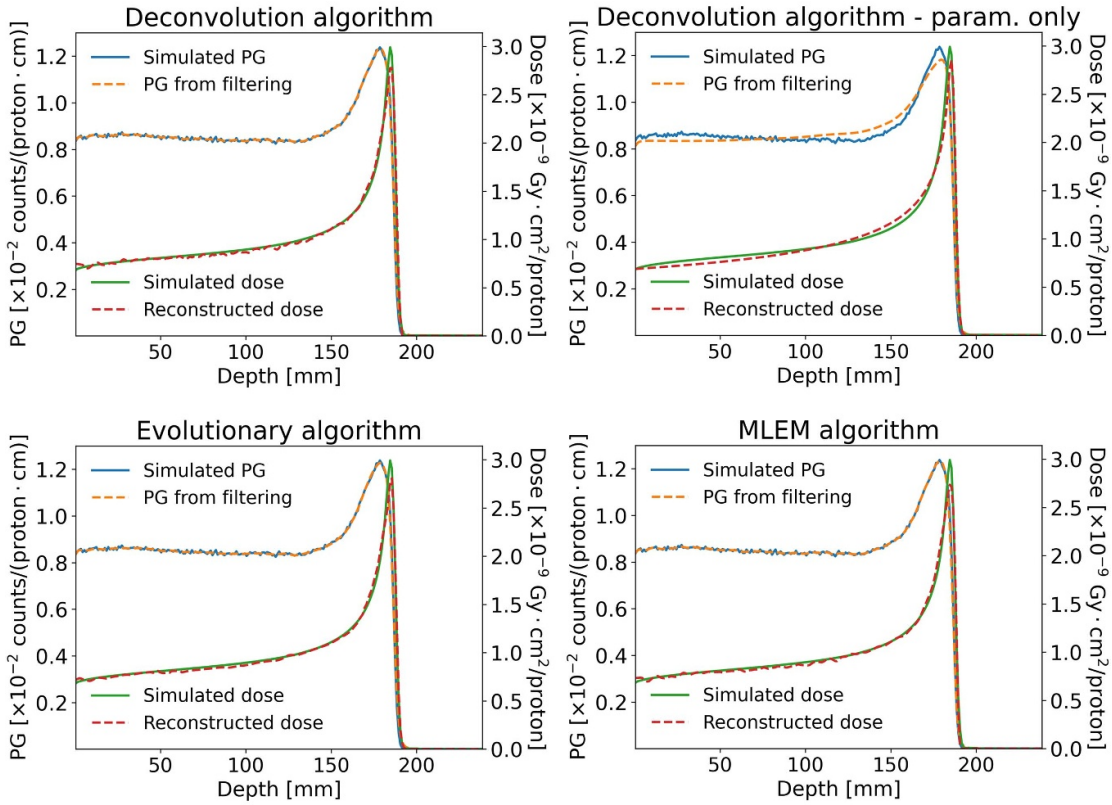

Figure 2. Pristine Bragg peak reconstruction in REFMAT, incident proton energy of 200 MeV. 5×10^7 primary protons were simulated. Upper row, left: deconvolution approach as in Remmele *et al* (2011). Upper row, right: deconvolution in terms of $\tilde{Q}_\nu$ function parameters only. Lower row, left: evolutionary algorithm. Lower row, right: MLEM.

Table 3. Metrics used for comparison between simulated dose (ground-truth) and reconstructed one with each method. Case study: pristine Bragg peak reconstruction in REFMAT, incident proton energy of 200 MeV, 5×10^7 primary protons simulated.

Metric	Deconvolution	Deconvolution (param. only)	Evolutionary	MLEM
λ	0.03	—	0.01	0.1
Iterations	112 119	311	1500	44
Time (s)	47	2	260	4
NRMSE ($\times10^{-2}$)	4.33	6.38	4.32	4.72
ΔR_{80} (mm)	-0.39 ± 0.01	-0.36 ± 0.00[a]	-0.25 ± 0.07	-0.40 ± 0.01
ΔR_{50} (mm)	-0.38 ± 0.01	-0.26 ± 0.02	-0.31 ± 0.03	-0.34 ± 0.00[a]
ΔR_{10} (mm)	-0.06 ± 0.05	-0.07 ± 0.06	-0.11 ± 0.05	0.32 ± 0.04

[a] Standard deviation lower than the shown digit.

Figure 3 shows instead the results for the reconstruction of multiple dose contributions, forming a SOBP. The values of the metrics used to compare the methods for this instance are shown in table 4. These metrics are calculated for the total SOBP curve, even if the reconstructions were performed separately for each contribution, to consider a realistic delivery spot by spot. The total number of iterations and the total duration of the reconstruction, summing over all calculated contributions, are also reported. R_{80}, R_{50} and R_{10} were calculated starting from the position of the dose maximum for the contribution with higher energy. It can be observed that the methods are able to reconstruct each contribution with its associated weight. The mean $\overline{\Delta w}$ among all the percentage differences between the weights of the reconstructed and simulated contributions are also reported in table 4.

3.2.2. Range shifts analysis

The results related to the range sensitivity analysis of the dose reconstruction algorithms are shown in figure 4. Dose was initially simulated at 200 MeV, which is the nominal energy, and then reconstructed coming from shifted PG signals simulated at energies from 198.5 MeV to 201.5 MeV in steps

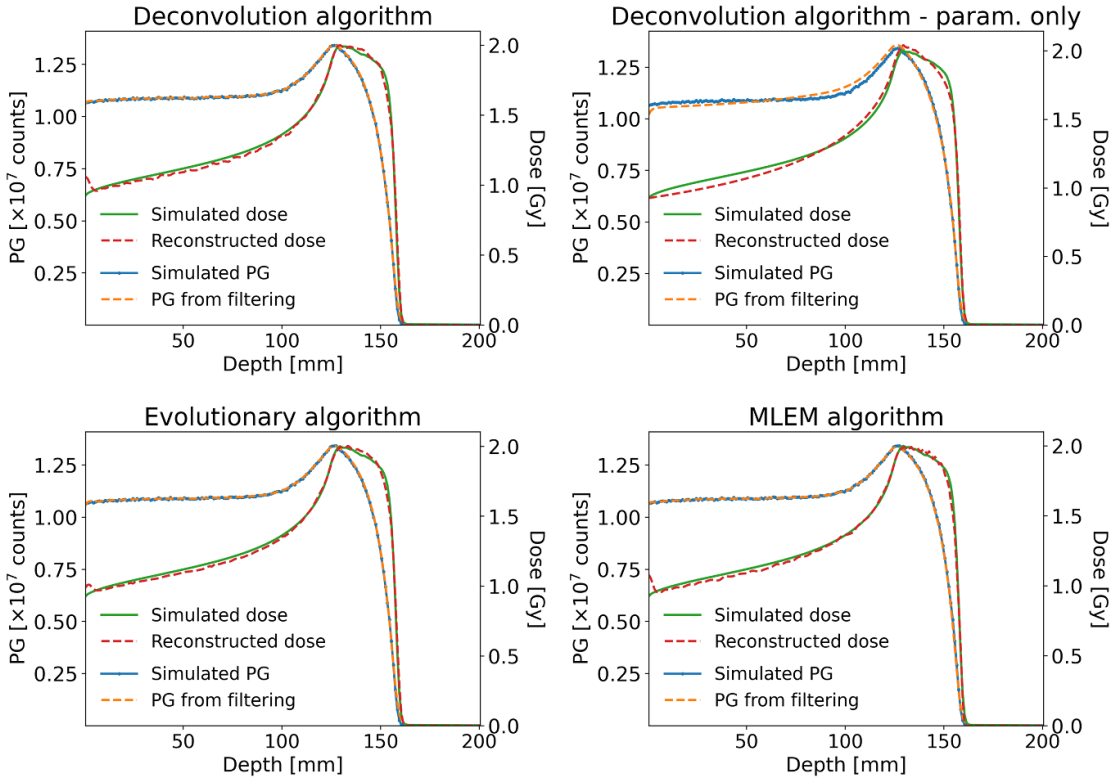

Figure 3. SOBP reconstruction in REFMAT. 5×10^7 primary protons were simulated per contribution. Upper row, left: deconvolution approach as in Remmele *et al* (2011). Upper row, right: deconvolution in terms of $\tilde{Q}_\nu$ function parameters only. Lower row, left: evolutionary algorithm. Lower row, right: MLEM.

Table 4. Metrics used for comparison between simulated dose (ground-truth) and reconstructed one with each method. Case study: SOBP reconstruction in REFMAT. 5×10^7 primary protons were simulated per contribution.

Metric	Deconvolution	Deconvolution (param. only)	Evolutionary	MLEM
Total iterations	2299 852	3260	16 500	1002
Total time (s)	821	18	2146	40
NRMSE ($\times10^{-2}$)	2.46	3.40	2.18	2.57
ΔR_{80} (mm)	-0.17 ± 0.01	-0.17 ± 0.00[a]	-0.07 ± 0.07	0.05 ± 0.01
ΔR_{50} (mm)	-0.24 ± 0.01	-0.06 ± 0.02	-0.20 ± 0.03	-0.16 ± 0.00[a]
ΔR_{10} (mm)	-0.12 ± 0.05	-0.19 ± 0.06	-0.18 ± 0.05	0.27 ± 0.04
Δpw (%)	1.5	2.1	1.0	1.5
$\overline{\Delta w}$ (%)	-3.30 ± 1.44	1.81 ± 0.58	4.16 ± 1.37	-4.86 ± 0.97

[a] Standard deviation lower than the shown digit.

of 0.5 MeV. The curves were zoomed in the distal area to better distinguish the shifts. The ΔR_{80} and ΔR_{10} between expected and unplanned energies, calculated both for the simulated and the reconstructed curves, are shown in the bottom part of figure 4. The remaining offset corresponds to the ΔR_{80} and ΔR_{10} between simulated and reconstructed dose at the nominal energy.

3.2.3. 1D reconstruction, Phantom1

Figure 5 shows the results for a pristine beam of protons incident on the heterogeneous Phantom1 with initial energy of 200 MeV, while figure 6 shows the results for the delivery of multiple contributions, forming a SOBP. In tables 5 and 6 the different metrics used to compare the algorithms performance are shown for the pristine Bragg peak and the SOBP, respectively. As for the homogeneous phantom, R_{80}, R_{50} and R_{10} in the SOBP case were calculated starting from the location of the dose maximum for the contribution with higher energy. Also, the methods are able to reconstruct each contribution with its associated weight. The mean $\overline{\Delta w}$ among all the percentage differences between the weights of the reconstructed and simulated contributions are also reported in table 6. The slight overestimation of the

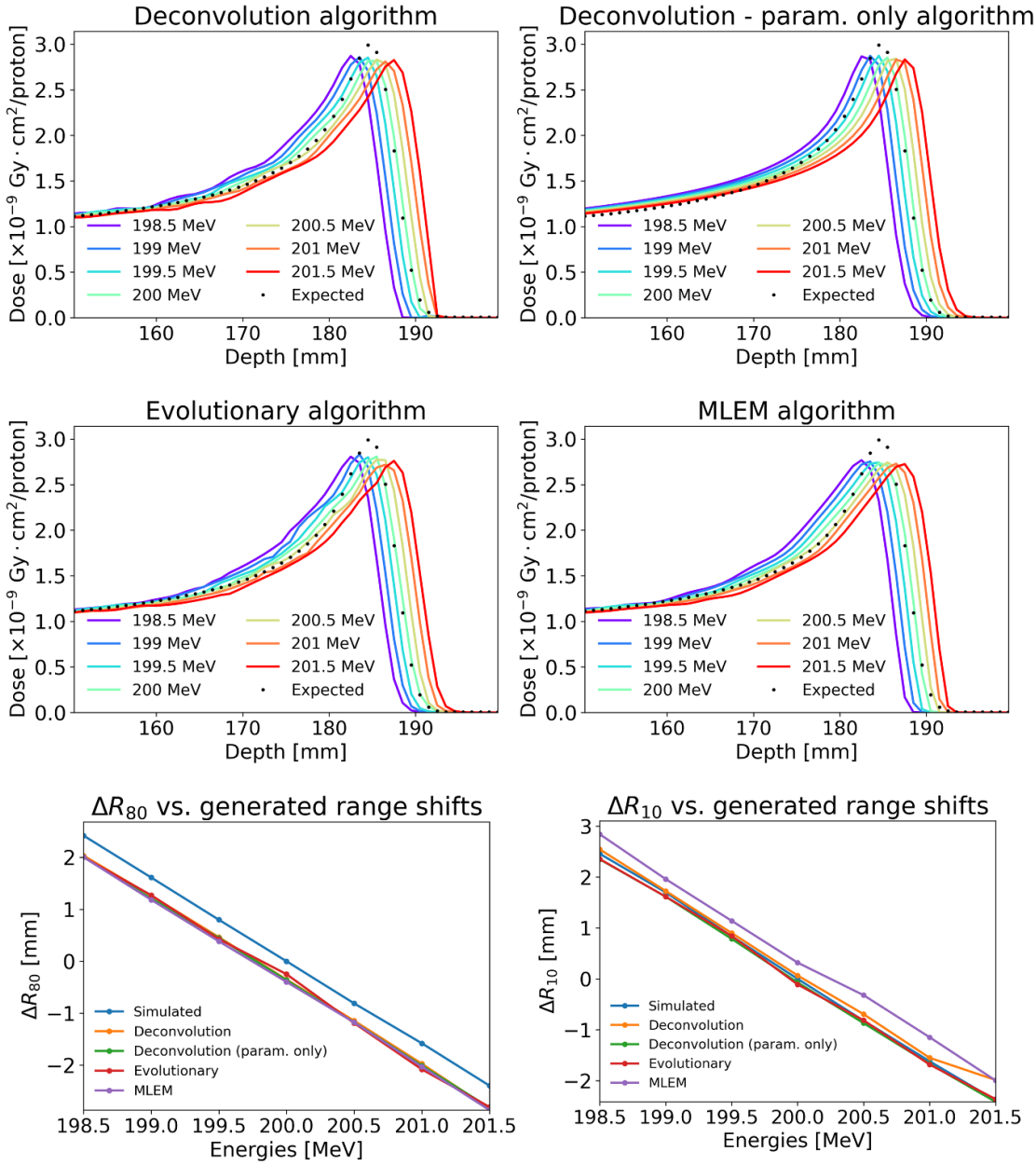

Figure 4. Range shifts analysis related to the reconstruction of a pristine Bragg peak (zoomed on the distal area) in REFMAT. Planned incident proton energy of 200 MeV. 5×10^7 primary protons were simulated per case. Upper row: deconvolution approach, implemented first as in Remmele *et al* (2011) (left) and then with the optimization of the $\tilde{Q}_\nu$ function parameters only (right). Middle row: evolutionary algorithm (left) and MLEM (right) are shown. Lower row: calculated metrics, ΔR_{80} on the left, ΔR_{10} on the right.

dose proximal to the Bragg peak for the pristine beam noticeable in figure 5 is summed over each contribution for the SOBP, which explains the final overestimation in the reconstructed plateau of the SOBP, more evident in figure 6.

3.2.4. 3D reconstruction, Phantom2

Phantom2 was used to verify the feasibility of all the dose reconstruction methods analyzed, when the 3D workflow is implemented, prior to application to clinical-like data. The case for a proton beam of σ of to 4 mm, incident angle of 45°, and initial energy of 150 MeV is reported as an example in figure 7. The values of the metrics used to compare the 3D reconstructed dose distributions with the ground-truth simulations are summarized in table 7. $\Delta R_{\%}$ metrics reach a maximum absolute value of the order of 1.6 mm, while the γ_{PR} metrics at the stricter condition of 1%/1 mm is always above 97%. For the $\Delta R_{\%}$ and γ_{PR} metrics, the error in the table is reported only for the evolutionary algorithm case. This

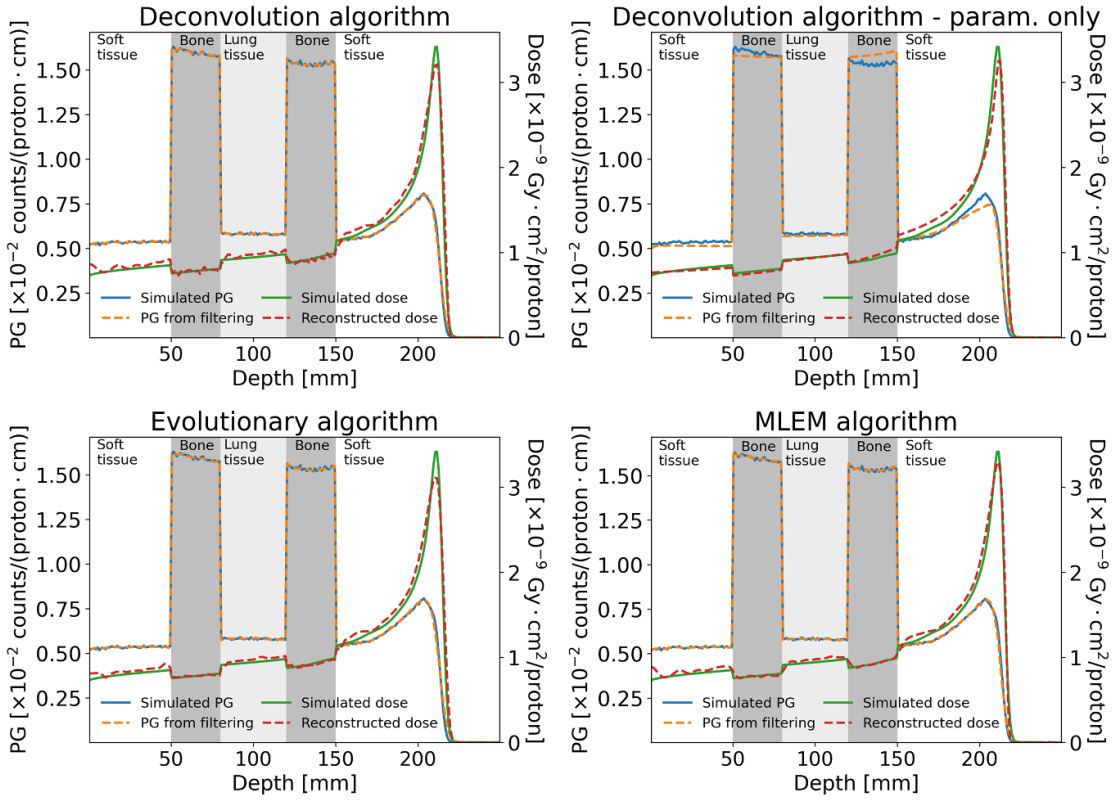

Figure 5. Pristine Bragg peak reconstruction in Phantom1, incident proton energy of 200 MeV. 5×10^7 primary protons were simulated. Upper row, left: deconvolution approach, implemented as in Remmele *et al* (2011). Upper row, right: deconvolution in terms of $\tilde{Q}_\nu$ function parameters only. Lower row, left: evolutionary algorithm. Lower row, right: MLEM.

is because only the evolutionary algorithm was repeated 30 times for Phantom2, because of what mentioned in section 2.7, and the average value and standard deviation of the calculated metric is reported in the table.

3.2.5. 3D reconstruction, H&N patient

To investigate the feasibility of the dose reconstruction algorithms toward a more relevant clinical case, two spots from a H&N plan were reconstructed. The spots differ on the level of heterogeneity of the traversed path. The results are shown in figures 8 and 9 and the metrics calculated to compare the 3D reconstructed dose distributions with the ground-truth simulations are summarized in table 8. $\Delta R_{\%}$ metrics reach a maximum absolute value of the order of 1.1 mm, while the γ_{PR} metrics at the stricter condition of 1%/1 mm is always above or of the order of 97%. Analogously to the Phantom2 case, for the $\Delta R_{\%}$ and γ_{PR} metrics, the error is reported only for the evolutionary algorithm case.

3.2.6. Consideration of different noise levels, patient case

For the H&N patient case, the explored dose reconstruction algorithms were applied to PG distributions presenting different noise levels, as described in section 2.6. The noise levels applied to the laterally integrated depth-PG profile related to each analyzed spot are defined by the considered number of primary protons before applying the Poisson random distribution, from 10^3 to 10^9. Average $|\Delta R_{80}|$ and $\gamma_{\mathrm{PR}}(1\%/1\,\mathrm{mm})$ are presented in figures 10 and 11.

4. Discussion

Three algorithms already presented in literature to reconstruct the dose profiles from PET signals and, in one case, PG signals were here applied to PG distributions: the deconvolution approach (in its initial formulation and an adapted one), the evolutionary algorithm and the MLEM algorithm. Only the evolutionary algorithm was applied to PG data so far, in the study carried out by Schumann *et al* (2016), Yao *et al* (2023). In the current work, the feasibility of the methods applied to PG emission distributions

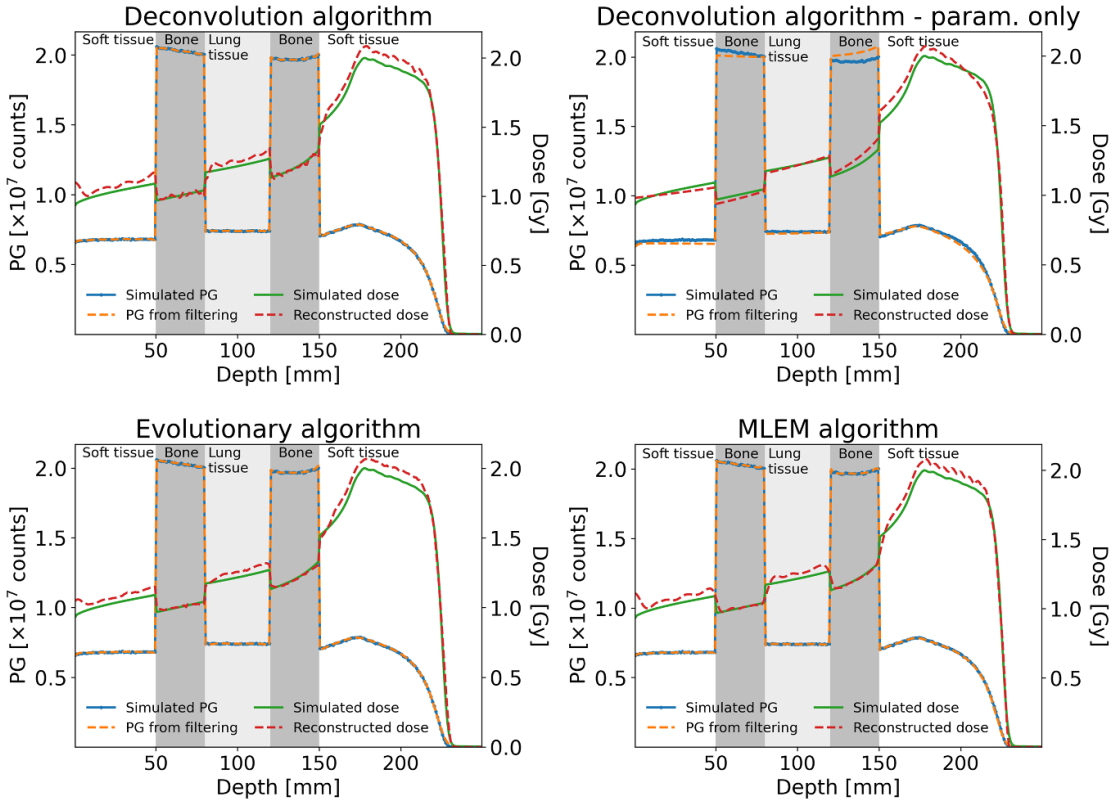

Figure 6. SOBP reconstruction in Phantom1. 5×10^7 primary protons were simulated per contribution. Upper row, right: deconvolution in terms of $\tilde{Q}_\nu$ function parameters only. Lower row, left: evolutionary algorithm. Lower row, right: MLEM.

Table 5. Metrics used for comparison between simulated dose (ground-truth) and reconstructed one with each method. Case study: pristine Bragg peak reconstruction in Phantom1, incident proton energy of 200 MeV. 5×10^7 primary protons were simulated.

Metric	Deconvolution	Deconvolution (param. only)	Evolutionary	MLEM
λ	0.02	—	0.01	0.2
Iterations	324 606	218	1500	21
Time (s)	122	2	231	4
NRMSE ($\times10^{-2}$)	6.46	7.18	6.94	5.63
ΔR_{80} (mm)	-0.58 ± 0.01	-0.70 ± 0.00[a]	-0.46 ± 0.07	-0.42 ± 0.01
ΔR_{50} (mm)	-0.74 ± 0.01	-0.56 ± 0.02	-0.76 ± 0.03	-0.36 ± 0.00[a]
ΔR_{10} (mm)	-0.25 ± 0.05	-0.26 ± 0.06	-0.89 ± 0.05	0.68 ± 0.04

[a] Standard deviation lower than the shown digit.

Table 6. Metrics used for comparison between simulated dose (ground-truth) and reconstructed one with each method. Case study: SOBP reconstruction in Phantom1. 5×10^7 primary protons were simulated per contribution.

Metric	Deconvolution	Deconvolution (param. only)	Evolutionary	MLEM
total iterations	2632 818	2850	16 500	455
total time (s)	971	22	25 951	13
NRMSE ($\times10^{-2}$)	3.92	3.48	3.73	3.87
ΔR_{80} (mm)	-0.05 ± 0.01	-0.44 ± 0.00[a]	-0.11 ± 0.07	0.19 ± 0.01
ΔR_{50} (mm)	-0.36 ± 0.01	-0.50 ± 0.02	-0.44 ± 0.03	-0.28 ± 0.00[a]
ΔR_{10} (mm)	-0.16 ± 0.05	-0.21 ± 0.06	-0.51 ± 0.05	0.50 ± 0.04
Δpw (%)	-2.84	-0.05	-0.71	-0.42
$\overline{\Delta w}$ (%)	2.22 ± 1.04	3.36 ± 1.13	-2.85 ± 2.64	3.76 ± 1.42

[a] Standard deviation lower than the shown digit.

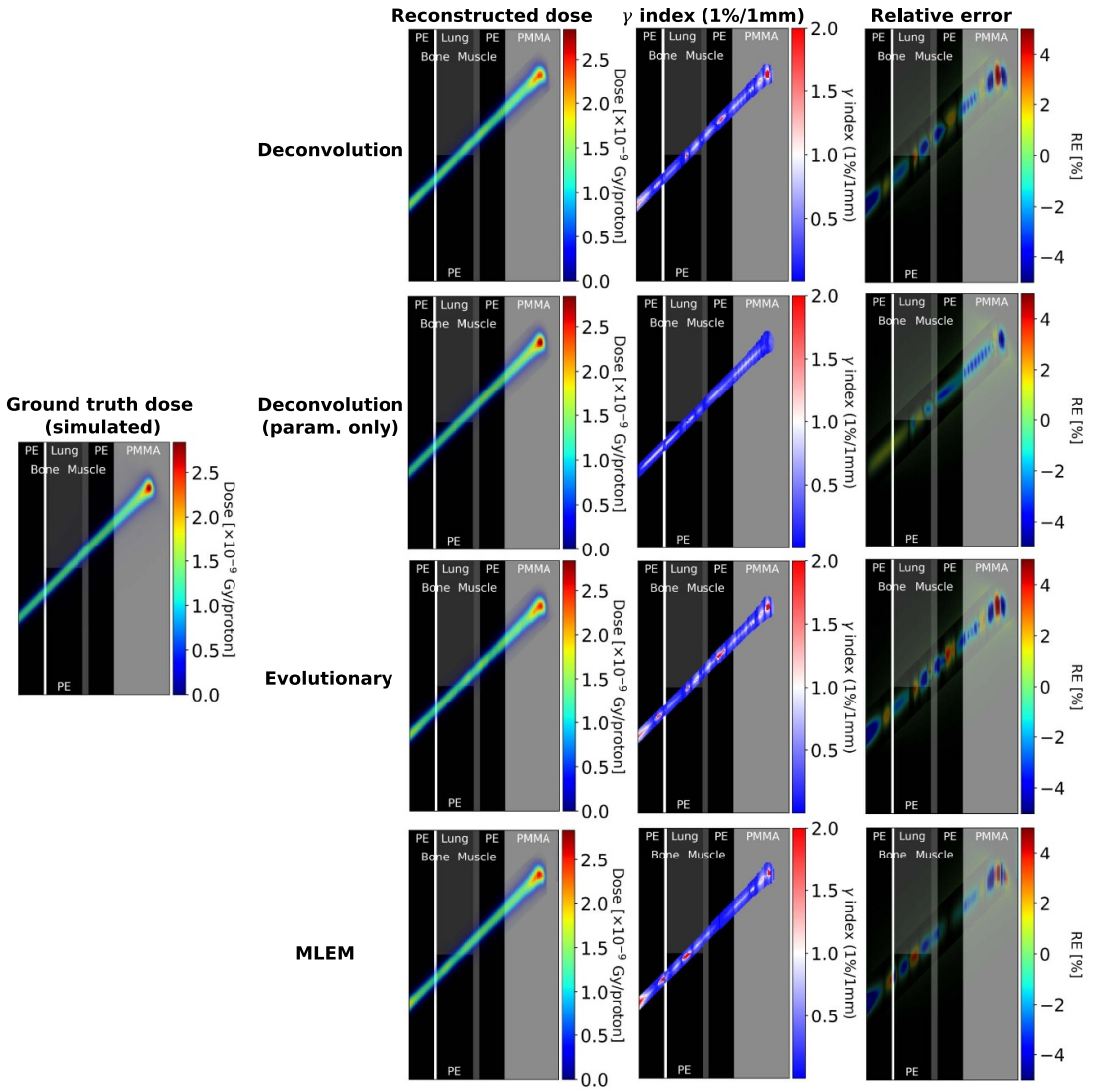

Figure 7. Results obtained after 3D reconstruction for Phantom2. A beam with initial energy of 150 MeV, 4 mm σ, irradiating the phantom with 45° angle of incidence is shown. 5×10^7 primary protons were simulated. First column from the left: simulated 3D dose distribution. Second column: reconstructed 3D dose distributions. Third column: $\gamma(1\%/1\,\text{mm})$ passing rate. Fourth column: relative error. First row: deconvolution algorithm. Second row: deconvolution algorithm in terms of $\tilde{Q}_\nu$ function parameters only. Third row: evolutionary algorithm. Fourth row: MLEM algorithm.

Table 7. Metrics used for comparison between 3D simulated dose (ground-truth) and 3D reconstructed one with each method. Case study: spot dose reconstruction in Phantom2. Protons irradiated the phantom with an initial energy of 150 MeV and an incident angle of 45°. 5×10^7 primary protons were simulated.

Metric	Deconvolution	Deconvolution (param. only)	Evolutionary	MLEM
λ	0.07	—	0.01	0.03
Iterations	351 943	251	1500	500
Time	118 min	18 s	216 min	0.6 s
ΔR_{80} (mm)	-1.33	-1.21	-1.42 ± 0.43	-0.96
ΔR_{50} (mm)	-1.55	-0.18	-0.38 ± 0.38	-0.43
ΔR_{10} (mm)	-1.50	-0.03	-1.26 ± 0.03	0.85
$\gamma_{PR}(3\%/3\,mm)$ (%)	99.61	100	99.74 ± 0.09	99.82
$\gamma_{PR}(2\%/2\,mm)$ (%)	99.42	100	99.53 ± 0.12	99.49
$\gamma_{PR}(1\%/1\,mm)$ (%)	97.54	99.71	97.24 ± 0.22	97.12

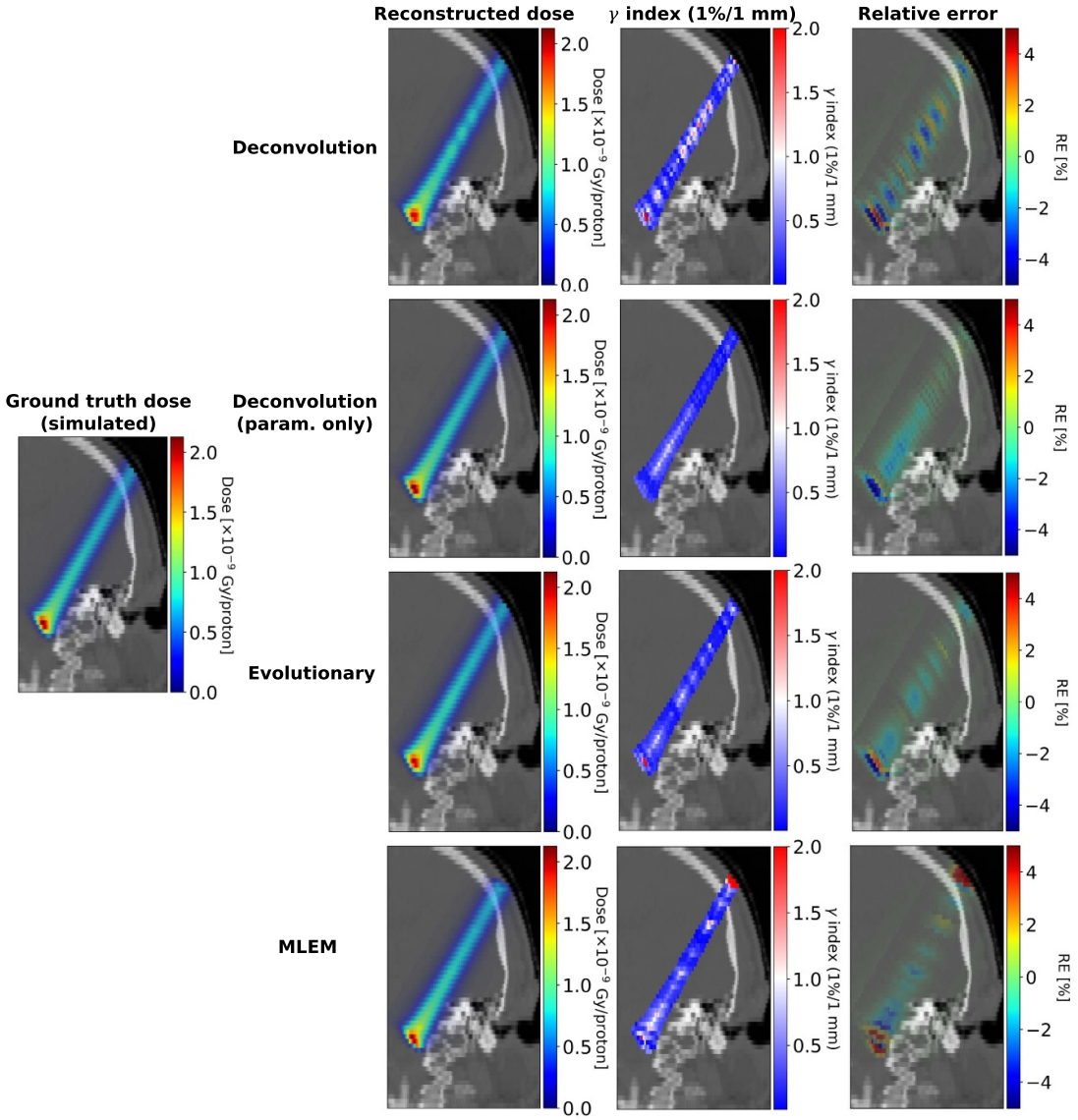

Figure 8. Results obtained after 3D reconstruction for the spot which crosses a more homogeneous path in the H&N patient. The initial energy of the beam is equal to 132.3 MeV. First column from the left: simulated 3D dose distribution. Second column: reconstructed 3D dose distributions. Third column: $\gamma(1\%/1\,\text{mm})$ passing rate. Fourth column: relative error. First row: deconvolution algorithm. Second row: deconvolution algorithm in terms of $\tilde{Q}_\nu$ function parameters only. Third row: evolutionary algorithm. Fourth row: MLEM algorithm. Plots are zoomed to provide a better view in the dose region. The complete anatomy can be visualized in the lower panel of figure 1.

was successfully verified. This feasibility was first shown exploiting simulated dose and PG 1D profiles in homogeneous and inhomogeneous phantoms for pristine Bragg peaks and SOBPs (as in sections 3.2.1 and 3.2.3). Then, 3D dose distributions were successfully reconstructed from PG distributions considering another inhomogeneous phantom and patient data (as in sections 3.2.4 and 3.2.5).

The chosen reconstruction methods are all based on the filtering approach, already verified for PG (Schumann *et al* 2016, Pinto *et al* 2020, Huang *et al* 2024). Modifications were made in this study (see appendix A of the supplementary material). Apart from fitting the PG signal with four $\tilde{Q}_\nu$ functions (as Hofmann *et al* 2019a did for filtering for carbon ion therapy) instead of three (as done so far in proton related studies), each filter corresponding to each element composing the target was used and optimized for the treatment energy window, from 110 MeV to 230 MeV. Moreover, the same filters can be used for every scenario and they can be applied to different target compositions than the one they were created for by stretching/compressing the range exploiting the stopping power relative to a reference, homogeneous material and by means of the local g_i-factors, as presented in section 2. This, for example, was not

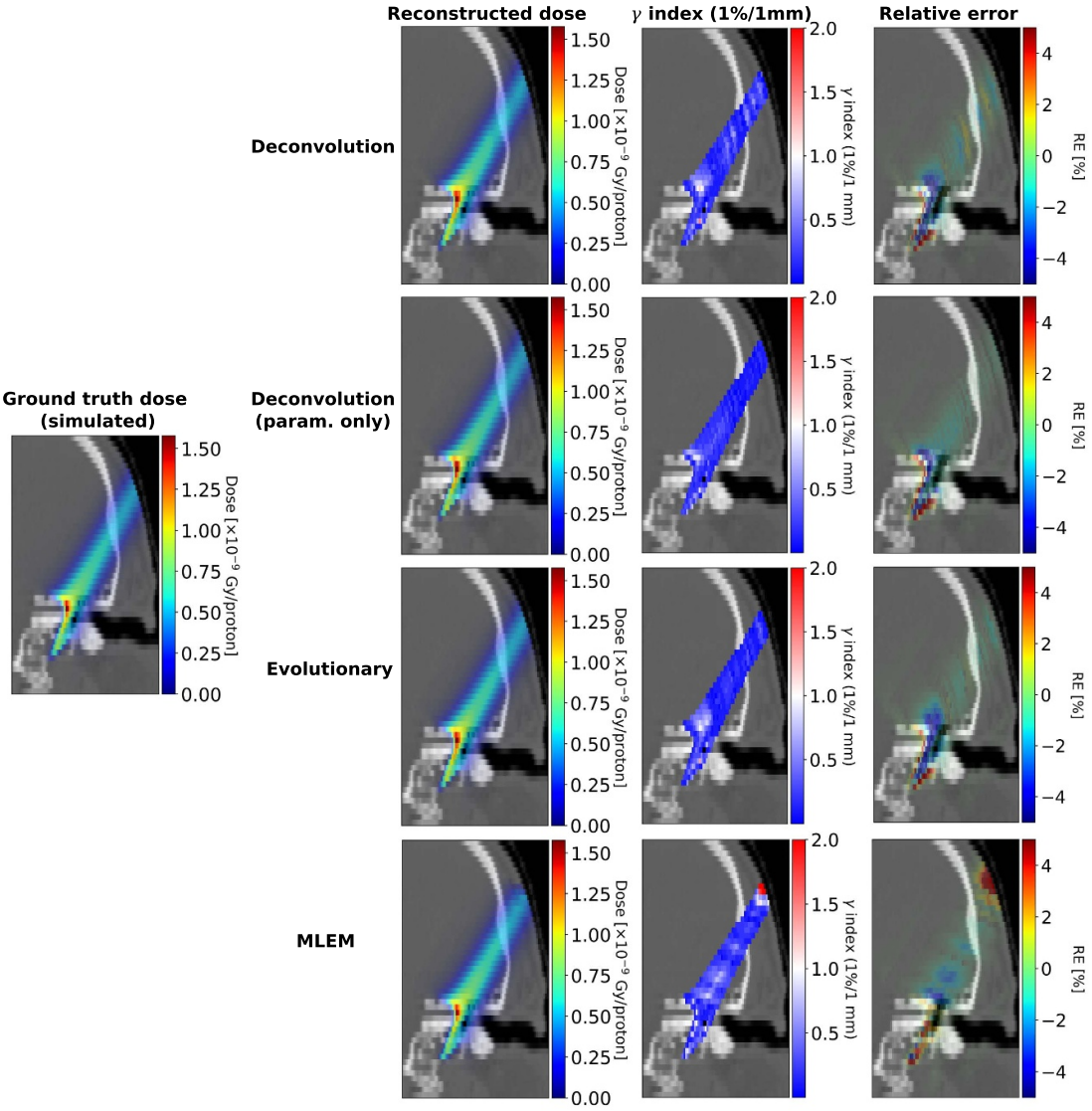

Figure 9. Results obtained after 3D reconstruction for the spot which crosses a more inhomogeneous path in the H&N patient. The initial energy of the beam is equal to 106.82 MeV. First column from the left: simulated 3D dose distribution. Second column: reconstructed 3D dose distributions. Third column: $\gamma(1\%/1\,\text{mm})$ passing rate. Fourth column: relative error. First row: deconvolution algorithm. Second row: deconvolution algorithm in terms of $\tilde{Q}_\nu$ function parameters only. Third row: evolutionary algorithm. Fourth row: MLEM algorithm. Plots are zoomed to provide a better view in the dose region. The complete anatomy can be visualized in the lower panel of figure 1.

foreseen in Schumann *et al* (2016), where the filter was only calculated for water, and it was supposed to depend on the material and on the proton energy. It also eliminates the need of a pre-calculation of the filters with the evolutionary algorithm for every specific case, as suggested by Yao *et al* (2023).

Some deviations are observed at the very beginning of the reconstructed 1D depth-dose profiles shown in sections 3.2.1 and 3.2.3. These deviations are an artifact of the reconstruction and happen at the interface that protons meet when entering the phantom. One reason for these artifacts may be the extension of the convolution beyond the available data at the edges. These deviations are more evident for the SOBP related results in figure 3 and in figure 6 than for the pristine Bragg peak reconstruction, because all the artifacts from each single reconstruction are summed up together. It was verified that, if the reconstruction is performed cutting the first few arbitrary points (about 5 mm in depth) from the PG signal used, these deviations disappear.

Another observed deviation is a slight overestimation of the dose proximal to the Bragg peak, which can be related to the fitting of laterally integrated depth-dose and depth-PG profiles with $\tilde{Q}_\nu$ functions, inherent in the forward filtering approach. In fact, PG profiles generated via Geant4 simulations present an increase in the region proximal to the Bragg peak , due to the transition of nuclear models, which is

Table 8. Metrics used for comparison between 3D simulated dose (ground-truth) and 3D reconstructed one with each method. Case study: spot dose reconstruction in the H&N patient for the two considered spots.

Metric	Deconvolution	Deconvolution (param. only)	Evolutionary	MLEM
SPOT WITH INITIAL ENERGY = 132.30 MeV (more homogeneous)				
λ	0.1	—	0.01	0.08
Iterations	261 213	258	1500	500
Time	60 min	16 s	188 min	0.7 s
ΔR_{80} (mm)	-0.18	-0.33	-0.28 ± 0.03	-0.77
ΔR_{50} (mm)	-0.06	-0.08	-0.09 ± 0.01	-0.68
ΔR_{10} (mm)	-0.13	-0.11	-0.17 ± 0.01	0.30
$\gamma_{PR}(3\%/3\,mm)$ (%)	99.86	100	99.98 ± 0.05	99.20
$\gamma_{PR}(2\%/2\,mm)$ (%)	99.74	100	99.88 ± 0.06	98.33
$\gamma_{PR}(1\%/1\,mm)$ (%)	98.61	99.98	99.70 ± 0.05	97.35
SPOT WITH INITIAL ENERGY = 106.82 MeV (more heterogeneous)				
λ	0.9	0 —	0.01	0.05
Iterations	255 366	337	1500	500
Time	75 min	20 s	367 min	0.86 s
ΔR_{80} (mm)	0.09	0.15	0.11 ± 0.01	-0.11
ΔR_{50} (mm)	0.08	0.35	0.12 ± 0.01	-0.11
ΔR_{10} (mm)	0.47	1.15	0.49 ± 0.01	0.14
$\gamma_{PR}(3\%/3\,mm)$ (%)	100	100	100 ± 0.01	99.53
$\gamma_{PR}(2\%/2\,mm)$ (%)	99.93	100	99.93 ± 0.01	98.93
$\gamma_{PR}(1\%/1\,mm)$ (%)	98.77	99.57	99.17 ± 0.12	96.94

not present in the PG profiles analytically calculated via filtering. The algorithms iteratively modify the dose profile such that the analytical PG profile matches the ground-truth, and the increase in the proximal region is then propagated in the reconstructed dose profile. The overestimation may be also due to the stretching/compressing of the actual space into reference space and vice versa, and to the application of the tissue dependent g_i factors. These operations can lead to artifacts at the interfaces, similar to the one at the entrance. The overestimation is more evident when the different pristine contributions are summed up to form a SOBP.

For all methods, except for the deconvolution in terms of $\tilde{Q}_\nu$ function parameters, where the regularization is not foreseen, one challenge in the 1D reconstruction is the flattening of the peak, present in all results where regularization is in use, and better noticeable from figures 2 and 5. Since the peak region is considered more important for range monitoring purposes than the rest of the profile, the flattening could be in principle reduced if, during the optimization, more weight is given to reconstruct the peak, and less to the rest of the curve. The expected result would be an overall noisier reconstructed curve, however better matching the peak and falloff region.

The effects of regularization and the artifact at the beginning of the 1D reconstructed curves are not observed when the deconvolution approach is applied only in terms of $\tilde{Q}_\nu$ function parameters. This method was introduced to gain in speed, and to eliminate the noise. It is indeed faster than the standard deconvolution approach. However, since the $\tilde{Q}_\nu$ function is itself an approximation of the signal, deviations from the ground truth curves are inherent, and the overall reconstructed 1D curve matches less with the simulated curve than the ones obtained with the other methods. This can be observed comparing for example the values of NRMSE from table 3 to table 6.

From the results in figure 4 on the range sensitivity analysis, it can be observed that the methods are indeed able to detect the presence of a range shift between 0.8 mm and 2.5 mm in absolute value in REFMAT, at least in the idealized conditions of our simulation study. These limit values are only due to the energies chosen to represent range shifts, and smaller shifts can be potentially detected. Sharper dose tails are noticeable for the deconvolution algorithm and, to a lesser extent, the MLEM algorithm, especially when the range shift is represented by an energy of 201.5 MeV. This behavior, deserving more investigations in future studies, comes from the changes at each iteration of the laterally integrated depth-dose profile corresponding to the nominal energy of 200 MeV protons in the attempt to match with its filtering (yielding the laterally integrated depth-PG profile) the ground-truth depth-PG profile, calculated for different 'actually delivered' energies.

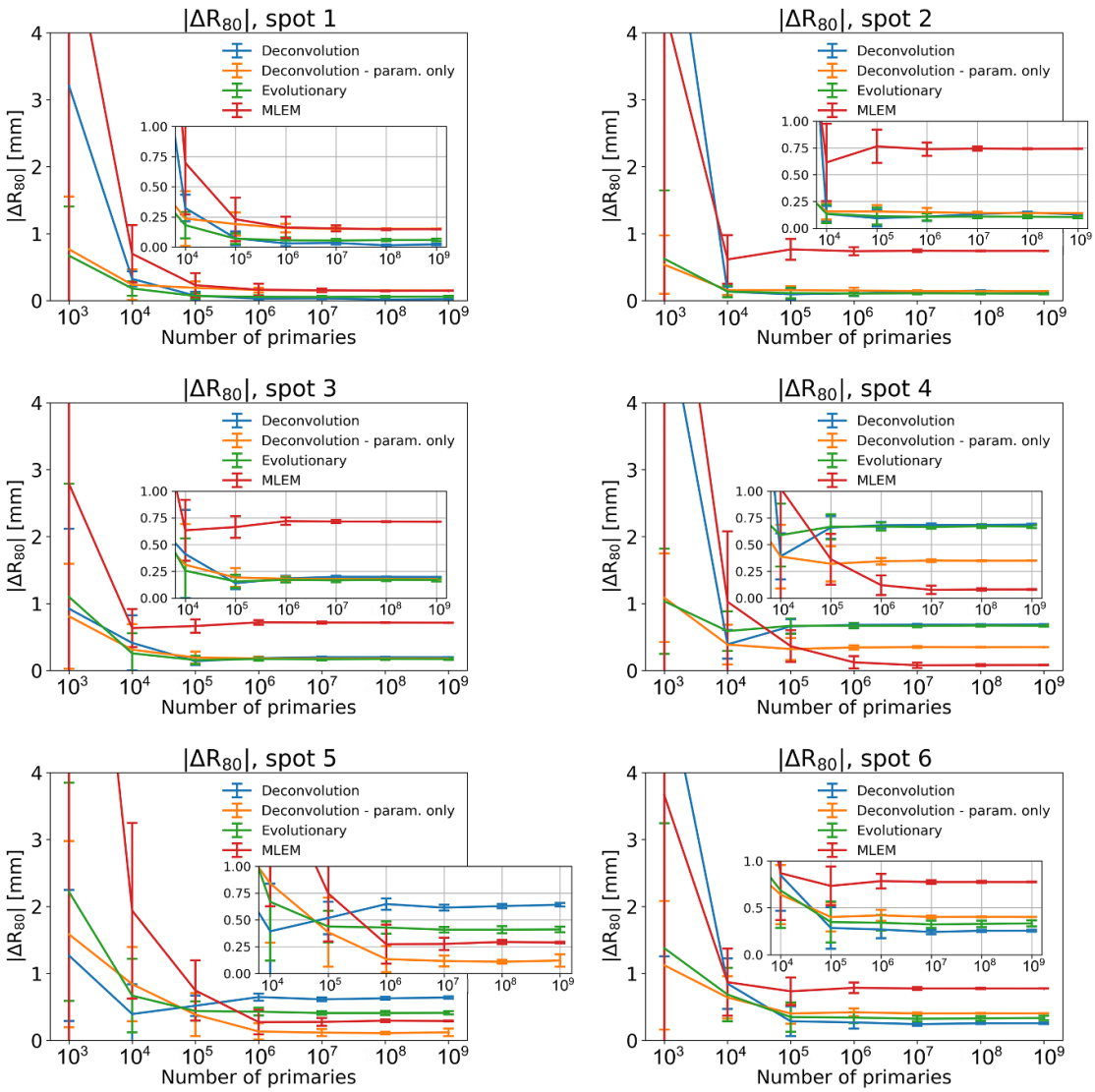

Figure 10. ΔR_{80} between simulated and reconstructed laterally integrated depth-dose profile for the H&N patient case. Six spots were analyzed. The dose was reconstructed applying different noise levels on the laterally integrated depth-PG profile, identified by the considered number of primaries on the *x*-axis. The standard deviation found for the case of deconvolution (blue) and MLEM (red) algorithms at 10^3 statistical level is, for most cases, larger than the range considered. The values found for the deconvolution algorithm were 3.2, 6.7, 0.9, 5.5, 1.3, 5.5 mm, for spot 1 through 6, respectively. The values found for the MLEM algorithm was 5.7, 4.4, 2.8, 7.7, 11.0, 3.7 mm, for spot 1 through 6, respectively. The range was limited for better visualization purposes. The inset aims at zooming in the range from 10^4 to 10^9 in the *x*-axis, and from 0 to 1.5 mm on the *y*-axis.

The above considerations remain valid also when the methods are applied to reconstruct 3D dose distributions, since the reconstruction of laterally integrated 1D depth-dose profiles is included in the 3D reconstruction workflow. However, the spread of laterally integrated profiles in 3D makes the artifacts and the peak reduction less evident.

Different metrics and criteria (i.e. number of iterations and tolerances), which were described in section 2.2, were applied to define the convergence of the algorithms. The choices were different only because the exploited parameters were sufficient for testing the feasibility of the algorithms. A comparative convergence analysis within the different algorithms, based on the same metrics and criteria, is foreseen. However, for the same reconstruction approach, the same parameters were kept when reconstructing for all case studies. For example, for the deconvolution on a bin per bin basis or only in terms of $\tilde{Q}_\nu$ function parameters, the same optimization algorithm Nelder–Mead was used, always with the same tolerance of 10^{-6}, to keep the comparison as fair as possible in this testing phase. The choice of a different optimization algorithm, such as Powell or the sequential least squares programming algorithm, combined with the fine-tuning of the chosen tolerance value, may be more suitable.

A first comparison among the methods could be performed, looking at the metrics included in this work. The duration of reconstruction is also reported in the results. From a first look, the MLEM

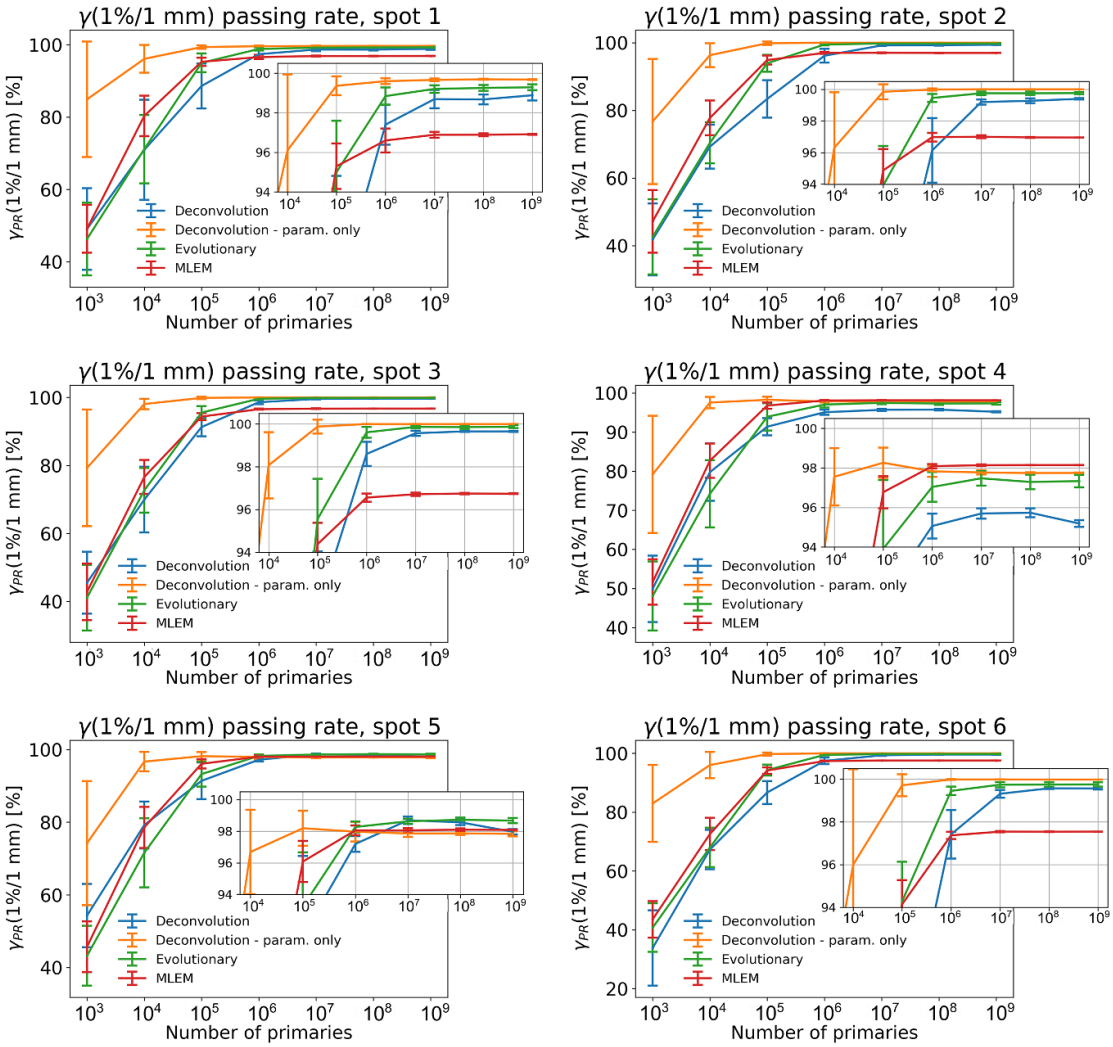

Figure 11. $\gamma_{PR}(1\%/1\,\text{mm})$ between simulated and reconstructed 3D dose distributions for the H&N patient case. Six spots were analyzed. The dose was reconstructed applying different noise levels on the laterally integrated depth-PG profile, identified by the considered number of primaries on the *x*-axis. The inset aims at zooming in the range from 10^4 to 10^9 in the *x*-axis, and from 95% to 100% on the *y*-axis.

algorithm seems to be the fastest. This is mainly due to the inherent consideration of tissue composition and range conversion within the system matrix, while for the other techniques these are considered at every iteration. The deconvolution approach applied only in terms of $\tilde{Q}_\nu$ function parameters is also fast, as expected because of the reduced number of parameters in the inherent optimization process. The MLEM algorithm could be the best candidate when compromising between accuracy and speed. Nevertheless, code optimization in terms of time is out of the scope for the present work addressing feasibility, and it will be addressed in future. All the dose reconstruction techniques could be improved in terms of speed, for example implementing lower level code optimizations, such as reducing I/O operations or moving towards parallel implementation on accelerators (GPU). That said, this work must be seen as a first step for dose reconstruction from PG distributions for mono- and polyenergetic beams (treated, in this work, as sum of monoenergetic contributions), in homogeneous and heterogeneous scenarios. It is too early to say which algorithm is the best performing. The showed results were obtained under idealized conditions considering PG at the emission point, without taking into account attenuation in the tissue or detector related factors, such as detector response, efficiency and positioning accuracy. Ongoing work, aimed to be published in a future manuscript, foresees the reconstruction from detected PG data instead of emitted ones. This would allow to study more realistic scenarios, where PG noise, which depends on proton statistics, has a major impact compared to simulations. At first, a simulated detection response is going to be implemented, following the behavior of already commercially available PG cameras. Measurements will be then used for algorithm validation. Once the methods are validated considering PG at detection level and with measurements, and once the codes go through an optimization process, a more thorough comparison among the methods could be carried out, especially

in terms of time. Duration of the algorithms permitting, the best approach could then be chosen for the final aim of real-time adaptive particle therapy applications.

On the verge of considering PG at detection level, a first assessment of the precision of the evaluated dose reconstruction algorithms is provided. Different levels of noise were considered introducing Poisson noise in a high-statistics simulation. This approach was preferred to independent simulations with different statistics, since it allows to produce different noise levels and different configurations for the same noise level in a shorter amount of time. Although PG are still scored at emission, the noise levels considered in section 3.2.6 and the related number of primaries allow for extrapolation when considering collimated detectors and their detection efficiency (lower than 10^{-3}, Krimmer *et al* (2018). The results shown in figures 10 and 11 can be potentially already exploited as a tool to estimate the result of a specific dose reconstruction algorithm when the number of primaries and the detection efficiency is known. However, such an extrapolation is only applicable to detection methods that roughly preserve the spatial PG distribution, such as multislit collimator cameras or Compton cameras. Knife-edge slit cameras would require a different analysis. A correlation between the optimal λ value and the noise level was identified. Except for the evolutionary algorithm, where the same λ value was kept to remove the sensitivity of the algorithm to random variations, it was noticed that higher values of λ were found to be optimal for high noise levels (corresponding to a lower number of protons), as expected since more regularization is needed when more noise is present.

The initial guess for the reconstruction was always the same, directly from simulations or a single spot from the treatment plan, for each case study. The initial guess was mutated only to evaluate the accuracy of the reconstruction, as described in section 2.8 for REFMAT case. In this scenario, almost no noise was present, and small deviations were introduced. This explains the low error values assigned and shown in tables from 3 to 6. It could be worthwhile to explore how the algorithms work not only in presence of different noise levels, as done in this study, but also in combination of larger variations of the initial guess, in case of a larger effect of uncertainties is present.

Other than the dose reconstruction techniques chosen for the purposes on this work, other ones exist in the literature which exploit deep learning approaches to reconstruct dose either from PET distributions, as in the work of Hu *et al* (2020), Zhang *et al* (2024), or from PG distributions, as done by Liu and Huang (2020). A comparison between analytical algorithms and machine learning methods is outside the scope of this work, but some considerations can still be made. While machine learning based frameworks could be easier to use and faster compared to kernel-based models, they are expected to be patient specific, depending on the anatomical region. Instead, the here presented analytical algorithms can be used independently on the examined site. Moreover, the same abovementioned factors such as noise and detector response should be taken into account in the networks' training, which is also subjected to the amount of data available. That said, it is still worthwhile to explore analytical reconstruction algorithms, also because physical relationships between PG emission and dose deposition are considered, and there is no need of such a large amount of data for validation as the one which would be needed for network training.

5. Conclusions

PG yields can be exploited to tackle the challenge of dose evaluation in modern proton therapy. Here, the analytical deconvolution approach, the evolutionary algorithm and the MLEM algorithm, which allow to reconstruct the dose from secondary radiation, were applied successfully to perform dose reconstruction from PG spatial distributions in 1D and 3D. Simulated proton beams in homogeneous and heterogeneous targets were considered in form of pristine Bragg peaks and SOBP and the algorithms were able to reconstruct 1D profiles as well as 3D dose distributions. With the use of these approaches it was also possible to discern range shifts even lower than 1 mm. This work is thought to be one step towards monitoring during real-time adaptive proton therapy.

Data availability statement

The data cannot be made publicly available upon publication because they are owned by a third party and the terms of use prevent public distribution. The data that support the findings of this study are available upon reasonable request from the authors.

Supplementary Data available at https://doi.org/10.1088/1361-6560/ae1fcb/data1.

Acknowledgments

This work is performed as part of the RAPTOR project, funded by the EU's Horizon 2020 MSCA, G.A. No. 955956. The authors also would like to thank the Heidelberg Ion Beam Therapy Centre for making their H&N CT data used in the former study from Pinto *et al* (2020) also available for this work.

ORCID iDs

B Foglia 0009-0003-7619-0955
C Gianoli 0000-0002-0106-7448
E De Bernardi 0000-0002-8394-0342
T Masuda 0000-0001-9648-764X
T Du 0009-0008-9539-8547
K Parodi 0000-0001-7779-6690
M Pinto 0000-0001-6835-2561

References

Agostinelli S *et al* 2003 Geant4—a simulation toolkit *Nucl. Instrum. Methods. Phys. Res. A* **506** 250–303
Albertini F, Matter M, Nenoff L, Zhang Y and Lomax A 2020 Online daily adaptive proton therapy *Brit. J. Radiol.* **93** 20190594
Attanasi F, Knopf A, Parodi K, Paganetti H, Bortfeld T, Rosso V and Guerra A D 2011 Extension and validation of an analytical model for *in vivo* PET verification of proton therapy—a phantom and clinical study *Phys. Med. Biol.* **56** 5079–98
Bortfeld T 1997 An analytical approximation of the Bragg curve for therapeutic proton beams *Med. Phys.* **24** 2024–33
Frey K, Bauer J, Unholtz D, Kurz C, Krämer M, Bortfeld T and Parodi K 2014 TPS $_{\text{pet}}$—a TPS-based approach for *in vivo* dose verification with PET in proton therapy *Phys. Med. Biol.* **59** 1–21
Golnik C *et al* 2014 Range assessment in particle therapy based on prompt γ-ray timing measurements *Phys. Med. Biol.* **59** 5399
Green P J 1990 Bayesian reconstructions from emission tomography data using a modified em algorithm *IEEE T-MI* **9** 84–93
Hofmann T, Fochi A, Parodi K and Pinto M 2019 Prediction of positron emitter distributions for range monitoring in carbon ion therapy: an analytical approach *Phys. Med. Biol.* **64** 105022
Hofmann T *et al* 2019 Dose reconstruction from PET images in carbon ion therapy: a deconvolution approach *Phys. Med. Biol.* **64** 025011
Hu Z, Li G, Zhang X, Ye K, Lu J and Peng H 2020 A machine learning framework with anatomical prior for online dose verification using positron emitters and PET in proton therapy *Phys. Med. Biol.* **65** 185003
Huang Z *et al* 2024 An experimental validation of a filtering approach for prompt gamma prediction in a research proton treatment planning system *Phys. Med. Biol.* **69** 155025
Hueso-González F, Rabe M, Ruggieri T A, Bortfeld T and Verburg J M 2018 A full-scale clinical prototype for proton range verification using prompt gamma-ray spectroscopy *Phys. Med. Biol.* **63** 185019
Jette D and Chen W 2011 Creating a spread-out Bragg peak in proton beams *Phys. Med. Biol.* **56** N131–8
Knopf A C and Lomax A 2013 *In vivo* proton range verification: a review *Phys. Med. Biol.* **58** R131–60
Krimmer J *et al* 2018 Prompt-gamma monitoring in hadrontherapy: a review *Nucl. Instrum. Methods. Phys. Res. A* **878** 58–73
Liu C and Huang H 2020 A deep learning approach for converting prompt gamma images to proton dose distributions: a Monte Carlo simulation study *Phys. Med.* **69** 110–9
Lomax A J 2020 Myths and realities of range uncertainty *Brit. J. Radiol.* **93** 20190582
Low D A, Harms W B, Mutic S and Purdy J A 1998 A technique for the quantitative evaluation of dose distributions *Med. Phys.* **25** 656–61
Masuda T, Nishio T, Kataoka J, Arimoto M, Sano A and Karasawa K 2019 ML-EM algorithm for dose estimation using PET in proton therapy *Phys. Med. Biol.* **64** 175011
Masuda T, Nishio T, Sano A and Karasawa K 2020 Extension of the ML-EM algorithm for dose estimation using PET in proton therapy: application to an inhomogeneous target *Phys. Med. Biol.* **65** 185001
Min C H, Kim C H, Youn M and Kim J 2006 Prompt gamma measurements for locating the dose falloff region in the proton therapy *Appl. Phys. Lett.* **89** 183517
Missaglia A, Bourkadi-Idrissi A, Casamichiela F, Mazzucconi D, Carminati M, Agosteo S and Fiorini C 2023 Prompt-gamma fall-off estimation with C-ion irradiation at clinical energies, using a knife-edge slit camera: a Monte Carlo study *Phys. Med.* **107** 102554
Newhauser W D and Zhang R 2015 The physics of proton therapy *Phys. Med. Biol.* **60** R155–209
Paganetti H, Botas P, Sharp G C and Winey B 2021 Adaptive proton therapy *Phys. Med. Biol.* **66** 22TR01
Parodi K 2016 On- and off-line monitoring of ion beam treatment *Nucl. Instrum. Methods. Phys. Res. A* **809** 113–9
Parodi K and Bortfeld T 2006 A filtering approach based on Gaussian-powerlaw convolutions for local PET verification of proton radiotherapy *Phys. Med. Biol.* **51** 1991–2009
Parodi K and Polf J C 2018 *In vivo* range verification in particle therapy *Med. Phys.* **45** e1036–50
Parodi K, Ponisch F and Enghardt W 2005 Experimental study on the feasibility of in-beam PET for accurate monitoring of proton therapy *IEEE Trans. Nucl. Sci.* **52** 778–86
Pinto M 2024 Prompt-gamma imaging in particle therapy *Eur. Phys. J. Plus* **139** 884
Pinto M, Kröniger K, Bauer J, Nilsson R, Traneus E and Parodi K 2020 A filtering approach for PET and PG predictions in a proton treatment planning system *Phys. Med. Biol.* **65** 095014
Press W H, Vetterling W T, Teukolsky S A and Flannery B P 2002 *Numerical Recipes in C++: The Art of Scientific Computing* (Cambridge University Press)
Remmele S, Hesser J, Paganetti H and Bortfeld T 2011 A deconvolution approach for PET-based dose reconstruction in proton radiotherapy *Phys. Med. Biol.* **56** 7601–19
Richter C *et al* 2016 First clinical application of a prompt gamma based *in vivo* proton range verification system *Rad. Onc.* **118** 232–7

Schumann A, Priegnitz M, Schoene S, Enghardt W, Rohling H and Fiedler F 2016 From prompt gamma distribution to dose: a novel approach combining an evolutionary algorithm and filtering based on Gaussian-powerlaw convolutions *Phys. Med. Biol.* **61** 6919–34

Stichelbaut F and Jongen Y 2003 Verification of the proton beam position in the patient by the detection of prompt gamma-rays emission *Talk at the 39th PTCOG Meeting* (unpublished)

Testa E, Bajard M, Chevallier M, Dauvergne D, Le Foulher F, Freud N, Létang J M, Poizat J C, Ray C and Testa M 2008 Monitoring the Bragg peak location of 73 MeV u^{-1} carbon ions by means of prompt γ-ray measurements *Appl. Phys. Lett.* **93** 093506

Verburg J, Shih H and Seco J 2012 Simulation of prompt gamma-ray emission during proton radiotherapy *Phys. Med. Biol.* **57** 5459–72

Vogel C R 2002 *Computational Methods for Inverse Problems* (Society for Industrial and Applied Mathematics) (https://doi.org/10.1137/1.9780898717570)

Wilson R R 1946 Radiological use of fast protons *Radiology* **47** 487–91

Xie Y *et al* 2017 Prompt gamma imaging for in vivo range verification of pencil beam scanning proton therapy *Int. J. Radiat. Oncol. Biol. Phys.* **99** 210–8

Yao Z, Xiao Y and Zhao J 2023 Dose reconstruction with Compton camera during proton therapy via subset-driven origin ensemble and double evolutionary algorithm *Nucl. Sci. Tech.* **34** 59

Zhang R *et al* 2024 Proton spot dose estimation based on positron activity distributions with neural network *Med. Phys.* **51** 7226–39