



Original paper

Prompt-gamma emission in GEANT4 revisited and confronted with experiment

Aleksandra Wrońska^{a,*}, Jonas Kasper^{b,*}, Arshiya Anees Ahmed^a, Achim Andres^b, Piotr Bednarczyk^c, Grzegorz Gazdowicz^a, Katrin Herweg^b, Ronja Hetzel^a, Adam Konefał^d, Paweł Kulesza^{c,1}, Andrzej Magiera^a, Katarzyna Rusiecka^b, Damian Stachura^b, Achim Stahl^a, Mirosław Ziębliński^c

^a Marian Smoluchowski Institute of Physics, Jagiellonian University, Kraków, Poland

^b Physics Institute 3B, RWTH Aachen University, Aachen, Germany

^c Institute of Nuclear Physics Polish Academy of Sciences, Kraków, Poland

^d Institute of Physics, University of Silesia in Katowice, Katowice, Poland



ARTICLE INFO

Keywords:

Proton therapy
Prompt-gamma imaging
Range verification
Monte-Carlo simulations

ABSTRACT

Purpose: The field of online monitoring of the beam range is one of the most researched topics in proton therapy over the last decade. The development of detectors that can be used for beam range verification under clinical conditions is a challenging task. One promising possible solution are modalities that record prompt-gamma radiation produced by the interactions of the proton beam with the target tissue. A good understanding of the energy spectra of the prompt gammas and the yields in certain energy regions is crucial for a successful design of a prompt-gamma detector. Monte-Carlo simulations are an important tool in development and testing of detector concepts, thus the proper modelling of the prompt-gamma emission in those simulations are of vital importance. In this paper, we confront a number of GEANT4 simulations of prompt-gamma emission, performed with different versions of the package and different physics lists, with experimental data obtained from a phantom irradiation with proton beams of four different energies in the range 70–230 MeV.

Methods: The comparison is made on different levels: features of the prompt-gamma energy spectrum, gamma emission depth profiles for discrete transitions and the width of the distal fall-off in those profiles.

Results: The best agreement between the measurements and the simulations is found for the GEANT4 version 10.4.2 and the reference physics list QGSP_BIC_HP.

Conclusions: Modifications to prompt-gamma emission modelling in higher versions of the software increase the discrepancy between the simulation results and the experimental data.

1. Introduction

The last decade has brought a vivid development of methods for online monitoring in proton therapy [1]. Such tools would enable to fully exploit the advantages offered by proton beams, at the same time minimizing the side-effects for patients [2]. Among the methods there are two groups: those allowing to verify the range of proton beams by providing one-dimensional information, and those providing 2d or even 3d information about the gamma vertex distribution, which can be translated to a deposited dose distribution [3]. In the first group of methods the most spectacular results were obtained using prompt-

gamma rays in clinical tests with a knife-edge shaped slit camera [4] and pre-clinical measurements with a spectroscopic setup [5]. Among the methods providing spatial distribution of deposited dose, very good results were obtained using a dedicated dual-head positron emission tomograph INSIDE [6]. Promising results of 3d imaging have also been obtained using methods involving prompt-gamma detection, e.g. in Ref. [7].

Monte-Carlo simulations play a crucial role in the design and optimization phase of detection setups, including those dedicated for prompt-gamma imaging (PGI) [8–12]. Not only the detector response is modelled this way, but also the characteristics of gamma quanta induced

* Corresponding author.

E-mail addresses: aleksandra.wronska@uj.edu.pl (A. Wrońska), jonas.kasper@rwth-aachen.de (J. Kasper).

¹ presently at GSI, Darmstadt

in the irradiated object by a proton beam, which depend on the description of the underlying processes in a given implementation of a Monte-Carlo engine. An example of a thorough simulation study of PG emission in proton, ^4He and carbon-ion therapy can be found in Ref. [13]. Other researchers have gone a step further and confronted the simulated PG emission with experimental data, though the data base of the latter is rather scarce. In [14] the authors compare Monte-Carlo codes GEANT4 and MCNP6 [15], and nuclear reaction codes TALYS [16] and Empire [17], reporting discrepancies up to a factor of 2 in emission of specific gamma lines in the simulation results and concluding that the nuclear reaction codes reproduce experimental data more consistently. Also the FLUKA simulation tool has been used to describe the PG emission and good agreement was found both in the spectrum shape, and in the energy-integrated PG depth profile [18]. Nevertheless, GEANT4 remains the most popular simulation framework used for that purpose, with its variety of physics lists, *i.e.* collections of processes and models considered during particle tracking, and a vivid development resulting in multiple versions [19,20].

Monte-Carlo simulations are used not only in the R&D phase, but in some approaches also in the application phase for range shift detection. There, the registered prompt-gamma distribution is compared with the simulated one and any offset between those two – if found – is interpreted as a range shift [7,21–23]. This method is sensitive to tissue elemental composition as well as cross sections of processes leading to the emission of gamma quanta. Therefore a correct description of such processes, resulting in prompt gamma emission yields matching the experimental ones, is a *conditio sine qua non* for the comparative methods.

Several published works focused on comparison of the GEANT4-simulated total, energy-integrated gamma yields in proton and carbon-ion therapy [24–26] with experimental results. In Ref. [24] the authors conclude that with the 9.4 version of GEANT4 for carbon-ion beams the best description of the integrated prompt gamma yields are found for the QMD model (Quantum Molecular Dynamics). Later, the same group, using GEANT4 version 10.01.p02 showed that a customized QMD reproduces well also the experimental data for proton therapy, while the BIC model (Binary Ion Cascade) overshoots the integral gamma yield by 40%. However, Vanstalle et al. performed also a spectroscopic analysis for carbon-ion therapy simulation with GEANT4 version 10.01. This study showed that QMD does not reproduce the features of the energy spectra of the simulated prompt-gamma radiation, in particular no discrete peaks are visible [26]. The authors point at the INCL++ (Liege Intranuclear Cascade) as the one which outperforms QMD or BIC in this case. Details of the GEANT4 physics models can be found in [27].

A thorough simulation study and a comparison of cross sections for individual discrete lines was presented by Verburg et al. [14], including the transitions most prominent at the end of the proton range resulting in emission of 4.44 MeV and 6.13 MeV gammas. The simulations were performed with GEANT4 9.5. The authors observed, that the standard settings strongly underestimated the second process. To fix this the authors applied the Fermi break-up model to the complete oxygen reaction and enforced per-event energy conservation. After these corrections, a reasonable agreement with experimental data was obtained for proton energies below 20 MeV. For higher energies the obtained cross sections were much too small. A similar feature was observed for the 4.44 MeV transition, where the discrepancies emerged at 10 MeV and 30 MeV respectively for the $^{12}\text{C}(p, p\gamma_{4.44 \text{ MeV}})^{12}\text{C}$ and $^{16}\text{O}(p, X\gamma_{4.44 \text{ MeV}})^{12}\text{C}$ reactions.

In this work we present results of simulations of prompt-gamma emission from phantoms irradiated with a proton beam, performed using several versions of GEANT4 and different physics lists. The results are confronted with the experimental data obtained in a series of measurements performed in the Heidelberger Ionenstrahl-Therapiezentrum (HIT) and the Cyclotron Centre Bronowice (CCB) in Kraków.

2. Materials and methods

2.1. Experiment

The goal of the performed measurements was to study the depth profiles of prompt-gamma emission from a phantom when irradiated with a proton beam of different energies spanning the full range used in proton therapy, *i.e.* 70–230 MeV. The detector gain was adjusted such as to register the gamma quanta of the energies up to 7 MeV, since the range 2–7 MeV is considered favourable for PGI. The experimental procedure has been described in detail in [28], so here we remind it only briefly. The heart of the setup, depicted in Fig. 1, was a HPGe detector oriented perpendicular to the beam axis at a distance of 47.2 cm, focused on a thin slice of material. The detector was equipped with an active Anti-Compton shield (ACS) and was efficiency-calibrated. The slice was a part of a multi-part target (phantom). The dimensions of its components were chosen such as to ensure a corridor of 50 mm for the beam. The peculiar construction of the phantom, with movable wedges and a 2-mm thin target slice, all made of poly(methyl methacrylate) PMMA (density 1.19 g/cm^3), allowed a geometrical selection of gamma quanta stemming from a certain depth range without using collimators. Extension blocks were used at higher beam energies to shift the range of scanned depths such as to cover the expected Bragg peak position. Moreover, additional measurements under the same conditions, but with the thin slice removed, allowed to describe and subsequently subtract the background stemming from other parts of the setup and environment. A Beam Current Monitor (BCM) in form of a scintillation telescope registered protons scattered on the exit window of the beam pipe. It was calibrated against the Faraday cup (CCB) or the indications from the EtherCAT system (HIT), and used to determine the total number of protons impinging on the target.

Proton beam energies and beam currents in the performed measurements are listed in Table 1. The measurement series at the energy of 130 MeV (or the closest available setting at HIT) was conducted in both centres to check the systematics. While in HIT the end of the ion pipe was equipped with a standard nozzle used in therapeutic rooms, the one in CCB had a simple titanium window of $50 \mu\text{m}$ thickness. Also the beam spatial and temporal characteristics were different due to different accelerators used in the two centres: a synchrotron at HIT and a cyclotron at CCB. Due to a different construction of the target holder in the two measurements, the setups were not identical. The distance between the exit window and the front face of the wedges was 37 cm (35.3 cm), and the distance between the exit window and the front face of the thin slice was 59.5 cm (63.3 cm) in HIT (CCB). Those small differences had their

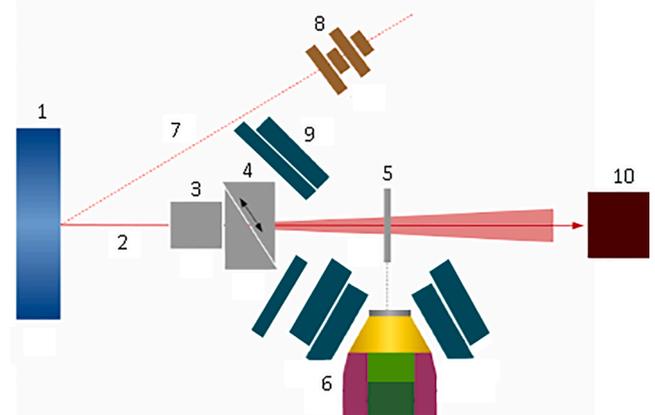


Fig. 1. Experimental setup used in CCB: 1 – exit window, 2 – proton beam trajectory, 3 – fixed extension, 4 – movable wedges, 5 – thin target slice, 6 – HPGe detector, 7 – sample trajectory of scattered protons, 8 – Beam Current Monitor, 9 – lead shield, 10 – beam dump.

Table 1

Overview of the conditions in the performed measurement series. In all cases the phantom material was PMMA.

Beam energy(MeV)	Proton range(mm)	Beam current(nA)	Facility
70.54	35.06	0.5	HIT
130.87	105.46	0.5	HIT
130	104.23	50	CCB
180	184.10	10	CCB
230	280.35	1.5	CCB

consequences: the beam cone leaving the thick part of the target, formed by angular straggling, was allowed to open more at CCB before reaching the thin slice than at HIT, which resulted in somewhat different shapes of gamma depth profiles.

2.2. Simulations

The aim in the design of the simulations was to produce data comparable with the measurement series, though without including the detector and shielding parts. These would need to be modelled in some way and could influence the simulation results. Therefore the geometry in the simulations contains only the elements that are traversed by the primary proton beam as described in Section 2.1. The changes in the target thickness in the simulations are realized like it was done in the measurements by re-positioning one of the target wedges. The proton beam conditions are modelled for each energy according to beam quality measurements at the respective facility. For 70 MeV the proton source was based on a realistic beam phase space input file generated downstream of the last element of the HIT beam line [29]. For the measurements at CCB the spatial spread was modelled according to measurements taken at the last beam profile monitor of the beam line and the proton energy spread was modelled with 0.2 MeV [30]. Simulations were performed for all conditions for which the experimental data existed (combinations of target thickness and beam energy). In each simulation the target was irradiated with $4 \cdot 10^8$ primary protons. As an output, all gamma quanta leaving the slice within the polar angle range covered by the detector in the real experiment were recorded. As no constraint on the azimuthal angle is imposed on the outgoing gammas, the solid angle of the simulations is larger than in the experiment. This is taken care of in the normalization stage, since the experimental yields are normalized by the detection solid angle.

To reduce the computation time, the production cuts for photons, electrons and positrons were set to 10 cm, while for protons they were set to $1 \mu\text{m}$. Those range cuts allow to reduce the computation time, but may lead to deficit in the number of generated low-energy gammas [31]. As they are irrelevant for the scope of this work, where we focus on the part of the spectrum above 2 MeV, the trade-off was acceptable.

GEANT4 is providing a big variety of reference physics lists with different use cases. The GEANT4 collaboration recommends the use of the QBBC reference physics list [32] for medical applications. However, other working groups tend to use rather the QGSP_BIC physics list [31,33]. That physics list is also used in a built-in advanced hadron-therapy related example in GEANT4 [34]. As mentioned in the introduction, also the QMD and INCL++ lists were used by other groups [24,26]. Our study is focusing on the list QGSP_BIC instead of the recommended QBBC for several reasons. Firstly, the QGSP_BIC is more commonly used by groups performing PG simulations. Secondly, there are only minor differences between the two [32], but the QGSP_BIC offers the advantage of including the High Precision model for neutrons and for light ions. Nevertheless, the QBBC list was also explored for all tested Geant4 versions and whenever differences between the lists occurred, they are reported. The AllHP variant of the QGSP_BIC list was tested only for the GEANT4 versions 10.6.3 and higher, as the earlier versions had known problems.

GEANT4 is a constantly evolving tool, which leads to changes in the models and their application regions. This, together with the constantly

updated databases on which some of the used models are based, leads to significantly different results for different release versions. The investigated reference physics lists and GEANT4 versions are listed in Table 2.

The simulations with the QGSP_BIC_HP physics list are indicated by * and referenced as star-simulations. The simulations with the QGSP_BIC_AllHP list are indicated with • and the simulations with the QBBC list are indicated with +. They are referenced as bullet- and cross-simulation, respectively. The GEANT4 versions are referenced with A, B, C and D for the 10.4 up to 10.7, respectively. In general, a reference physics list in GEANT4 comprises a purely hadronic part that uses different models and sets of cross sections to describe hadronic elastic, inelastic and capture processes, and an electromagnetic part describing the electromagnetic interactions of the simulated particles. The electromagnetic part in all investigated cases was set to EMZ which stands for the *electromagnetic standard option 4*, delivering the most precise results for electromagnetic interactions [32]. For hadron-nucleus processes, the QGSP_BIC_HP list invokes the Quark-Gluon String model (QGS), the Fritiof parton model (FTF), the Bertini Intranuclear Cascade model, the Binary Cascade model (BIC) and Precompound models. The Binary Light Ion Cascade (BIC) is used to simulate inelastic nucleus-nucleus scattering up to a certain threshold energy. The same process is modelled by the FTF model for energies above another energy threshold, with a combination of the models in the intermediate energy range. The energy regimes for the different models are version-dependent [32,35–37]. For protons and neutrons in the range between 0 and ∞ (5 GeV) the Binary Cascade model is used. For higher energies the FTF model and the QGS models are used. Again, the threshold energy differs for different release versions. Remnant nuclei from processes of these three models are de-excited by the use of the Precompound model (P). It is a combination of the Fermi breakup, neutron and light ion evaporation and photon evaporation models [32]. The investigated prompt-gamma production is described in the photon evaporation model as a combination of continuum gamma transitions using a dipole approximation and discrete gamma transitions which are data driven [27,38]. The loaded neutron High-Precision model (HP) is invoked for the processes of neutrons of 20 MeV and lower. This model is data driven and based on the G4NDL database which is derived from [38]. The QGSP_BIC_AllHP list is similar to the QGSP_BIC_HP list with the extension that inelastic interactions of protons and light ions (*d, t, ^3He , and α*) of energies below 200 MeV are treated by the ParticleHP model. The cross sections for this model are derived from the TENDL database [27,39]. The QBBC list invokes the same models as the QGSP list besides the QGS model. The ranges in which the different models are applied differ for the two lists and for different GEANT4 versions. Additionally, the lists use a different parameterisation of inelastic cross sections [32].

2.3. Analysis methods and foci

The analysis of experimental data covered several points. In the first step, the spectra have been energy calibrated using characteristic peaks visible in them. They have also been normalized by the number of

Table 2

Investigated combinations of physics lists and GEANT4 versions.

Physics list	GEANT4 version	Label	Reference
QGSP_BIC_HP	10.4.2	A*	star-simulations
	10.5.1	B*	
	10.6.3	C*	
	10.7.1	D*	
QGSP_BIC_AllHP	10.6.3	D•	bullet-simulation
	10.7.1		
QBBC	10.4.2	D+	cross-simulation
	10.5.1		
	10.6.3		
	10.7.1		

impinging protons and corrected for dead time. At this stage, the shapes of the experimental and simulated spectra were qualitatively compared, in particular the presence of peaks originating from discrete transitions. Special attention was paid to the 4.44 MeV and 6.13 MeV lines, as they are two most pronounced structures in the experimental spectrum. The first is formed by the gamma quanta stemming from the processes $^{12}\text{C}(p, p\gamma_{4.44 \text{ MeV}})^{12}\text{C}$ and $^{16}\text{O}(p, X\gamma_{4.44 \text{ MeV}})^{12}\text{C}$, and the second stems from $^{16}\text{O}(p, p\gamma_{6.13 \text{ MeV}})^{16}\text{O}$.

To estimate the quality of the internal broadening of the two peaks in GEANT4, the position and width for these two lines at a reference beam energy and target thickness were determined. In the simulations both peaks are background-free and have a Gaussian shape, thus their widths were determined by a Gaussian fit. For the experimental data, first the background around the peak was described by a parabola and subtracted. Subsequently, the oxygen line was fitted with a Gaussian and the carbon line, which has a complex, angle-dependent shape [40], was approximated as a combination of two Gaussian peaks, as shown in Fig. 2.

After the normalization, we constructed the dependencies of the gamma yield on the depth in the phantom, *i.e.* *depth profiles*. For each measurement series two separate profiles have been constructed for each of the investigated lines, like in our previous work [28]. To obtain the yields, the spectra were integrated within the limits given by the width of the peak. The integration of the complete peak in the simulated spectrum is necessary for a valid verification of the simulated yield (and thus the gamma production cross section) by a comparison with the experimental data.

All the integrated yields $N_{\gamma, \text{det}}$, both experimental and simulated, have been normalized by the thin slice thickness (*field of view*, FOV), the detection solid angle $\Delta\Omega$ and the number of impinging protons N_p . The experimental yields were additionally corrected for dead-time (ϵ_{DT}) as well as the energy- and rate-dependent detection efficiency (ϵ_{det}), including the ACS suppression factors. The formula is given by (1).

$$N_{\gamma}/FOV/\Delta\Omega/10^9 \text{ protons} = \frac{10^9 \cdot N_{\gamma, \text{det}}}{FOV \cdot \Delta\Omega \cdot \epsilon_{\text{DT}} \cdot \epsilon_{\text{det}}} \cdot N_p. \quad (1)$$

The yields calculated this way were plotted against the target thickness. It should be stressed that we use an effective target thickness, which includes not only the target material (extension, wedges, slice), but also

the material budget of the ion pipe exit window and the air gaps, recalculated to the corresponding thickness of PMMA yielding the same energy loss. In order to eliminate in the experimental profiles the contribution stemming from outside of the thin slice, a subtraction of the profiles obtained from the corresponding series measured without the slice was performed.

In the last step we quantify the width of the distal fall-off in the gamma depth profiles and analyze its dependence on the beam energy. For this purpose that part of the depth profile was fitted with the so-called associated error function

$$f(x) = a \cdot \left(1 - \int_0^{\frac{x-x_0}{\sqrt{2}\sigma}} e^{-t^2} dt\right), \quad (2)$$

where x_0 denotes the depth on which the gamma yield reaches half of its maximum value a , and σ is a measure of the fall-off width. To be consistent with the standard definition as the thickness on which the profile drops from 90% to 10% of its maximal value, we multiply the obtained σ parameter by a factor 2.56.

3. Results

3.1. Experimental results

In Fig. 3, the spectra registered from the target thickness a few millimetres upstream of the Bragg peak are presented. The target thickness, measured to the middle of the thin slice, is given with respect to the proton range (BP). Two features are immediately visible: the continuous background at CCB is higher than at HIT, and the contribution of that continuum increases with beam energy. Beside the 4.44 MeV and 6.13 MeV lines and their Compton structures and escape peaks, the following lines are visible in all spectra: 2.61 MeV and 3.00 MeV from activation of lead shielding and germanium material, as well as 2.21 MeV from neutron capture on hydrogen and 2.74 MeV from ^{16}O deexcitation.

Gamma depth profiles, obtained in the way described in Section 2.3, are presented in Fig. 4. Different panels correspond to different energies of the proton beam: 70, 130, 180 and 230 MeV. The lower horizontal axis represents the target thickness minus the proton range. The upper horizontal axis represents the mean energy of protons at the corresponding position. The data from HIT are represented by triangles and from CCB by circles. The statistical uncertainties are typically in the

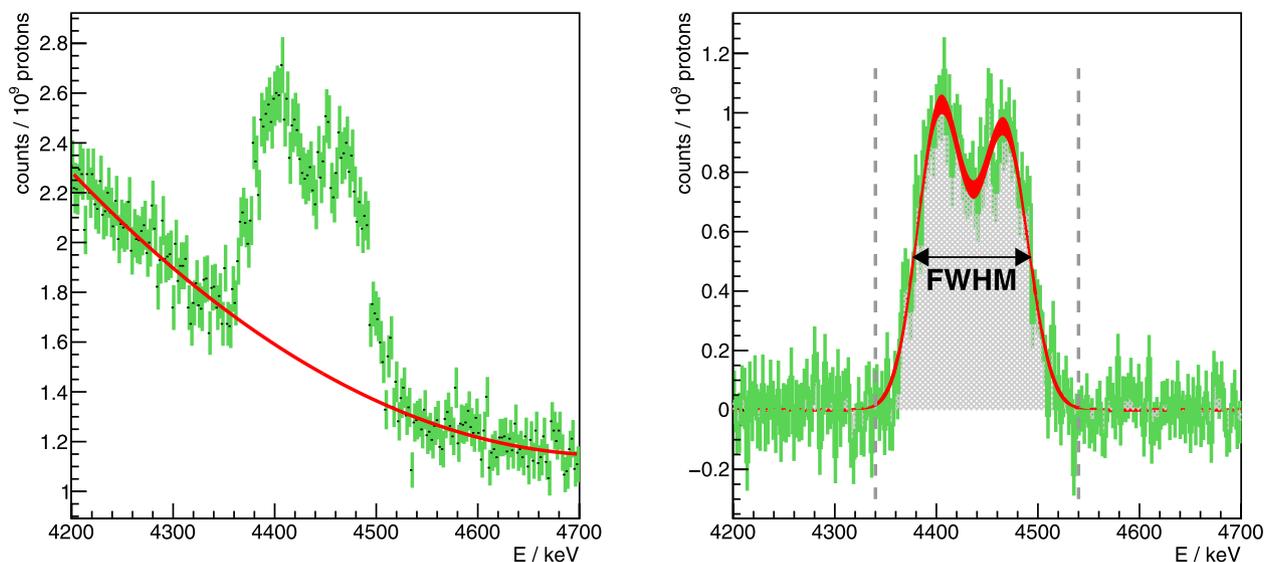


Fig. 2. (Color online) Determination of the 4.44 MeV experimental peak width. Left: zoom to the part of the spectrum containing the line, with a parabolic fit describing continuum component (background). Right: background-subtracted spectrum, with a double-Gaussian fit describing the line shape. The red band represents the 1σ confidence interval used to determine the line FWHM and its uncertainty. The grey area represents the peak integral between the dashed lines used to construct the gamma depth profiles. Note that there a different background subtraction method was applied though (see text).

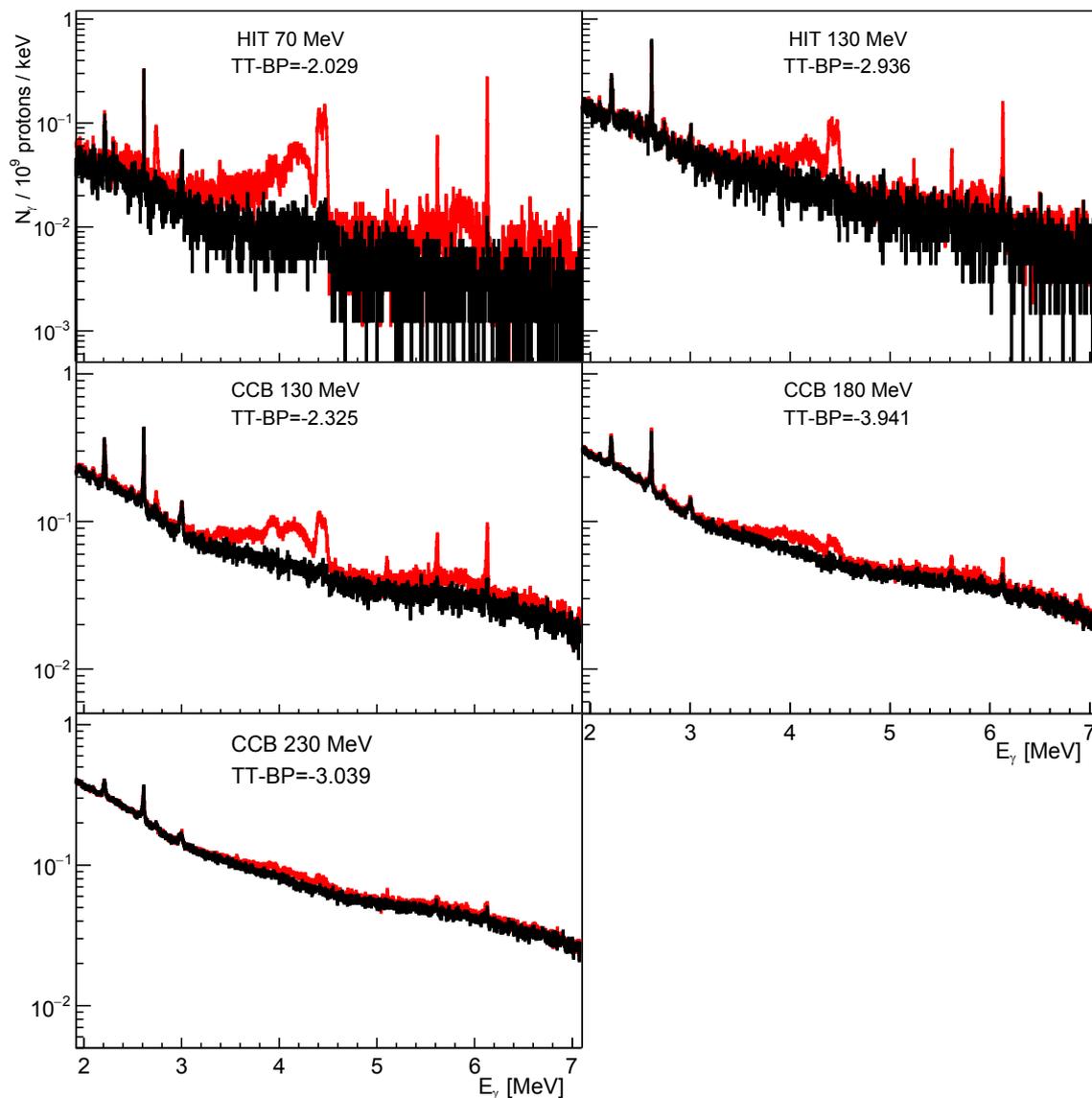


Fig. 3. (Color online) Gamma energy spectra recorded at CCB and HIT for different beam energies with a PMMA target slice (red) and without (black). The target thickness (TT) setting in mm with respect to the expected Bragg peak position (BP) is indicated in each panel. Please note, that the lower limit of the vertical axis in the upper-row panels is different than in other panels.

order of several percent, thus the error bars representing them are mostly contained in the marker symbols. Each profile is a subject of additional uncertainty in absolute normalization estimated to be about 15%, resulting from the statistical uncertainty of the BCM readings and the uncertainty of the BCM calibration input data. The uncertainty leaves space to vertical profile scaling, but without changing its shape. Similarly, the target thickness calibration, though precise to a few tens of microns point-to-point, can have an offset of up to 2 mm for the whole series. This systematic has not been corrected for in the presented results. The upper right panel of Fig. 4 contains the profiles obtained at nearly the same beam energy (130 MeV) at HIT and CCB, allowing verification of consistency of the experimental procedure. Although in general similar and consistent in magnitude, the profiles exhibit some differences in shape, which we attribute to different beam profiles and small differences in the setup geometry, mainly the bigger distance between the wedges and the thin slice at CCB. In the following analysis, in particular in the discussion of gamma depth profiles presented in Section 3.3, the CCB data for that beam energy have been used.

3.2. Results of GEANT4 simulations

It needs to be stated that for GEANT4 versions 10.4.2 and 10.5.1 no differences between QBBC and QGSP_BIC_HP were visible, these started to emerge since version 10.6.3. As neither for QBBC nor for QGSP_BIC_AllHP a difference between 10.6.3 and 10.7.1 results could be observed, only the results for 10.7.1 for both lists are included in the comparison. Similarly as for the experimental data, first the energy spectra obtained in simulations using different physics lists and GEANT4 versions have been compared. Fig. 5 presents the energy spectra of gamma quanta emitted from the slice located 2.325 mm upstream of the Bragg peak position at the beam energy of 130 MeV. The fits described in Section 2.3, used to parametrize the shape of the lines, are superimposed as red curves. The resulting integration ranges are indicated by the red (4.44 MeV line) and the green (6.13 MeV line) areas.

The parameters of the Gaussian fits, corresponding to the intrinsic peak widths of the physics lists and their mean energies are listed in Table 3.

The Doppler broadening in the bullet-simulation is clearly over-estimated. To check how this effect evolves with the mean proton

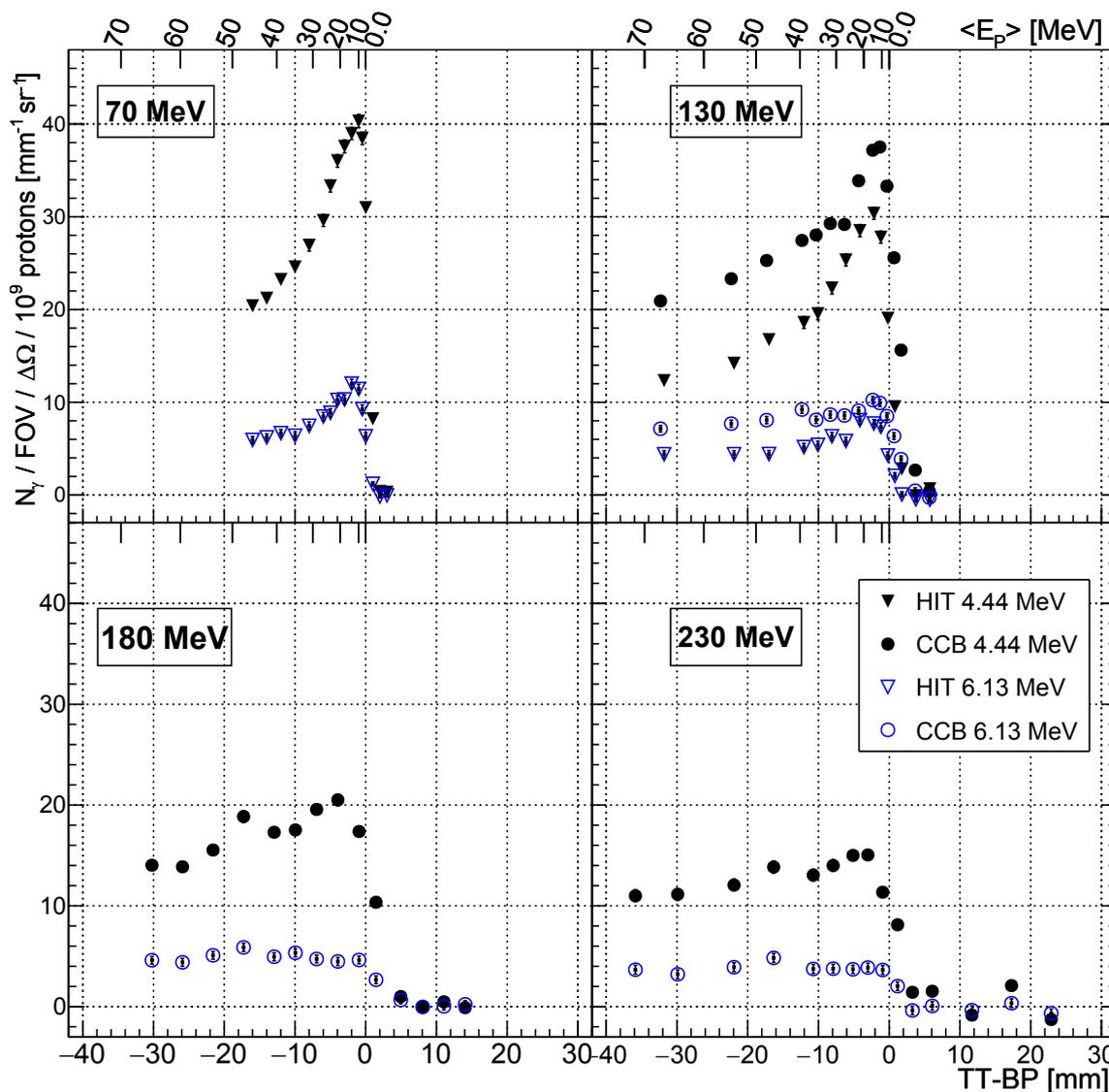


Fig. 4. (Color online) Measured depth profiles for the 4.44 MeV (black) and the 6.13 MeV (blue) gamma emission. Triangles represent data points of the HIT experiment, circles of the CCB one.

energy, we plotted the gamma energy spectra at different depths in the phantom for the D^* simulation. They are presented in Fig. 6.

In order to identify the processes leading to the creation of the discrete structures in the spectrum, the spectra of A^* have been decomposed based on the process the gamma quanta originate from. Access to this information required a modification of the GEANT4 source code itself. Simulations before and after these changes were compared to ensure that the results are not affected by these changes. Fig. 7, left shows such a stacked spectrum, with the right panels zooming into the regions of interest. For the A^* simulation, the integral of the 6.13 MeV line is strongly influenced by the overlapping 6.05 MeV peak. Therefore the integration of this line for that particular simulation is performed only for events stemming from the $^{16}\text{O}(p, p\gamma_{6.13 \text{ MeV}})^{16}\text{O}$ interaction.

The integration limits to obtain the yields are given in Table 4. The different integration ranges for experiment and simulations and even different ranges for different simulation settings were necessary due to the above described differences in the peak positions and widths of the peaks of interest. The ranges are kept constant within one simulation- or experimental series. For the star-simulations and the cross-simulation, the peak integration ranges were set to $\pm 3\sigma$ from the peak centre, where σ denotes the peak width determined in the reference spectra simulated at 130 MeV beam energy and a target thickness of $TT - BP =$

-2.325 mm (see Fig. 5). For the bullet-simulation, the integration range has been reduced to 2σ on the right-hand side to exclude events originating from a different line.

3.3. Experimental versus simulated depth profiles

The experimental depth profiles presented in Section 3.1 are compared to the depth profiles derived from the different simulation settings. The results for the four investigated beam energies for the carbon line 4.44 MeV are shown in Fig. 8 and for the oxygen line 6.13 MeV in Fig. 9. It should be stressed again that we use different integration ranges for the experiment and for each simulation.

Additionally to the shape of the profiles, the width of the distal fall-off visible in the experimental gamma depth profiles has been compared with the value extracted from the simulations, according to the procedure described at the end of Section 2.3. For clarity, only the A^* results were considered, as the obtained gamma depth profiles are most similar to the experimental ones. The results of such an analysis are shown in Fig. 10.

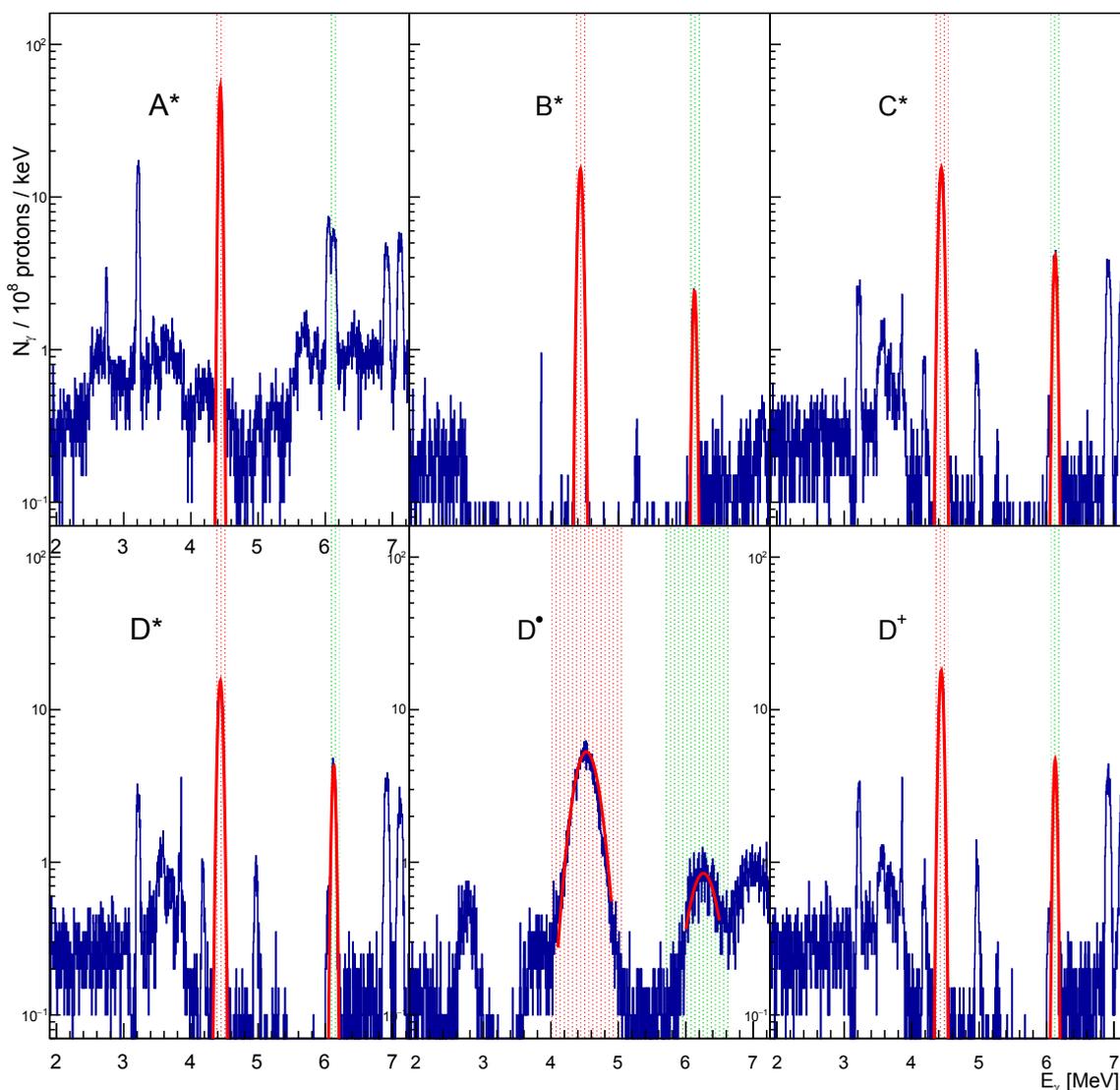


Fig. 5. (Color online) GEANT4 energy spectra obtained using different GEANT4 versions and physics lists for beam energy 130 MeV and PMMA target $TT - BP = -2.325$ mm.

Table 3

Widths of the two investigated lines 4.44 MeV and 6.13 MeV. Uncertainties in the last valid digits are given in brackets. Experimental data are included as reference. Values extracted for the beam energy of 130 MeV and PMMA target $TT - BP = -2.325$ mm.

Data set	Mean value 4.44 MeV line [MeV]	FWHM 4.44 MeV line [MeV]	Mean value 6.13 MeV line [MeV]	FWHM 6.13 MeV line [MeV]
Experiment	4.433(1)	0.101(3)	6.128(1)	0.019(1)
A*	4.4383(2)	0.0520(4)	–	–
B*	4.4386(5)	0.0767(6)	6.130(1)	0.059(2)
C*	4.4390(5)	0.0772(7)	6.1265(8)	0.063(1)
D*	4.4383(5)	0.0777(7)	6.1257(8)	0.064(1)
D°	4.53(3)	0.41(7)	6.260 (8)	0.47(2)
D+	4.4378(1)	0.0738(3)	6.1268(5)	0.058(1)

4. Discussion

4.1. Features of PG energy spectra

A striking feature of the set of experimental depth profiles presented in Fig. 4 is that their maxima get reduced with the increasing beam

energy. The origin is two-fold: smearing of the visible structures due to energy straggling as well as an increasing role of angular straggling. The latter leads to a loss of beam intensity in the slice acceptance for the beams penetrating larger thickness of the target at larger energies.

The energy spectra presented in Fig. 5 differ significantly from each other. Even within the same QGSP_BIC_HP physics lists (star-simulations), their shape varies strongly with the GEANT4 version. In particular, the changing features are the shape and magnitude of the continuum component, and the presence and intensity of peaks associated to discrete transitions. E.g. the two peaks 6.92 MeV and 7.12 MeV from ^{16}O deexcitation appear in versions A*, C* and D*, but are absent in B*. Striking for all produced spectra is the absence of the internal structure in the 4.44 MeV peak, which is clearly visible in the experimental spectra and has a well understood origin [40]. Regardless of that, all the star-simulations reproduce the expected position of the 4.44 MeV peak, though the obtained widths are smaller than the experimental result, which cannot be explained by the energy resolution of the detector, which is of a few keV. Larger differences are observed for the 6.13 MeV oxygen peak: version A* shows a double peak at that position (discussed below), other versions produce a single peak, but of a reduced intensity. Its centre for the star-simulations agrees with the expected value within a few standard deviations, but the width is overshot by a

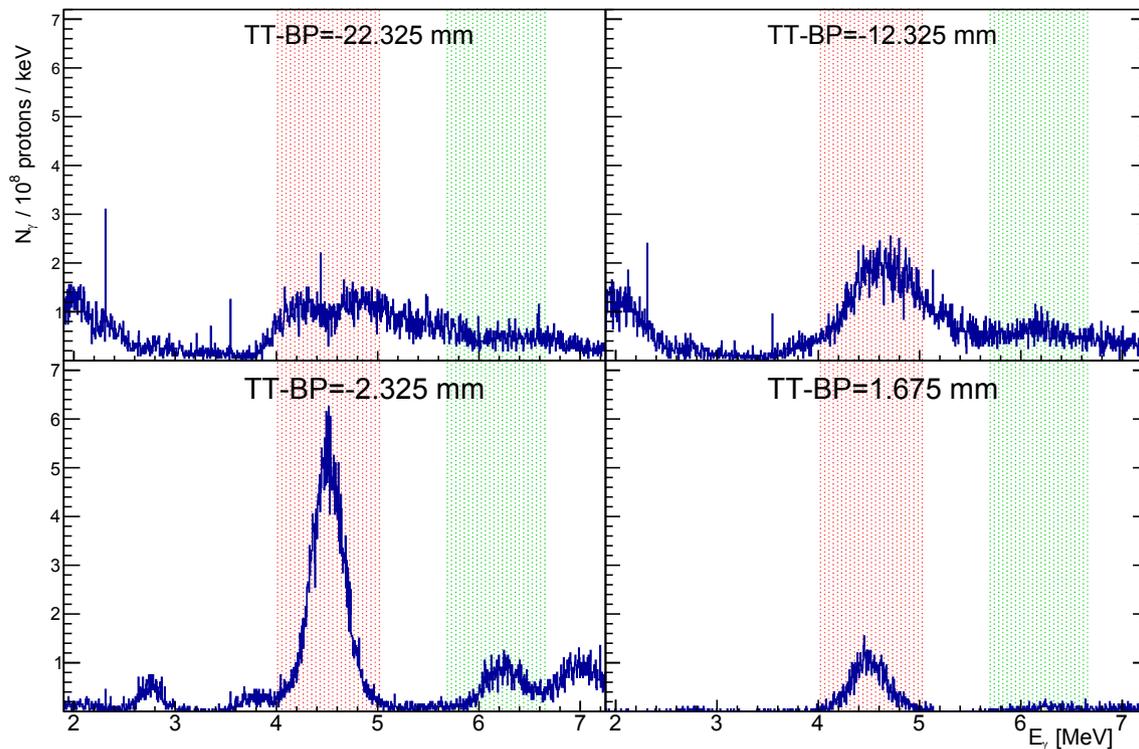


Fig. 6. (Color online) Energy spectra obtained using D^* simulation for a beam energy of 130 MeV and four settings of target thickness (indicated in the panels). The red areas indicate the integration range for the carbon line and the green area for the oxygen line.

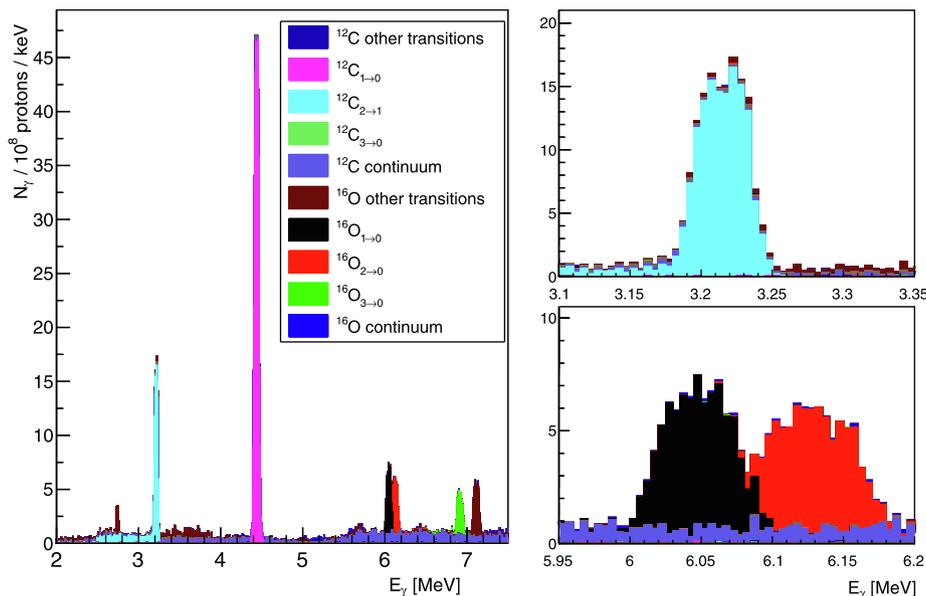


Fig. 7. (Color online) GEANT4 energy spectrum decomposed origin-wise, obtained using A^* for a beam energy of 130 MeV and target thickness of $TT-BP = -2.325$ mm.

factor of about three, indicating that the Doppler broadening is overestimated. It also shows the necessity to use individual integration ranges for the simulations.

Another peak, prominent in A^* , C^* , D^* and absent in other simulations, is located at 3.21 MeV. Fig. 7 allows to identify that the underlying process is the deexcitation in carbon $7.65 \rightarrow 4.44$ MeV. However, the carbon excited state of 7.65 MeV is the so called Hoyle state, thanks to which elements heavier than helium can be produced in stars. It is a ${}^8\text{Be} + \alpha$ resonant state (*i.e.* above the threshold for a strong ${}^{12}\text{C}$ decay),

and thus almost entirely disintegrates into 3α with the branching ratio of only $4.12 \cdot 10^{-2}$ into isomeric transitions [41]. This is an indication that the discussed peak is an unphysical structure in the simulated spectrum and explains why the transition is not visible in the experimental data. Another unphysical structure is the peak at 6.05 MeV. In the simulations, it corresponds to a radiative deexcitation of the first excited state of oxygen. Such a process would be a $0^+ \rightarrow 0^+$ transition, which is strongly suppressed by spin-parity selection rules, and the state deexcites predominantly via e^+e^- pair creation. The cross section for processes

Table 4
Integration ranges for the two investigated lines 4.44 MeV and 6.13 MeV

Data set	Integration range [MeV] 4.44 MeV line	Integration range [MeV] 6.13 MeV line
Experiment	4.34–4.54	6.07–6.15
A*	4.37–4.50	6.07–6.18 + Transition Condition
B*	4.34–4.54	6.05–6.21
C*	4.34–4.54	6.05–6.21
D*	4.34–4.54	6.04–6.21
D [•]	4.02–5.04	5.66–6.66
D ⁺	4.34–4.53	6.05 and 6.20

leading to the excitation of higher oxygen energy levels (6.92 MeV and 7.12 MeV) are overshoot as well, the experimental spectra do not show such prominent signals from their deexcitation. These strongly suppressed transitions are gone in the higher versions.

The features discussed above for the star-simulations are also valid for the cross-simulation. The latter reproduce the peak positions similar to the star-simulations and show the same overshoot in the broadening of the lines.

As for the bullet-simulation, it is apparent from Fig. 5 that the physics

list QGSP_BIC_AllHP largely overestimates the peak widths. Based on data from Table 3 one can see that for the 4.44 MeV the width is about 4 times too large, while for the 6.13 MeV about 20 times too large compared to the experimental results. Also the mean values of the peaks are shifted towards higher energies for the bullet-simulation. This behaviour is even stronger for spectra simulated further upstream of the Bragg peak, see Fig. 6. For target thickness far upstream of the calculated Bragg peak the 4.44 MeV line is losing a Gaussian shape and extends towards higher energies. This effect is so strong, that its tail is partially integrated for the 6.13 MeV line which has also an influence on the determined depth profiles for this bullet-simulations. For this reason the upstream parts of the depth profiles are not directly related to the gamma production cross section, as they contain contributions from the neighbouring lines.

4.2. Features of PG depth profiles

It is striking for Fig. 8 that none of the tested GEANT4 version and physics list combinations is able to reproduce the experimental results for all beam energies or in all parts of the depth profiles. The best agreement to the shape of the experimental depth profiles is found for the profile obtained with the QGSP_BIC_AllHP physics list (bullet-

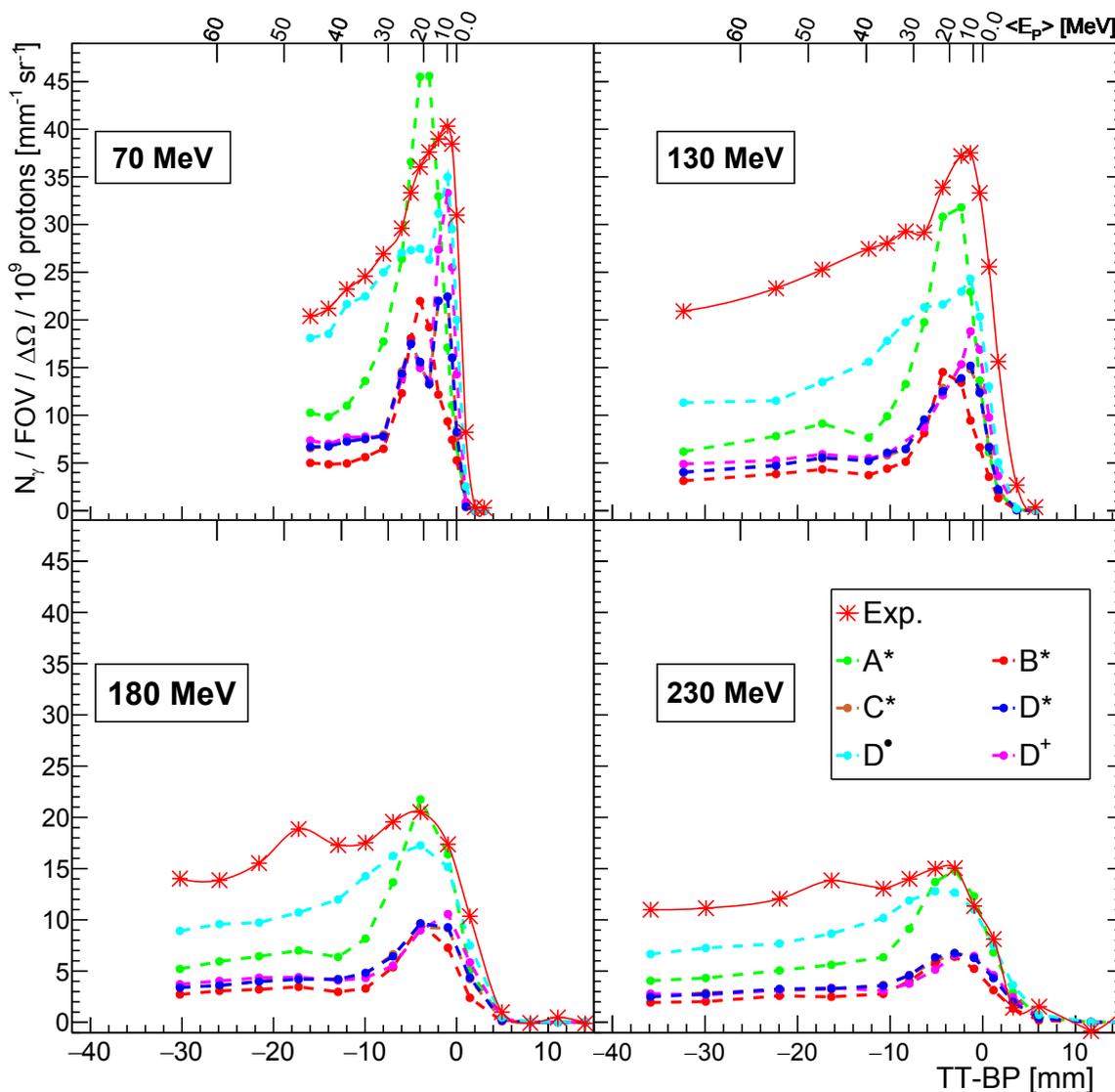


Fig. 8. (Color online) Comparison of gamma emission depth profiles for the 4.44 MeV line obtained from the simulations and the experiments for the beam energies 70 MeV, 130 MeV, 180 MeV and 230 MeV.

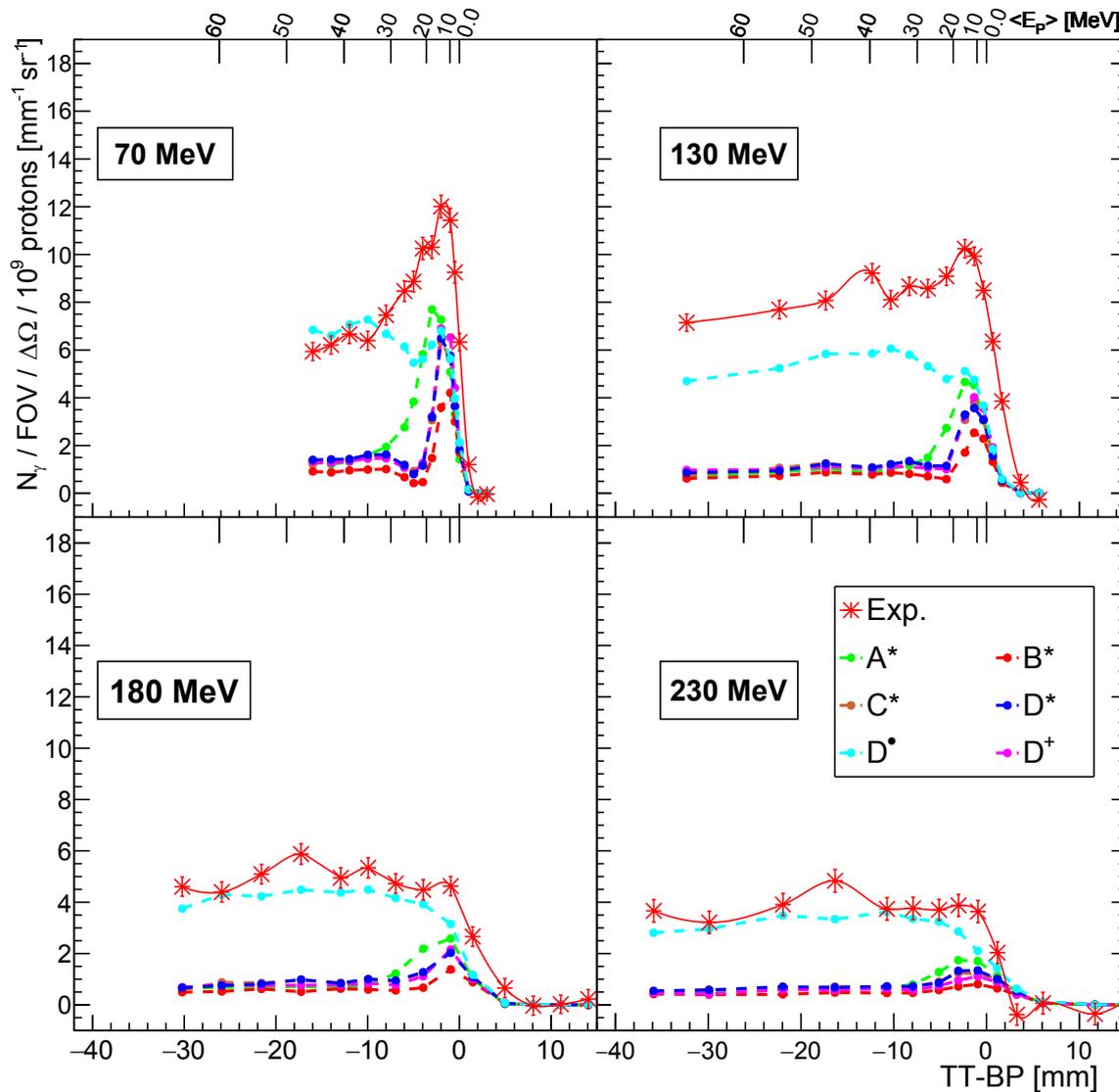


Fig. 9. (Color online) Comparison of depth profiles for $^{16}\text{O}_{6,13}$ obtained from the simulations and the experiments for the beam energies 70 MeV, 130 MeV, 180 MeV and 230 MeV.

simulation). For all beam energies this physics list reproduces the shape of the plateau region, the increase when approaching the Bragg peak region and the shape of the distal fall-off of the experiment. However, it should be kept in mind that in view of the finding presented at Fig. 5, the upstream points do not represent the full integral of the carbon line, thus the shape consistency of the simulated and experimental profiles is rather accidental.

Unlike for the bullet-simulation, for all the star-simulations and the cross-simulation the profiles represent full gamma yields enabling a meaningful comparison with experimental data. Here, the A* shows the highest gamma yield, consistent with the experimental results in the Bragg peak region for all beam energies, given the overall 15% experimental uncertainty. B* simulations deliver the lowest yields for all beam energies, only slightly increasing for C* and D*, which overlap completely. The yields in the Bragg peak regions for these three simulation settings are underestimated by a factor of about 2 for all beam energies. For all the star-simulations, the yield in the plateau region overshoots the experimental data even more than in the peak region, similarly to the results obtained in our previous work with TALYS calculations [28] and the findings of [14].

The profiles of the cross-simulation D⁺ are similar to the profiles of D*, with the largest differences occurring at lower beam energies in the

Bragg peak region.

The discussion presented above for the 4.44 MeV line remains valid also for the oxygen line profiles, though here the yield is underestimated for all simulated profiles. Simulation A* results in the highest yield of prompt gammas, though it is still a factor of about 2 below the experimental data. In case of this particular simulation the deficit could be compensated if the prompt gammas that originate from the unphysical radiative deexcitation of the first excited state of oxygen were redirected to populate the 6.13 MeV line.

For this line, in contrast to the results for the 4.44 MeV line, the results of the cross-simulation are fully consistent with D*.

The depth profiles derived for the oxygen line within the bullet-simulation are strongly affected by the spread of the carbon line into the integration range of the oxygen line. This is visible in the fact, that the maximum yield is not reached close to the Bragg peak but somewhere in the plateau region. Thus, although the profiles describe the plateau region fairly well, it is a rather accidental agreement keeping in mind that the plateau region has a large contribution of the smeared-out 4.44 MeV line.

Since the centres of the investigated peaks in the star-simulated spectra are correct, and the integration ranges for the lines have been adjusted to the simulated peak widths, it can be concluded that the cross

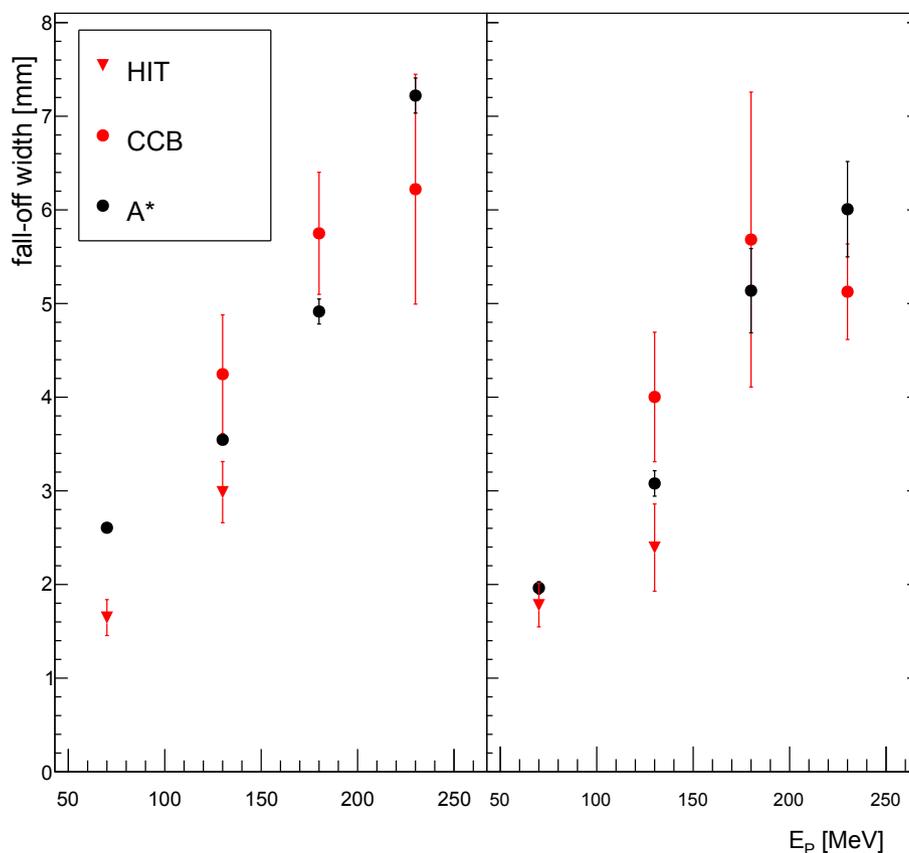


Fig. 10. (Color online) Distal fall-off width as a function of proton beam energy for the 4.44 MeV line (left) and the 6.13 MeV line (right). Red triangles and circles represent experimental points (HIT and CCB, respectively), black dots represent results of the GEANT4 simulation using QGSP_BIC_HP physics list of version 10.4.2.

sections for the interactions are not correctly modelled over the whole proton energy range. For the Bragg peak region, the yields of 4.44 MeV gammas obtained in the A* simulation are consistent with the experimental data, but in the plateau region they are much too small. The yields in both regions are even smaller in higher GEANT4 versions, leading to an even larger mismatch. For the 6.13 MeV line, the cross sections in the Bragg peak region are a factor of about 2 too low for the A* simulation and the negative trend in the development of the cross section for higher versions is also visible here. For the cross-simulation the investigated discrete peaks are also well visible in the spectra, so the profiles are representative for the cross sections. From the analysis of depth profiles it can be concluded that the cross sections for the Bragg peak region for D⁺ are slightly closer to experimental results than for D*. Still, the discrepancy is larger than for A*. Drawing meaningful conclusions for the bullet-simulation results is hindered by the large broadening of the gamma lines leading to their overlaps.

The shape of the gamma depth profile is a convolution of the beam attenuation (particularly important at the distal fall-off) with the proton energy dependence of the gamma production cross section, since the mean proton energy in the slice and its distribution are depth-dependent. Thus, a comparison of distal fall-off widths presented in Eq. (2) tests how well those two aspects are described in the simulations. Although the error bars are rather large, which results from a limited number of data points, it can be seen that the trends of the simulation and experimental points are consistent, which indicates that close to the Bragg peak, the beam attenuation as well as the dependence of the gamma production cross section on proton energy are correctly modelled. This comparison is insensitive to the absolute value of the mentioned cross section though.

5. Conclusion

We have presented results of systematic measurements of prompt-gamma emissions from PMMA targets irradiated with proton beams of 130 MeV, 180 MeV and 230 MeV. These results, complemented by the earlier experimental data obtained at 70 MeV proton beam energy, have been compared to simulations performed with the GEANT4 versions 10.4.2, 10.5.1, 10.6.3 and 10.7.1. For all versions the reference physics list QGSP_BIC_HP and QBBC have been tested. The reference physics list QGSP_BIC_AllHP has been tested for the versions 10.6.3 and 10.7.1. The analysis was focused on the 4.44 MeV line stemming from the $^{12}\text{C}(p, p\gamma_{4.44 \text{ MeV}})^{12}\text{C}$ and $^{16}\text{O}(p, X\gamma_{4.44 \text{ MeV}})^{12}\text{C}$ interactions and the line at 6.13 MeV stemming from the $^{16}\text{O}(p, p\gamma_{6.13 \text{ MeV}})^{16}\text{O}$ interaction. The experimental results obtained at the 130 MeV beam energy are consistent with results presented before in [28]. The comparison of simulated and measured spectra shows that the QGSP_BIC_AllHP physics list does not reproduce the features of the experimental spectra even qualitatively: the simulated spectra show an unrealistic shape evolution of the discrete lines along the proton path.

The simulations obtained with the QGSP_BIC_HP physics list reproduce the two investigated lines in all tested GEANT4 versions. Versions 10.4.2, 10.6.3 and 10.7.1 (A*, B*, D*) lead to the extraneous spectral line at 3.21 MeV corresponding to the deexcitation of the 7.65 MeV state of a ^{12}C nucleus to the 4.44 MeV state, while this state deexcites fully via α emission. Additionally, in A* also a 6.05 MeV line corresponding to a strongly suppressed $0^+ \rightarrow 0^+$ transition in ^{16}O is visible. Both of those structures are absent in the experimental spectrum. The second of those nonphysical lines disappears in versions of GEANT4 higher than 10.4, but at the same time also the yields of the investigated lines are decreased. Simulations obtained with the QBBC physics list show small difference with respect to the results obtained with QGSP_BIC_HP for

GEANT4 versions higher than 10.6.3. The cross sections in the Bragg peak region are slightly better modelled than with QGSP_BIC_HP in these versions, but still worse than in version 10.4.2 with QGSP_BIC_HP. We conclude that the best description of the two studied discrete transitions at 4.44 MeV and 6.13 MeV is provided in the A* simulation, though only for small proton energy, i.e. in the Bragg peak region. For realistic simulations even this version needs to be patched by redirecting the 6.05 MeV transition in oxygen nuclei into the 6.13 MeV one and elimination of the radiative deexcitation of the Hoyle state responsible for the appearance of the 3.21 MeV peak.

Acknowledgements

This work was supported by the Polish National Science Centre (grant number 2017/26/E/ST2/00618). The experimental part of the presented research was possible thanks to the HORIZON2020 Programme of the European Commission ENSAR2 (grant agreement no. 654002).

References

- Wrońska A, Dauvergne D. In: Radiation detection systems, vol. 2. Ed. by Iniewski K, Iwańczyk J. preprint at <https://hal.archives-ouvertes.fr/hal-03085504/document>. CRCPress/Routledge; 2021. Chap. 6.
- NuPECC report 2014: Nuclear Physics for Medicine. <http://www.nupec.org/pub/npmed2014.pdf>. Accessed: 2021-07-13.
- Liu CC, Huang HM. A deep learning approach for converting prompt gamma images to proton dose distributions: A Monte Carlo simulation study. *Phys Med* 2020;69:110–9. <https://doi.org/10.1016/j.ejmp.2019.12.006>.
- Richter C, Pausch G, Barczyk S, et al. First clinical application of a prompt gamma based in vivo proton range verification system. *Radiother Oncol* 2016;118:232–7. <https://doi.org/10.1016/j.radonc.2016.01.004>.
- Hueso-González F, Rabe M, Ruggieri TA, Bortfeld T, Verburg JM. A full-scale clinical prototype for proton range verification using prompt gamma-ray spectroscopy. *Phys Med Biol* 2018;63:185019. <https://doi.org/10.1088/1361-6560/aad513>.
- Ferrero V, Fiorina E, Morrocchi M, et al. Online proton therapy monitoring: clinical test of a Silicon-photodetector-based in-beam PET. *Sci Rep* 2018;8:4100. <https://doi.org/10.1038/s41598-018-22325-6>.
- Draeger E, Mackin D, Peterson S, et al. 3D prompt gamma imaging for proton beam range verification. *Phys Med Biol* 2018;63:035019. <https://doi.org/10.1088/1361-6560/aaa203>.
- Krimmer J. Development of a Compton camera for medical applications based on silicon strip and scintillation detectors. *Nucl Instrum Methods Phys Res A* 2015; 787:98–101. <https://doi.org/10.1016/j.nima.2014.11.042>.
- Solevi P. Performance of MACACO Compton telescope for ion-beam therapy monitoring: first test with proton beams. *Phys Med Biol* 2016;61:5149–65. <https://doi.org/10.1088/0031-9155/61/14/5149>.
- Gutierrez A, Baker C, Boston H, et al. Progress towards a semiconductor Compton camera for prompt gamma imaging during proton beam therapy for range and dose verification. *J Instrum* 2018;13:C01036. <https://doi.org/10.1088/1748-0221/13/01/C01036>.
- Cambraia Lopes P, Crespo P, Simões H, Ferreira Marques R, Parodi K, Schaart DR. Simulation of proton range monitoring in an anthropomorphic phantom using multi-slat collimators and time-of-flight detection of prompt-gamma quanta. *Phys Med* 2018;54:1–14. <https://doi.org/10.1016/j.ejmp.2018.09.001>.
- Wrońska A, Hetzel R, Kasper J, et al. Characterization of components of a scintillation-fiber-based Compton Camera. *Acta Phys Pol B* 2020;51:17–25. <https://doi.org/10.5506/APhysPolB.51.17>.
- Zarifi M, Guatelli S, Qi Y, Bolst D, Prokopovich D, Rosenfeld A. Characterization of prompt gamma ray emission for in vivo range verification in particle therapy: A simulation study. *Phys Med* 2019;62:20–32. <https://doi.org/10.1016/j.ejmp.2019.04.023>.
- Verburg JM, Shih HA, Seco J. Simulation of prompt gamma-ray emission during proton radiotherapy. *Phys Med Biol* 2012;17(57):5459–72. <https://doi.org/10.1088/0031-9155/57/17/5459>.
- Werner CJ (editor). MCNP Users Manual - Code Version 6.2. LA-UR-17-29981: https://laws.lanl.gov/vhosts/mcnp.lanl.gov/pdf_files/la-ur-17-29981.pdf. Accessed: 2021-07-13. 2017.
- The TALYS package. https://tendl.web.psi.ch/tendl_2019/talys.html. Accessed: 2021-07-13.
- Herman M, Capote R, Carlson B, et al. EMPIRE: Nuclear reaction model code system for data evaluation. *Nucl Data Sheets* 2007;108:2655. <https://doi.org/10.1016/j.nds.2007.11.003>.
- Battistoni G, Bauer J, Boehlen TT, et al. The FLUKA code: An accurate simulation tool for particle therapy. *Front Oncol* MAY 2016;6:11. <https://doi.org/10.3389/fonc.2016.00116>.
- Allison J, Amako K, Apostolakis J, et al. Recent developments in GEANT4. *Nucl Instrum Methods Phys Res A* 2016;835:186–225. <https://doi.org/10.1016/j.nima.2016.06.125>.
- Incerti S, Brown JM, Guatelli S. Advances in Geant4 applications in medicine. *Phys Med* 2020;70:224–7. <https://doi.org/10.1016/j.ejmp.2020.01.019>.
- Golnik C, Hueso-González F, Müller A, et al. Range assessment in particle therapy based on prompt -ray timing measurements. *Phys Med Biol* 2014;59:5399–422. <https://doi.org/10.1088/0031-9155/59/18/5399>.
- Nenoff L, Priegnitz M, Janssens G, et al. Sensitivity of a prompt-gamma slit-camera to detect range shifts for proton treatment verification. *Radiother Oncol* 2017;125: 534–40. <https://doi.org/10.1016/j.radonc.2017.10.013>.
- Kohlhase N, Wegener T, Schaar M, et al. Capability of MLEM and OE to detect range shifts with a compton camera in particle therapy. *IEEE Trans Radiat Plasma Med Sci* 2020;4:233–42. <https://doi.org/10.1109/TRPMS.2019.2937675>.
- Pinto M, Bajard M, Brons S, et al. Assessment and improvements of Geant4 hadronic models in the context of prompt-gamma hadrontherapy monitoring. *Phys Med Biol* 2014;59:1747–72. <https://doi.org/10.1088/0031-9155/59/7/1747>.
- Pinto M, Dauvergne D, Freud N, Krimmer J, Létang JM, Testa E. Assessment of Geant4 prompt-gamma emission yields in the context of proton therapy monitoring. *Front Oncol* 2016;6:10. <https://doi.org/10.3389/fonc.2016.00010>.
- Vanstalle M, Mattei I, Sarti A, et al. Benchmarking Geant4 hadronic models for prompt- monitoring in carbon ion therapy. *Med Phys* 2017;8(44):4276–86. doi: 10.1002/mp.12348.
- Physics Reference Manual, Release 10.7. <https://geant4-userdoc.web.cern.ch/UsersGuides/PhysicsReferenceManual/fo/PhysicsReferenceManual.pdf>. Accessed: 2021-07-13.
- Kelleter L, Wrońska A, Besuglow J, et al. Spectroscopic study of promptgamma emission for range verification in proton therapy. *PhysMed* 2017;34:7–17. <https://doi.org/10.1016/j.ejmp.2017.01.003>.
- Tessonnier T, Marcelos T, Mairani A, Brons S, Parodi K. Phase Space Generation for Proton and Carbon Ion Beams for External Users' Applications at the Heidelberg Ion Therapy Center. *Front Oncol* 2016;5:297. <https://doi.org/10.3389/fonc.2015.00297>.
- Eickhoff H, Dr. Weinrich U, and Alonso J. Design Criteria for Medical Accelerators. In: Ion beam therapy: Fundamentals, technology, clinical applications. Ed. by Linz U. Springer, 2012:325–43.
- Schumann A, Petzoldt J, Dendooven P, et al. Simulation and experimental verification of prompt gamma-ray emissions during proton irradiation. *Phys Med Biol* 2015;60:4197–207. <https://doi.org/10.1088/0031-9155/60/10/4197>.
- Guide For Physics Lists, Release 10.7. <https://geant4-userdoc.web.cern.ch/UsersGuides/PhysicsListGuide/fo/PhysicsListGuide.pdf>. Accessed: 2021-07-13.
- Zacharatos Jarlskog C, Paganetti H. Physics settings for using the Geant4 toolkit in proton therapy. *IEEE Trans Nucl Sci* 2008;55:1018–25. <https://doi.org/10.1109/TNS.2008.922816>.
- Cirrone G. Hadrontherapy example. <https://twiki.cern.ch/twiki/bin/view/Geant4/AdvancedExamplesHadrontherapy>. Accessed: 2021-07-13.
- Guide For Physics Lists, Release 10.4. <https://geant4-userdoc.web.cern.ch/UsersGuides/PhysicsListGuide/BackupVersions/V10.4/fo/PhysicsListGuide.pdf>. Accessed: 2021-07-13.
- Guide For Physics Lists, Release 10.5. <https://geant4-userdoc.web.cern.ch/UsersGuides/PhysicsListGuide/BackupVersions/V10.5-2.0/fo/PhysicsListGuide.pdf>. Accessed: 2021-07-13.
- Guide for physics lists, Release 10.6. <https://geant4-userdoc.web.cern.ch/UsersGuides/PhysicsListGuide/BackupVersions/V10.6c/fo/PhysicsListGuide.pdf>. Accessed: 2021-07-13.
- Rose PF. ENDF-201: ENDF/B-VI summary documentation. 1991. doi:10.2172/10132931. url: <https://www.osti.gov/biblio/10132931>.
- Koning A, Rochman D, Sublet JC, Dzysiuk N, Fleming M, and van der Marck S. TENDL: Complete nuclear data library for innovative nuclear science and technology. *Nucl Data Sheets* 2019;155. Special Issue on Nuclear Reaction Data: 1–55. doi: 10.1016/j.nds.2019.01.002.
- Rusiecka K, Wrońska A, Magiera A, et al. Shape of the spectral line and gamma angular distribution of the $^{12}\text{C}(p, p'4.44)^{12}\text{C}$ reaction. *Acta Phys Pol B* 2018;49: 1637–52. <https://doi.org/10.5506/APhysPolB.49.1637>.
- NDS IAEA Interactive chart of nuclides. <https://www-nds.iaea.org/relnsd/vcharthtml/VChartHTML.html>. Accessed: 2021-07-13.