Accelerated Clinical Prompt Gamma simulations for Proton Therapy

Brent Huisman

Accelerated Clinical Prompt Gamma simulations for Proton Therapy

A PhD thesis by Brent Fokke Bram Huisman
mail@brenthuisman.net

Supervision: David Sarrut & Étienne Testa

August 13, 2017

L'Institut National des Sciences Appliquées de Lyon

Numéro d'ordre: Année 2017

THÈSE

présentée devant

L'Institut National des Sciences Appliquées de Lyon

pour obtenir

LE GRADE DE DOCTEUR

ÉCOLE DOCTORALE N° 160 : ÉLECTRONIQUE, ÉLECTROTECHNIQUE, AUTOMATIQUE
FORMATION DOCTORALE : IMAGES ET SYSTÈMES

par

Brent Fokke Bram Huisman

Accelerated Clinical Prompt Gamma simulations for Proton Therapy

Soutenue le 19/05/2017

		Président
Peter Dendooven	Professor, Helsinki Institute of Physics, Helsinki Associate Professor, KVI-CART, Groningen	Rapporteur
Sébastien Jan	Ingénieur Chercheur, CEA, Orsay	Rapporteur
Magdalena Rafecas	Professor, Universität zu Lübeck, Lübeck	Examinateur
David Sarrut	Directeur de Recherche, CNRS, Creatis, Lyon	Directeur de thèse
Edmond Sterpin	Associate Professor, KU Leuven - Department of Oncology, Leuven	Examinateur
Étienne Testa	Maître de Conférences, HDR, Université de Lyon 1, IPNL , Lyon	Co-directeur de thèse

Abstract

After an introduction to particle therapy and prompt gamma detection, this doctoral dissertation comprises in its first part of original work the development, analysis and release of a variance reduction method for the simulation of prompt gammas (PGs) in clinical proton therapy simulations, achieving a gain of about a factor 10^3 with respect to analog Monte Carlo. Secondly, the method is used to study clinical deviations and the associated PG detection, which helps illuminate if and how PG cameras might be used in clinical scenarios. Spot-by-spot PG verification with two contemporary collimated PG cameras is found to be unlikely to produce usable results, but a spot grouping method is proposed that does not adversely affect the fall-off position (FOP) precision and makes millimetric FOP measurements possible.

The variance reduction method (named vpgTLE) is a two-stage track length estimation method developed to estimate the PG yield in voxelized volumes. A PG production database is precomputed once and can be reused for every patient-source combination. As primary particles are propagated throughout the patient CT, the PG yields are computed as function of the current energy of the primary, the material in the voxel and the step length. The result of this first stage is a voxelized image of PG production yield. The second stage uses this intermediate image as a source to generate and propagate the number of PGs throughout the rest of the scene geometry, e.g. into a detection device, corresponding to the number of primaries desired. For both a geometrical heterogeneous phantom and a complete patient CT treatment plan with respect to analog MC, at a convergence level of 2% relative uncertainty on the PG yield per voxel in the 90% yield region, a gain of around 10^3 was achieved. The method agrees with reference analog MC simulations to within 10^{-4} per voxel, with negligible bias.

The second major study conducted in this PhD program was on PG FOP estimation in clinical simulations. The number of protons (spot weight) required for a consistent FOP estimate was investigated for the first time for two optimized PG cameras, a multi-parallel slit (MPS) and a knife edge design (KES). By studying recent treatment plans from various proton clinics, we observe very few spots with weights over 10^8. Trends in increased precision in treatment planning seems to point towards more and small spots, at least for certain clinical conditions where tolerances are low. A CT and replanning CT (RPCT) was obtained in the Centre Léon Bérard clinic in Lyon. Initial treatment was created based on the CT image, and after the patient had suffered from significant weight loss, replanning was deemed necessary and the RPCT created. We created a proton treatment plan for the CT image, and studied the difference in detected PG signal by simulating the treatment with both the CT and RPCT. This is the closest a PG simulation study has come to studying the effect due to morphological change to date. Three spots were selected for an in depth study, and at the prescribed spot weights were found to produce results of insufficient precision, rendering usable clinical output on the spot level unlikely. When the spot weight is artificially increased to 10^9 primaries, the precision on the FOP reaches millimetric precision. On the FOP shift the MPS camera provides between 0.71 - 1.02 mm (1σ) precision for the three spots at 10^9 protons; the KES between 2.10 - 2.66 mm.

Grouping iso-energy layers was employed in passive delivery PG detection for one of the PG camera prototypes. Here we apply iso-energy grouping for both cameras, and compare it with a new spot

grouping method based on the notion of iso-depth. In iso-depth grouping, enabled by active delivery, spots with similar distal dose fall-offs are grouped so as to provide well-defined fall-offs as an attempt to sidestep range mixing. While with spot measurements the measured shift was in agreement with the PG emissions shift, for both grouping methods this was not the case. The precision of the shift improved with respect to iso-energy grouping for the KES camera (1σ = 3.05 and 2.41 mm, iso-energy and iso-depth resp.), but not for the MPS (1σ = 1.00 and 1.04 mm, iso-energy and iso-depth resp.). It is shown that grouping spots does not necessarily negatively affect the precision compared to the artificially increased spot, which means some form of spot grouping can enable clinical use of these PG cameras. With all spots or spot groups the MPS has a better signal compared to the KES, thanks to a larger detection efficiency and a lower background level due to time of flight selection.

Preface

This document is the record of the doctoral research program conducted by Brent Huisman from September 2013 to May 2017 in the CREATIS (Biomedical Imaging Research Lab) and IPNL (Institute of Nuclear Physics of Lyon) laboratories at the University of Lyon. Entitled "Towards real-time treatment control in protontherapy using prompt-radiation imaging: simulation and system optimization", work consisted of various studies of the characteristics of PG (Prompt Gamma) imaging for particle therapy. Most notable outcomes are the implementation of a variance reduction technique for PG simulation, available to the general public, and a study on clinical deviations during treatment and the possible use of PG imaging in detecting the change.

Acknowledgements

Towards the end of my time in Lyon I realized that during the interview for the position I did not actually realize what prompt gammas could be used for. In my master I helped develop a proton CT prototype, and in the beginning the difference between a (proton) CT and a prompt gamma signal were not clear. David and Étienne must have noticed this, but selected me for the position anyway. For that I am thankful, because I have learned a lot about the subject, I have learned how difficult running multi-year projects are, how immensely complex the structure of science supporting our field is and, hopefully, I have learned how to make a small but meaningful contribution. Without David and Étiennes decision to give me a chance, I might have never accomplished any of that.

Colleagues over the years have helped me learn or understand things, but since that list is too long to detail here, I have compiled an alphabetical list. Adrien, Alessio, Benoît, Caterina, Catherine, Cyril, David, Denis, Eduardo, Étienne, Evgenia, Fabrice, Floriane, François, Gloria, Ilaria, Jan, Jean-Michel, Jochen, Marc, Marco, Marie, Marie-Claude, Mattia, Mauro, Micaela, Nils, Olga, Simon, Thomas 1, Thomas 2, Valérian, thank you. In addition, thanks to the others that supported me with scientific questions, at conferences, next to coffee machines, in corridors or behind desks. Thanks to those that helped with administrative issues or gave advice about life in Lyon.

Lastly, I want to mention friends, family and my loved one. Your help and support during this time is appreciated. Dank jullie wel!

Table of Contents

Chapter 1

Particle Therapy

Cancers are malignant anomalous cell growths that limit normal human functioning and may result in permanent disability or death. Most cancers are environmental (pollution, lifestyle choices, infections) while some are due to an inheritable predisposition (International Agency for Research on Cancer, 2014). With age various risk-factors accumulate, which explains the higher incidence of cancer among the elderly. With continuous advances in health care, cancer is increasingly the leading cause of death. Preventative measures may stop or slow the development of certain kinds of cancers. If a cancer manages to develop, various types of countermeasures exist to combat it. Most common treatments include surgery, anti-cancer chemicals or drugs, possibly aided by biomarkers, and radio therapy (RT). Often, various modalities are combined to improve the chances for recovery.

In RT, ionizing radiation deposits dose in tissue, where dose quantifies the damage to the tissue. Radiation may be applied by introducing radio-active sources into the body, through ingestion, injection or placement through surgery. External beam radio therapy (EBRT), the most common form of radiotherapy, is applied by directing beams of ionizing radiation through the body. Localized cancers are commonly referred to as (malignant) tumors, and EBRT requires the cancer to be localized, because it targets spatial regions. About half of all tumors involve treatment with EBRT. Various kinds of ionizing radiation are or have been used, but by far the most common is X-ray: photons with energies in the 1-20 mega electron volt (MeV) range. X-rays have another useful feature: they can be used to image the body because different tissues attenuate the beam differently.

Particle therapy is a more recent type of EBRT, and it refers to the medical use of accelerated ions. *Hadron* or *ion therapy* would more correctly differentiate it from X-ray therapy from a physics point of view, but since the therapy involves multiple disciplines (physics, biology, medicine) the word *particle* appears to be the most widely understood term to differentiate it from X-ray therapy. Outside of experimental facilities, two species of ion are used in medical particle therapy: protons (about 90% of patients treated) and carbons (10%). In the rest of this thesis, the term *particle therapy* will be used.

Particle therapy is seeing an increasing number of patients treated (figure 1.1). Particles are accelerated in synchro- or cyclotrons and then 'shot' into the tumor. This modality can be considered a spin-off of high energy physics, which uses particle beams to probe matter and physical interactions. Medical accelerators are conceptually no different from experimental physics accelerators. There is a very appealing physics reason to choose particles over photons in EBRT: these ions exhibit a definite range in matter, as opposed to an exponential dose decrease through absorption as seen with X-rays (figure 1.2). Moreover, right before the end of their range, the deposited dose spikes: the Bragg Peak (BP). By placing this spike at the site of the tumor, for a given level of dose in the tumor, the surrounding tissue receives a reduced dose with respect to treatment with X-rays.

EBRT targets a spatial region, and for photons this means all tissues in the path of the beams. To deconcentrate the dose in the healthy tissue surrounding the tumor, X-ray treatments usually involve

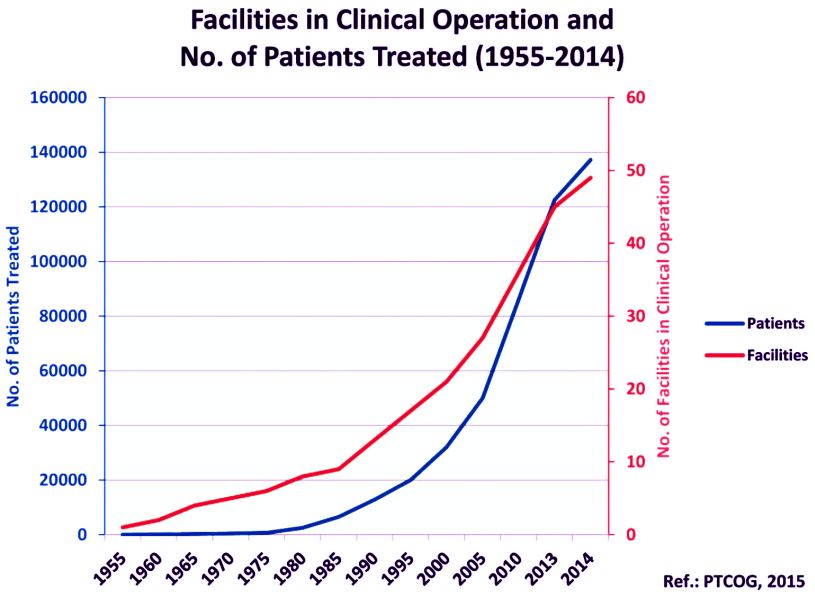


Figure 1.1: Particle clinics and patients treated with particle therapy as function of time (Jermann, 2015).

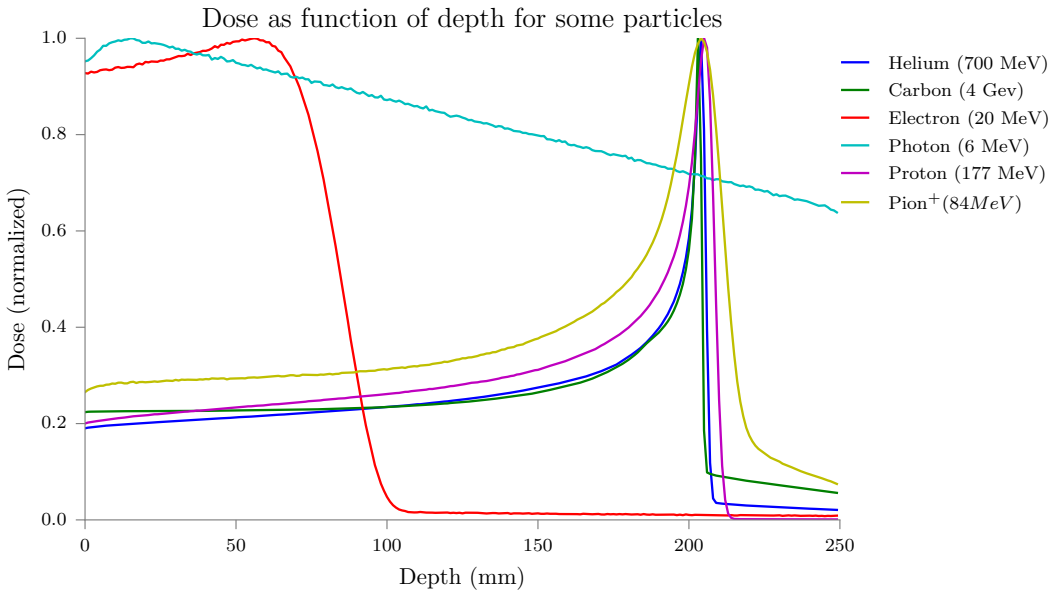


Figure 1.2: Depth-dose profiles for common particles in matter. Note the tails past the Bragg Peak for the heavier ions. Obtained by simulation with Gate/Geant4.

beams from multiple directions, fields, by rotating a gantry or moving the patient-table (intensity-modulated radiation therapy, IMRT). In particle therapy the number of fields can be reduced to one or two, whereas IMRT may have more than 10.

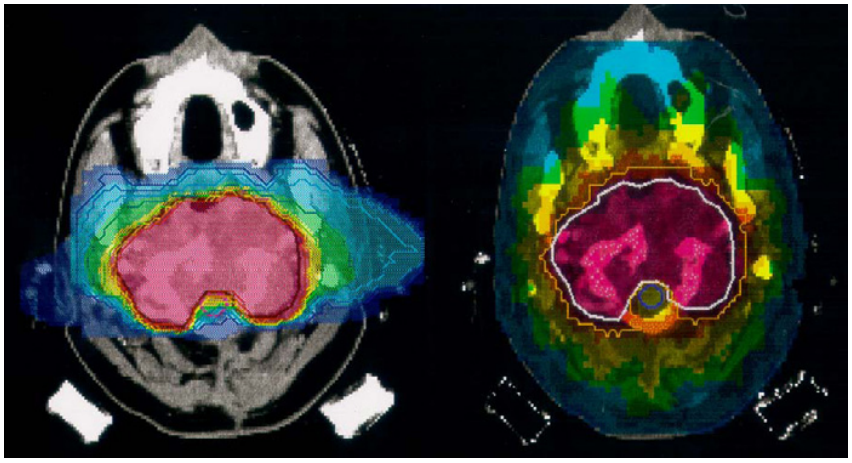


Figure 1.3: On the left a carbon irradiation using 2 fields. On the right a treatment plan for an X-ray irradiation using 9 fields. Using fewer fields the carbon plan manages a much more confined dose delivery. Note that in addition the carbon plan deposits almost no dose in the optical nerves, which is a well known type of tissue extremely sensitive to radiation. Source: Amaldi and Kraft (2005).

In summary, particle therapy appears to be an improvement over X-ray therapy (fig. 1.3). A significant disadvantage at this time is cost of treatment: particle therapy requires expensive particle accelerators. A proton treatment installation will cost, at time of writing, between one and two orders of magnitude more than an X-ray treatment installation, and facilities capable of accelerating heavier ions such as carbon even more. Moreover, a significant expansion of the hospital itself is often required to host the accelerator, and also more staff must be enlisted to operate the accelerator. Work is being done to make more compact and affordable medical particle accelerators possible. Dubbed "single room solutions", the Varian ProBeam and the IBA Proteus One are examples of such designs.

1.1 Interactions in matter

Particles lose energy through interactions with matter, thereby slowing down. The particle undergoes such interactions in rapid succession. My mental picture is similar to a game of *Sjoelen*, a traditional Dutch tabletop game resembling a mix of billiards and pinball, where wooden pucks are shuffled through sometimes dense 'forests' of other wooden pucks, with unpredictable and random trajectories as result. Collisions are mostly elastic, while the pucks lose energy through friction with the table. As in *Sjoelen*, mastery of the game of particle therapy reduces the unpredictability and may lead to victory.

1.1.1 Microscopic view

Collisions between particles, are either elastic or inelastic. For *elastic* collisions, the kinetic energy of the involved particles is preserved: any transfer of (kinetic) energy puts other particles in motion. An *inelastic* collision is where part of the kinetic energy is converted, putting target atoms or nuclei into excited states, or even the incident particle itself. Upon relaxation, new secondary particles may be created. These are often photons and sometimes positrons, which then may recombine with

Figure 1.4: A *sjoelbak*, a game that resembles particles traversing matter on a macroscopic scale. Image CC-BY-NC-SA licensed.

electrons and generate two back-to-back photons, for which PET detection (Positron Emission Tomography) was specifically designed. These secondaries may exit the target and give us hints as to what went on in the target; more on this later.

An incoming proton knocking an electron out of orbit into a trajectory away from its atom, ionization, requires a minimum amount of energy to liberate the electron from its atom, even if most of the transferred energy ends up in the form of kinetic energy of the electron (so called δ-rays). All in all, heavier ions ionize the target matter much more than lighter ions (with the same velocity), see figure 1.5.

Particle physics distinguishes four elementary interactions of forces: electromagnetic force, strong and weak nuclear force, and the gravitational force, each with an associated strength which has a force-specific long distance behaviour and range. For the purposes of accelerated particles in (baryonic) matter, the dominating force is the electromagnetic force, because it is much stronger than the gravitational force and the two nuclear forces have finite ranges on the order of subatomic scales. Only when an accelerated particle gets *very* close to a target nucleus, will the strong nuclear force become a factor. For the electromagnetic force, it is useful to distinguish the primary particles based on their electric charge: neutral particles, such as photons and neutrons, show very different interactions in matter than charged particles, such as protons, electrons and heavier ions. Incidentally, charged particles are much easier to accelerate than neutral particles. Photons being the only bosons in this list, are a bit of an exception, and are technically not the primary particle in X-ray generators. Electrons are accelerated, by creating a strong electric field, and then collide with a metal anode, where the photons (in medicine called X-rays) are created. For all other particles the same principle holds true as for the electron acceleration in X-ray generators: because of their charge, the protons or ions accelerate due to the externally applied electric field.

Neutral particles

Neutrons can only interact through the weak and strong nuclear forces, which means that only when they get very close to target nuclei they will scatter or excite a target nucleus. The electric fields of the nuclei and their electrons leave them unperturbed, so neutrons do not scatter in interatomic space. Photons, being the carriers of the electromagnetic force, *do* interact with the electromagnetic distortions that electric charges cause. As such, they undergo various interactions, including with target nuclei, where they may penetrate easily because they are insensitive to the weak and strong nuclear forces. Since nuclei are composed of (many) charged particles, held close together by the

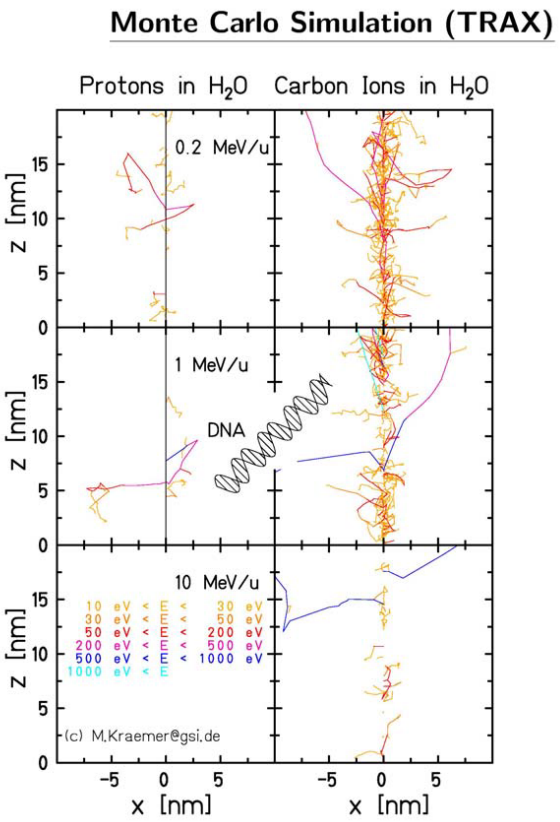


Figure 1.5: Carbon ions and protons have a very different ionization potential along their paths. Because carbon has much densely ionized clusters, the disrupted volume reaches the dimensions of a DNA strand, which then can be broken clean in two, greatly disabling the host cancer cells ability to recuperate. Image courtesy of Dieter Schardt.

strong nuclear force, a photon with sufficient energy, once in a nucleus, is likely to put the nucleus in an excited state. The nucleus may then decay under emission of another photon or light particles.

A bit further from the nucleus, but still close enough, the net electric charge may cause a passing photon of sufficient energy (>1 MeV) to undergo pair-production: the energy of the photon is converted into particle and its anti-particle, often an electron-positron pair. The likelihood of pair-production increases with photon energy (up to a certain point) and also approximately with the square of atomic number of the nearby atom. The inverse of pair-production, electron-positron annihilation, is the principle method behind PET (Positron Emission Tomography) imaging, as will be discussed in a later chapter.

The two other main interactions of photons with matter are the photoelectric effect and the Compton effect. In the photoelectric effect the photon energy is transferred to an electron, which, depending on the atomic orbital of the electron and the energy of the photon, may be enough to put the electron into an excited state or liberate it from its atom altogether. Elsewhere, the electron may then be reabsorbed and emit another photon. The Compton effect is an inelastic scattering of a photon by a charged particle, in regular matter most likely an electron. It results in a reduced photon energy which is converted into scatter of both the photon and the now free electron (momentum must be conserved).

Charged particles

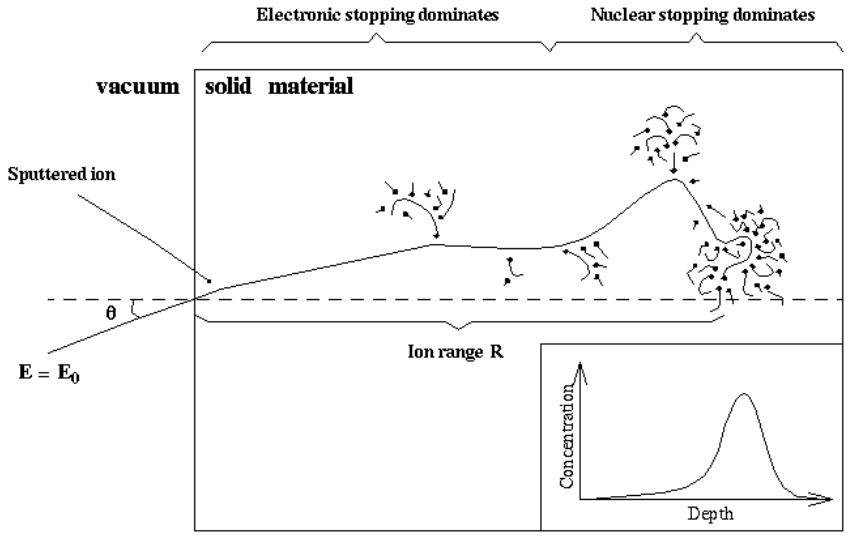


Figure 1.6: Particles in matter slow down through various interactions with the target matter. Electronic energy loss (interactions with target electrons) dominates along almost the entire path of ions in matter. Nuclear energy loss (energy transfer to target nuclei) becomes significant only at the very end of the incident ions path (the last tens of μm).

Charged particles traversing matter will interact with the target electrons or nuclei, both predominantly through electromagnetic interactions. Materials can be characterized by a *stopping power*: how much are traversing particles decelerated (through electromagnetic interactions)? For ions (not electrons), we find that electronic stopping is the main component slowing the particle down (fig. 1.6). Note that electronic or nuclear stopping refers to stopping due to electrons or nuclei in the target, *not* to the fundamental (strong nuclear, weak nuclear or electromagnetic) forces. Naturally, the incident particles properties (type, energy, mass, charge) affect the stopping power, thus stopping power is a

property for a certain particle moving through a certain target material.

The stopping power is an aggregate effect of the Coulomb force exerted on the primary particle by electrons and nuclei in the target. Depending on primary energy and mass difference with target particles, the primary is deflected because of these Coulomb forces. Since the number of target particles is of the order of the number of Avogadro, the number of such small angle scatters is large, which is why the effect is called Multiple Coulomb scattering (MCS), sometimes also referred to as elastic Coulomb scattering. The resulting spread of the beam in the transverse plane can be described in terms of angle θ_0 using equation 1.1 from Molière's theory (expressed in the Highland-Lynch-Dahl formula):

$$\theta_0 = \frac{13.6MeV}{\beta cp} z\sqrt{x/X_0}\,[1 + 0.038\ln(x/X_0)] \tag{1.1}$$

where x is the traversed depth through a certain material and p, βc and z are the momentum, velocity and charge number of the incident particle respectively, and X_0 radiation length. An illustration of the eventual spread for various particles and various energies can be seen in figure 1.7. Electrons, due to their lower mass, will be deflected over larger angles and therefore exhibit a more rapid loss of forward direction than hadrons. Deflection of particles in each others electromagnetic fields may also cause the production of photons: bremsstrahlung. Again, because of the mass difference, deflection of electrons is large close to nuclei, and the bremsstrahlung effect greatest.

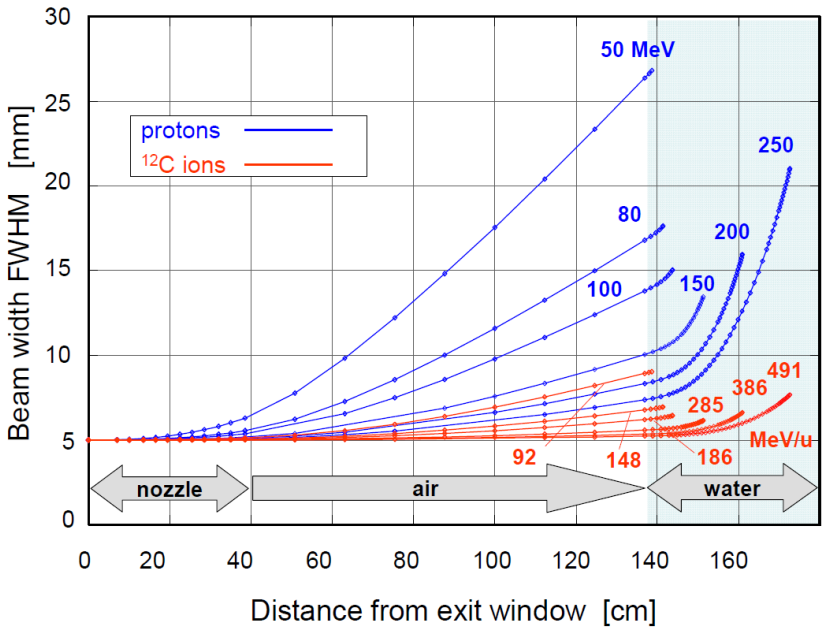


Figure 1.7: The beam spread for different types of hadrons at different energies are shown for a certain beam and a certain setup. The trend visible is that particles with higher energies and/or heavier ions provide a smaller spot size. This effect in combination with the former two results in a spread of the Bragg peak over a small volume, which is a few millimeters in depth (figure 1.2) a bit more in the transverse plane. Note that the distance from the exit window need not be this large, and the relative performance of the (lower energy) protons can be increased. Image courtesy of U. Weber of Rhön-Klinikum.

Nuclear interactions

Because of their mass and the large number of nuclei in the target, there is a non-negligible probability that the incident and target nuclei can get close enough for more complex nuclear interactions. Head-on collisions leads to nuclear fragmentation: a broken up ion translates to various new lighter nuclei, generally traveling at a similar speed to the parent ion. Heavier ions can break up into heavier parts (heavier than protons), and these parts in turn can break up again. If the incident particle is a low energy proton, a head-on collision may also lead to fusion: the creation of a heavier and excited isotope of the former target nucleus. Excited nuclei will decay, often by the production of secondaries, one of which is the emission of a photon. Such photons are produced within 1 ns of the collision, giving them their name Prompt Gammas (PGs). Detecting such a PG gives a very direct indication of where a collision took place, hence a likely location of dose delivery. Another possible secondary is the positron, which is emitted after β^+ decay, and thus not as prompt as PGs. After a short range it annihilates with the abundant electrons in the target, producing the signature head-to-head photons signal that lies at the core of PET. Note that other secondaries such as neutrons, protons and alphas radiation may be produced, each in turn possibly repeating the process at other target nuclei. Lighter incident nuclei, in particular protons, tend to not produce heavy (hadronic) secondaries, so for proton therapy in particular PGs are almost exclusively produced in primary-target interactions.

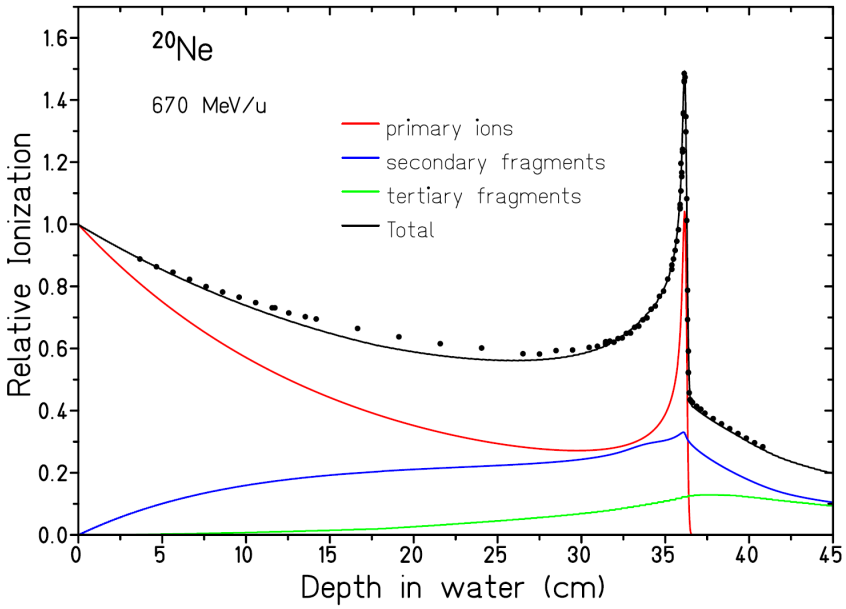


Figure 1.8: Dose curve for 670 MeV/u (^{20}Ne) ions in water measured at GSI and calculated contributions of primary ions, secondary and tertiary fragments. Image courtesy of Schardt and Elsässer (2010).

In collisions that are not quite head-on, where the primary and target ions graze, *fireballs* are produced. Only a few constituent protons and neutrons may be separated from target and primary, but they may inherit much of the velocity of the primary. These fireballs are responsible for the dose-tail past the Bragg-peak seen when heavier ions are used as primaries (see figure 1.9). Figure 1.8 shows for ^{20}Ne which fraction of primaries and secondary fragments contribute to the ionization. Heavier ions thus come at the cost of a delocalized dose deposition. Because protons can not break up, they don't suffer from nuclear fragments and the associated downsides.

An important consideration for particle treatment is neutron dose. When protons and host nuclei collide, free neutrons are one of the possible fragments. They tend to keep the speed and direction of the incident particle. Fast neutrons seem to have a small contribution (1% in actively scanned

carbon-ion therapy) to the effective dose (Gunzert-Marx et al., 2008). Neutrons do have a higher RBE (Relative Biological Effectiveness) than protons, and they are linked to induced secondary cancers. Brenner and Hall (2008) argue that since most neutrons are created in passive beam shaping objects, any adverse effect is easily remedied by using better accelerator designs and active scanning techniques.

1.1.2 Macroscopic view

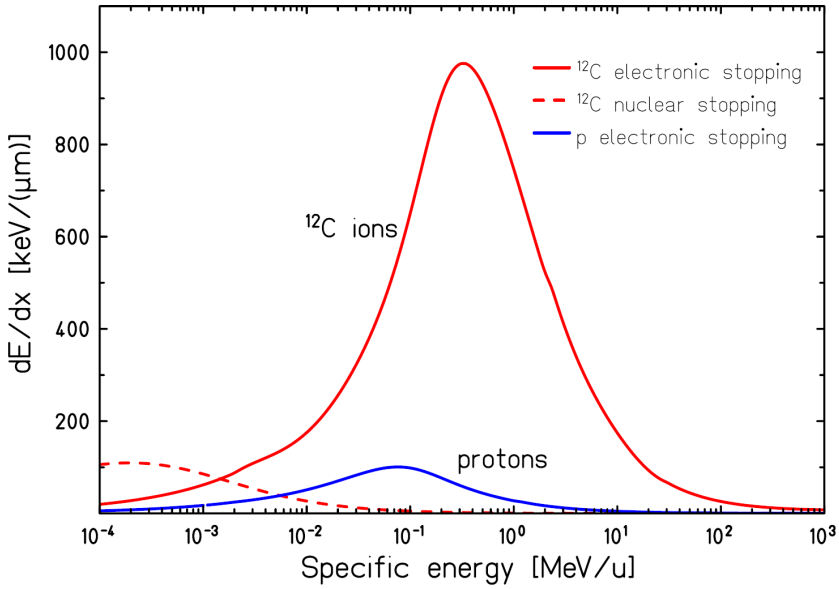


Figure 1.9: dE/dx curve, or stopping power, for protons and carbon ions as function of initial energy. Assuming the initial energy is in the range of a few 100 MeV/u, it is clear how the Bragg-peak comes about: the deceleration accelerates itself Schardt and Elsässer (2010).

As mentioned, the multitude of interactions and their infinite combinations can be aggregated in a macroscopic behavior. The macroscopic description for particles in matter is defined as the energy loss per unit distance traversed:the *stopping power*. It is quantified as $S(E)$:

$$\begin{aligned} S(E) &= S_{nucl}(E) + S_{rad}(E) + S_{elec}(E) \\ &= -\langle\frac{dE}{dx}\rangle_{nucl} - \langle\frac{dE}{dx}\rangle_{rad} - \langle\frac{dE}{dx}\rangle_{elec} \end{aligned} \tag{1.2}$$

where the energy loss due to nuclear interactions, radiative processes (bremsstrahlung) and electronic interactions are separated. All physics processes are reduced to a statistical average energy loss per unit distance traveled, and naturally the expression is different for different particles and circumstances. The Bethe-Bloch equation is one of the first things a young particle physicist will learn, because it is the expression for $S_{elec}(E)$, which is for light charged particles such as protons the main component of the full stopping power:

$$-S_{elec}(E) = \left\langle\frac{dE}{dx}\right\rangle_{elec} = \frac{4\pi}{m_e c^2}\cdot\frac{nz^2}{\beta^2}\cdot\left(\frac{e^2}{4\pi\varepsilon_0}\right)^2\cdot\left[\ln\left(\frac{2m_e c^2\beta^2}{I\cdot(1-\beta^2)}\right)-\beta^2\right] \tag{1.3}$$

where c is the speed of light and ε_0 the vacuum permittivity, $\beta = \frac{v}{c}$, e and m_e the electron charge and rest mass respectively, v the speed of the primary with charge z (in multiples of the electron charge) and energy E, traveling a distance x into a target of electron number density n and mean

excitation potential I. For protons, only at very low energy scales (< 1 MeV) does the nuclear component of the stopping power reach the same magnitude as the electronic component, which has next to no effect on the total distance traversed. Therefore, most sources on proton stopping power implicitly talk about the electronic stopping power.

The electronic and nuclear stopping powers for protons and carbon ions were shown in figure 1.9. Naturally, the precise curve is different for each combination of particle and target matter, but the main features are similar. At energies over ~100 MeV/u, the stopping power shows a wide minimum, a region often referred to as the minimum ionizing range. Going down from that minimum to lower energies, we observe a peak. Here, the energy deposition increases tremendously (note the logarithmic energy scale), and hence this is the clinical range. After all, the goal of particle therapy is to disrupt the target material, and the more energy can be deposited, the more damage is done. At very low energies the trend is downwards again, but at such energies, the particle range is so low it is almost immediately stopped, see fig. 1.10.

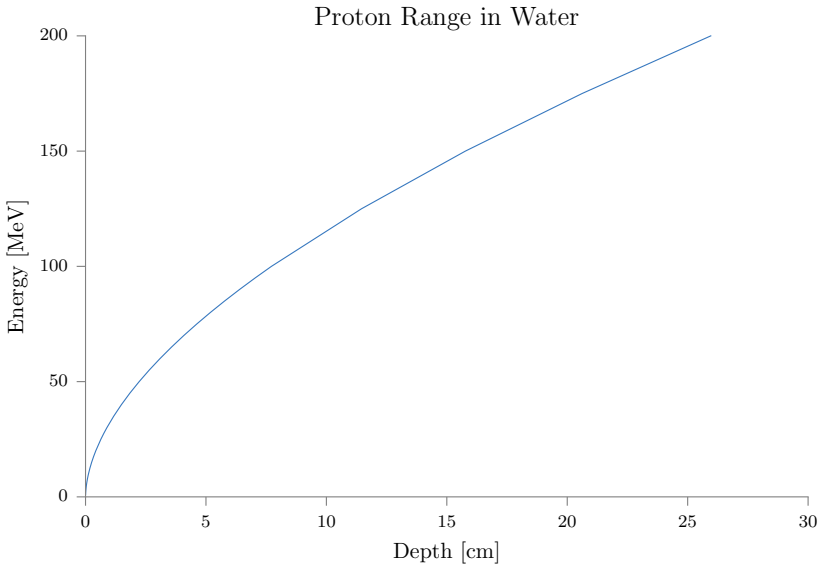


Figure 1.10: The range of protons in water computed in the CSDA approximation (Berger, 1995).

Now, with this macroscopic view in mind, shooting a particle into a block of matter is greatly simplified in comparison to simulating all the particular interactions described before. In certain cases the output of the physics processes are not of interest, but the average slowing down behavior of the particle and its deposited energy is. This is called the Continuously Slowing Down Approximation, or CSDA. If no other physics is considered, one can simply integrate the inverse stopping power of the energy of the particle (assuming the energy is low enough it does not pierce the target), and obtain the average or CSDA range r_{CSDA}:

$$r_{CSDA} = \int_0^{E_{\text{inc}}} \frac{1}{\rho} \frac{1}{\text{S(E)}} \text{dE} \tag{1.4}$$

where ρ is the density of the target material and S(E) the stopping power of the particle in that material, and E_{inc} the energy of the incident particle.

1.2 Dosimetry

Treating cancer with radiation is based on the idea that particles traversing tissue do damage. A definitive and comprehensive model for the quantification of damage due to radiation has not been found, and the field of radiobiology exists to answer this question. Along the particle track, as described by the stopping power (fig. 1.9), kinetic energy is lost and considered 'deposited'. In the most general terms, damage is regarded as receiving a *dose* D:

$$D[Gy] = \frac{E}{m}[\frac{J}{kg}] \tag{1.5}$$

where $\frac{E}{m}$ is the deposited energy per unit mass, with S.I.-unit Gray. The unit, an amount of energy per unit mass, implies that energy confined to a mass, at least for a certain period, does damage. Ionization may be regarded as such a confinement: it occurs along the primaries path and depends on the energy of the incident particle and the particular reactions that took place. In dosimetry the electronic stopping power is often called (unrestricted) Linear Energy Transfer or LET (Seltzer et al., 2011). Since energy is conserved, to first degree all the kinetic energy of the primary particle is transferred to the target. However, since some of the energy is converted into secondaries, which in turn may travel elsewhere, the unrestricted LET does not coincide with the deposited energy. The concept of deposition requires a criterion: how local and for how long must the energy remain to be considered deposited? After all, physically all energy disperses eventually. The concept of the restricted LET adds a δE threshold, under which the energy liberated is considered too low to escape through the production and travel of secondaries. In particular, the energy deposited by δ-electrons is controlled with this parameter. The threshold is arbitrary, thus a matter of context: in dosimetry studies it is usually set somewhere in the 10 keV range, while in microdosimetric studies the range is 100 eV (Shortt et al., 2006).

Regardless of the precise δE threshold chosen, LET is found to be a good proxy for damage due to ionization, and is easily integrated to attain the deposited dose from the energy loss, also known as absorbed or physical dose. However, this is still not the final measure of damage, as one of the first publications on it warned us for (Rossi, 1959). The stopping power and thus the LET are macroscopic quantities, describing energy loss through energy transfers. Of course, these transfers occur in discrete physical processes, and to understand the damage we must understand the quantity and location of these energy transfer points. Depending on the quantity of dose, the particle species and its energy (referred to as radiation quality), the transfer points may be distributed differently and have different average magnitudes. The LET describes just the macroscopic energy loss along the track. To incorporate the differences in damage of the various radiation qualities, a weight is added to convert the physical dose into an effective, equivalent or biological dose H:

$$H[Sv] = \frac{E}{m}[\frac{J}{kg}] \tag{1.6}$$

where the unit is Sievert to distinguish it from the physical dose.

In low dose environments, where the context of possible damage is radiosafety, such as nuclear power plants, people working with radiation sources or with machines generating radiation (X-ray machines), the equivalent dose is defined as:

$$H = w_F \times D \tag{1.7}$$

where w_F is the weight factor that is defined by the International Commission on Radiological Protection. As of 1990, for photons $w_F = 1$, for protons $w_F = 2$ and for heavier ions $w_F = 20$, and independent of the particle energy, except for neutrons which have a more complex weight factor.

In applications where radiation is directly imparted on a human for the specific purpose of damaging tissue, treatment, the equivalent dose is defined as:

$$H = RBE \times D \tag{1.8}$$

where RBE stands for the Relative Biological Effectiveness, relative to photon treatments. An important assumption in this equation is that the stochastic effects of a given type of radiation can be scaled to the reference radiation. The RBE for photons is thus 1 by definition and for protons is was concluded to be on average 1.1 (Paganetti et al., 2002). This means that to reach the default 2 Gy dose per treatment session, or fraction, proton beams actually deliver a physical dose of 1.82 Gy. For carbon ions, the RBE used varies between treatment centers (depending on the radio-biological model used), generally around 3 but in at least one case it is set to 1 (Durante and Paganetti, 2016). When doses are compared, people always speak about the (iso)effective dose H, which for photons is equivalent to physical dose D and therefore sometimes leads to H being expressed in GyE (Gray-Equivalent) instead of Sv (Fossati et al., 2012).

Various models attempt to predict RBEs based on a derivation and experimental inputs. In Europe, various versions of the Local Effect Model (LEM) are in use, while the NIRS (Japanese National Institute of Radiological Sciences) models are used in Japan. Colleagues recently published a new model, NanOx Cunha et al. (2017). Because studying the nanoscopic effects of particles in cells or organs is unfeasible, an aggregate biological effect can be obtained by counting the surviving fraction of cells in a petri-dish after exposing it to a particular radiation quality, under a particular circumstance. This is a standard technique in radiobiology. These curves are well described by the so-called Linear Quadratic (LQ) model:

$$S(D) = e^{-\alpha D - \beta D^2} \tag{1.9}$$

where S is the surviving fraction under dose D. α and β are free parameters, where α may be interpreted as lethal events and β as sublethal events. Instant cell death is a double DNA-strand break (DSB), while most other damage, even breaking one of the DNA strands, can often be repaired. Note that other interpretations of α and β exist. Lethal damage through non-DSB damage is a matter of piling on enough sublethal events to render a cell too damaged to repair. β therefore depends on the dose rate (how fast the dose is delivered) and the number of fractions for a given experiment, because the time the cancer cells have to recover determines the severity of the nonlethal events. A brief but nice overview of the links between α and β with biological types of damage is provided by Chapman and Gillespie (2012).

It is useful to consider the dissipation of energy in biological context. In dosimetry, direct and indirect damage (Podgorsak and Kainz, 2006) can be distinguished. Direct damage is then the instant alteration of the tissue by the primary and secondaries, while indirect damage are the delayed effects induced by oxidative species produced during irradiations (radiolysis). Often cells are modeled as bags of water (about 80%) containing a nucleus and a number of organelles. Atoms are bound in molecules, and excited atoms may cause molecules to break up or change function: the molecules are destroyed or damaged. The changed or destroyed molecules can in turn change the behavior of organelles or even the cell as a whole, which in turn can instigate all kinds of biological processes. In addition, in Cunha et al. (2017) it is found that about 20% of all energy is converted to heat: molecular vibrations, the interaction between electrons and water phonons and the reattachment of electrons to atoms.

In cell survival studies, the matter under investigation is just a few cell diameters thick. Over this range, the LET (or electronic stopping power) can be assumed to be constant, and that is why the radiation qualities can be ranked in terms of their LET. A higher LET translates to a higher potential for ionization. Photon treatments and lower LET radiation such as protons ionize less strongly, and

probably rely on the accumulation of effects to produce a lethal event, and will have a larger β, sometimes called the shoulder in cell survival curves, see figure 1.11. The α/β-ratio is thus an indication of the LET and the main instrument for determining the RBE.

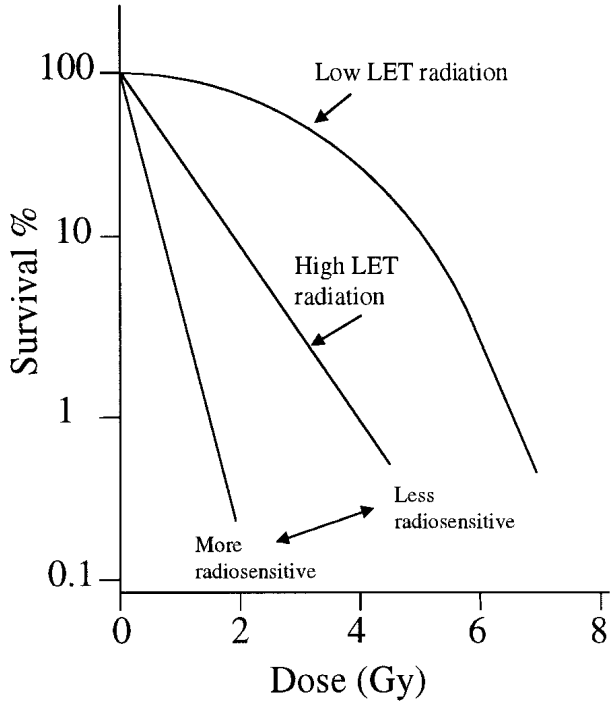


Figure 1.11: The typical shapes of cell survival curves for low and high LET particles (Pouget and Mather, 2001), where lower LET radiative qualities demonstrate the shoulder: an increased β component in eq. 1.9. These curves tend to describe experimental data well.

A clinical way to establish the efficacy of a treatment is to study the Tumor Control Probability (TCP) and the Normal Tissue Complication Probability (NTCP), see fig. 1.12. The surest way of destroying a tumor is to deliver as much dose as possible, but that leads to too many complications to surrounding tissues. Fortunately, normal tissue tends to be more resistant to radiation than tumorous tissue (erratic cells tend to be worse at damage reparation), which presents us with the so-called clinical window: the dose that allows us to eradicate (most of) the tumor but spare (most of) the surrounding tissues.

1.3 Particle beams in medicine

1.3.1 Accelerators

Performing a particle treatment is about putting particles in the right place. First, the particles must be accelerated. Cyclo- and synchrotrons are the main two types of devices that provide particle beams. The cyclotron is an older technique that has the advantages of being compact, providing high and continuous beam currents (fluence) at the cost of providing just one extraction energy. Two half cylinder shaped ends have internal magnetic fields keeping the particles in their (expanding) orbit, while between the two halves an electric fields provides the acceleration. The frequency of this field determines the micro time structure of the beam, which is an important property for detectors further down the line. The net effect is that particles are delivered in small 'bunches', with 'dead time' between them: the duty cycle. Techniques such as the synchro-cyclotron and more recently the isochronous cyclotron may vary the magnetic field (as function of time or distance to the center

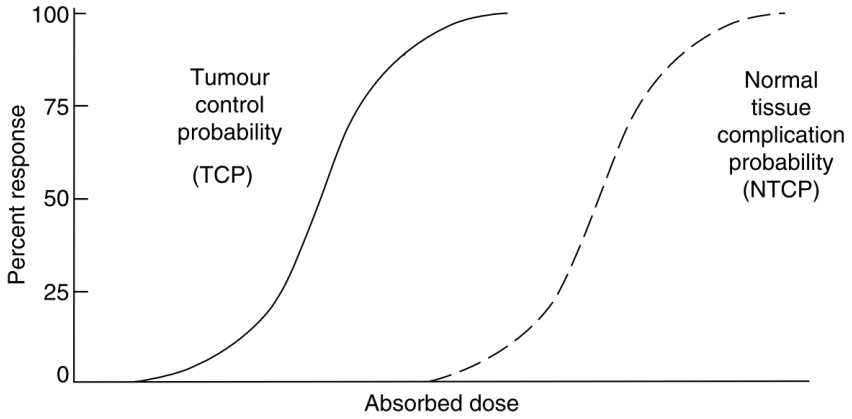


Figure 1.12: The TCP and NTCP plotted as function of absorbed dose. The region between the curves is the therapeutical window: ideally a treatment meets or exceeds the TCP line, but stays below the NTCP line.

respectively) to deal with relativistic effects, enable smaller or cheaper cyclotron geometries, provide higher energies and so on, at the cost of a lower duty cycle, see table 1.1. In addition, there is now a macro time structure to consider: since the fields change, the cyclotron can no longer be loaded continuously, but once per magnetic or electric cycle.

Table 1.1: Main parameters of three particle accelerators. Data collected from manufacturers presentations.

	Cyclotron	Synchrotron	Synchro-cyclotron
Example	IBA C230	HIT	IBA S2C2
Energy at extraction	230 MeV	221 MeV	230 MeV
Macro time structure			
- period	-	≈ 10 ns	1 ms
- pulse length	-	≈ 10 ns	1 ms
Micro time structure			
- period	≈ 10ns	≈ 100 ns	10 ns
- bunch length	≈ 2 ns	≈ 10 ns	2 ns
Mean beam current (at nozzle)	3 nA	max. ≈ 1.5nA	max. ≈ 2 nA
Current during macropulse	-	max. ≈ 3 nA	max. ≈ 300 nA
Duty factor	0.2	0.05	0.0014
Protons per bunch	≈ 200	≈ 2000	≈ 20k

The synchrotron is made up of multiple linear accelerators, usually placed in a circular setup so that a single accelerator segment can accelerate the bundle multiple times by increasing their electric field strength in each successive cycle. Synchrotrons are larger installations, but can extract particles at various energy levels, obviating the passive (and scatter and secondary inducing) slowing down methods required with cyclotrons. Due to most synchrotrons being circular in order to reuse accelerator components, they also have a macro time structure: the ring is filled once and then the particles are cycled up to the requested energy and extracted. Macro time structure means extra dead time with respect to the micro time structure, translating to a lower duty factor, as can be seen

in table 1.1.

In the end, the choice of the accelerator design depends mainly on the goal of the facility, financial factors and building size. Most research oriented facilities and multiple ion facilities favor a synchrotron design, because of its flexibility (multiple ions basically exclude cylcotron-based designs). Smaller facilities or facilities with a purely clinical focus tend towards the cheaper and smaller cyclotron designs.

1.3.2 Beam shaping

Figure 1.13: A photo of the gantry at the Heidelberg Ion-Beam Therapy Center (HIT), in the horizontal position. The support structures are painted blue and the bending magnets orange. Photo courtesy of Universitätsklinikum Heidelberg.

A beam line guides the particle bundle, again by way of electric and magnetic fields, towards the patient. Some treatment centers have a gantry right before the bundle reaches the patient: a rotatable beam line so that the bundle can impinge from different angles without having to move the patient. Since the beam line components are heavy, even moreso when the bundle needs to bend, which is the case in a gantry, such installations are usually very large (see fig. 1.13 for an example). Many particle therapy centers therefore have fixed beam lines, which are cheaper, lighter, require less space and are easier to calibrate, but require the patient to be repositioned if a beam must be delivered from multiple angles.

Once the particle bundle exits the beam line, it must be shaped so that all particles stop exactly inside the tumor. Shaping the beam is done in one of two ways: passive or active shaping (fig. 1.14). Passive shaping is an older method, and is often used in combination with cyclotrons which output particles with a single energy. A range shifter element decelerates the particles such the bundle now consists of a spectrum of energy. A scatterer widens the beam and a collimator and a compensator put the beam in its final shape, so that the tumor and only the tumor is covered. All these elements are put in place and to deliver the treatment we only have to turn on the cyclotron and wait. Since

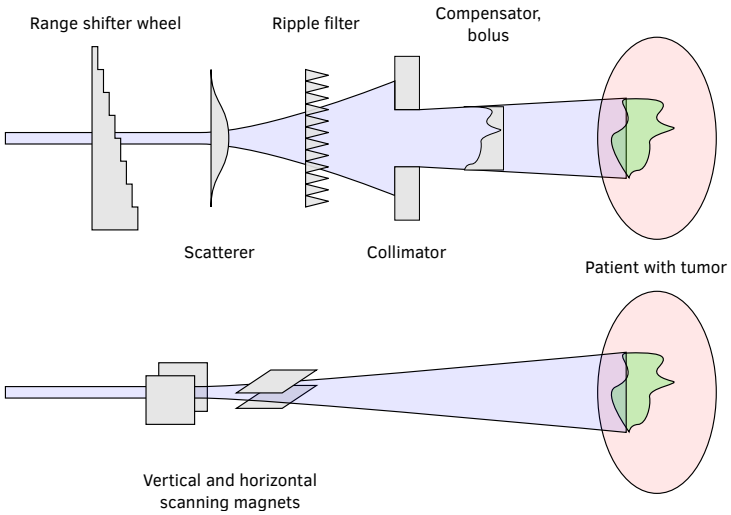


Figure 1.14: The two concepts applied for beam shaping in particle therapy. On top: passive beam shaping, which irradiates the full field at once. The beam's energy and direction are manipulated by inserting high-Z objects in the beam. The range shifter is a wheel that spins, and has a stepped or continuous slope such that the beam is slowed down over a range of energies. The scatterer widens the beam to a diverging bundle, such that we can cover the whole field of view at once. Sometimes a second scatterer is introduced to provide a flat particle fluence, within a certain aperture. An optional ripple filter can be present, in particular for particles with sharp BPs. A collimator is then used to adjust the beam to cover the contours of the tumor. A patient specific compensator or bolus then adjusts the range, so that in each position in the transverse plane the particles energy is tuned to stop at the distal point of the tumor. On the bottom: active shaping. An optional ripple filter or range shifter wheel may be present for the same reasons, and then a pair of magnets steer the bundle into a certain position in the transverse plane. Note that this is an active process, so the tumor volume is not irradiated at once, but point for point.

these elements are high-Z elements (they adjust the trajectories of the particles), there are many effects that increase the uncertainties of the bundle position and energy. Optical effects such as chromatic aberration and interface effects such as the West-Sherwood effect occur, but also secondary production in the elements, of which neutron production is the most problematic.

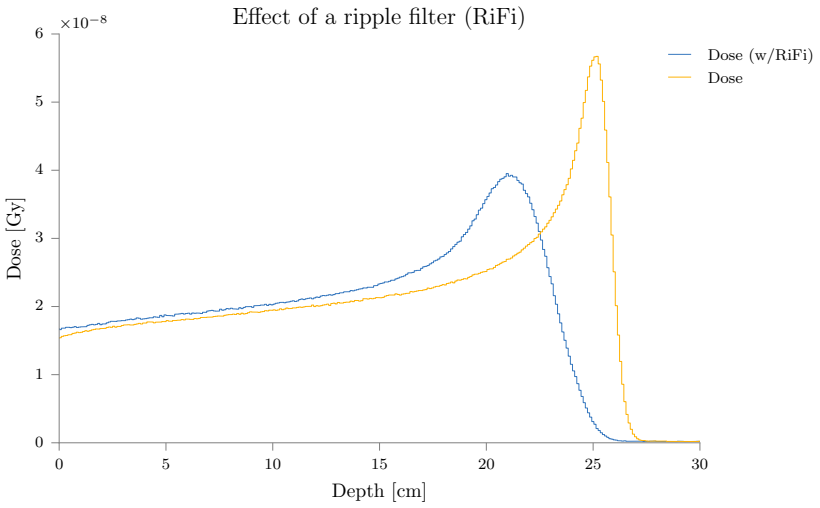


Figure 1.15: The effect of a Ripple Filter on the range of 200MeV protons in water. Ripple Filters (RiFi) are used to introduce artificial smearing of the BP for heavier ions: their Bragg Peaks are so sharp that the number of energies that the synchrotron would have to deliver to provide a smooth SOBP would be too great.

On the other hand, active scanners use magnetic fields to steer the beam and are always employed in tandem with an accelerator that can deliver the particles over a certain energy range. This means most of the passive elements are no longer required to cover the tumor volume. For heavier particles such as carbon, the ripple filter is often used as it is in passive shaping (see fig. 1.15). An important difference with passive delivery is that in an active setup, the magnets can only steer the beam into one particular position (it does not widen the beam), and the synchrotron delivers a single energy at a time. This means the volume must be scanned over, both with the magnets and synchrotron.

The typical dimensions of tumors are significantly larger than the typical width of the BP. As mentioned, the ripple filter is sometimes employed to widen the BP, but to reach across the full depth of the typical tumor, multiple energies are required (each delivering ions at a certain depth as per fig. 1.10). Therefore, peaks of various energies must be overlaid to scan over the full range of the tumor. These different energies are obtained either by the synchrotron or by inserting slabs of meterial as range shifters into the beam. This leads to the Spread Out Bragg Peak (SOBP), also known as peak summing effect (fig. 1.16). The entrance dose is enlarged, making the apparent advantage of particle therapy decidedly less pronounced. Multiple beams may help spread out this entrance dose, as is done in X-ray treatments.

1.3.3 Particle species

Several ions are or have been considered for medical treatment: H, D, He, Li, Be, C and Ne. The upside to heavier ions is less lateral scattering and an increased RBE compared to protons, predominantly due to the nucleus being heavier and increased ionization respectively. The downside is fragmentation: the debris stops before or after the Bragg peak and makes the dose deposition less accurate. This is compensated for by a local increase of the RBE in the BP region: the RBE may in fact be a func-

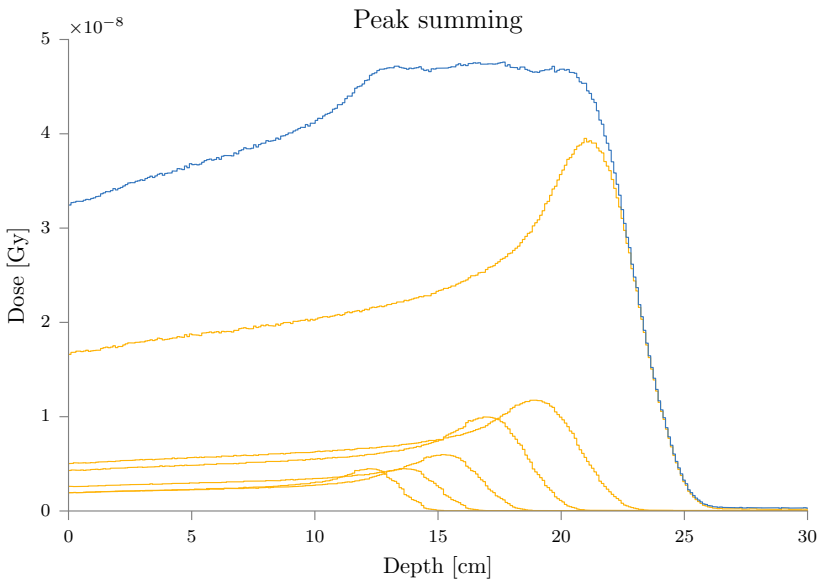


Figure 1.16: In order to irradiate over the tumor volume, multiple hadron energies are summed. This has the unfortunate side effect of an increased dose before the tumor, to healthy tissue. Plot produced with Gate/Geant4.

tion of the particles energy (Tommasino and Durante, 2015). Consensus over this effect has however not been reached, which introduces a downside: the RBE increase translates to an effective biological range extension. If the effect exists, and is not taken into account when creating treatment plans, there is some dose deposited after the supposed end of the range.

TSome research has been done on pion radiation (Goodman et al., 1990; Wisser, 2004) that concluded that the performance has no clear advantages nor disadvantages compared to protons or carbon ions. Even less research has been performed on antiprotons which is summarized by Amaldi and Kraft (2005) as having 'non-obvious advantages'. Both have one clear disadvantage: the production of these exotic particles is costlier compared to normal matter. Moreover, evaluating the RBE of such exotic species will be even more difficult than it already is for relatively ordinary protons and carbon ions. Up to 2014, about 120k patients were treated worldwide with protons, and about 15k worldwide with carbon ions, and no other species were used outside of research settings (Jermann, 2015). However, helium seems to reach a good optimum between fragmentability, reduced lateral scattering, and a potential increase of RBE around the BP. These benefits with respect to protons (Durante and Paganetti, 2016) lead to a renewed clinical interest in this particle species (Fuchs et al., 2015; Krämer et al., 2016).

1.4 Monte Carlo tools

The physics in particle physics experiments and particle therapy are described by a vast and complex body of models and data, each with their own assumptions and trade-offs. Naturally, physics *does* combine all the processes underlying these models. A single particle slowing down under the CSDA assumption is a relatively simple system, compared to the *Sjoelbak* physics referenced before. When the multiplicity of all possible states of a system explodes, and therefore (analytical) modeling complexity, MC tools provide a way out. Large experiments dedicate large efforts towards the tabulation of probabilities or cross-sections of all possible interactions, and a Monte Carlo (MC) simulation generates a particular history using random number generation and these probabilities (see fig. 1.17

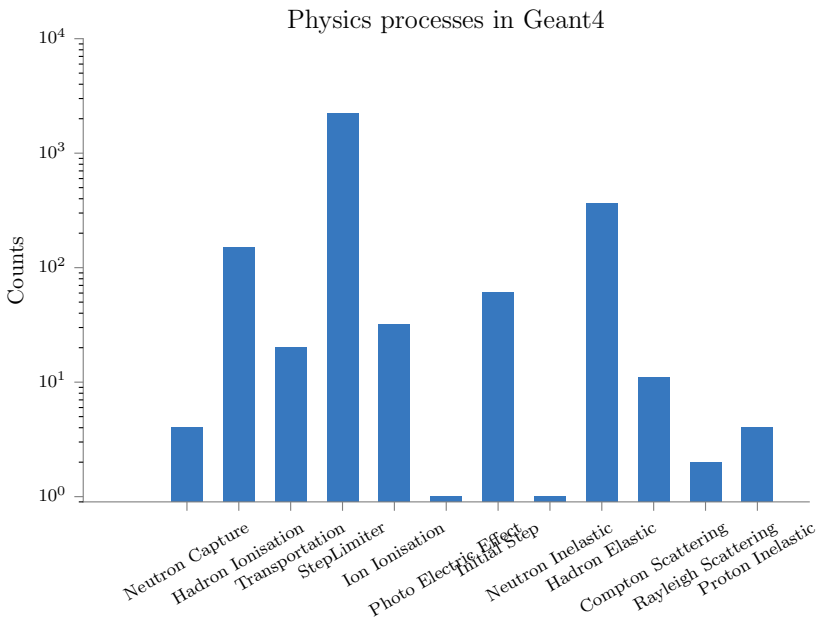


Figure 1.17: Models of processes used during the simulation by Geant4 of 10 proton trajectories through a box of water.

for an illustration). In that context, MC tools provide a way to perform experiments, without the need to put a patient at risk.

Some common MC packages in particle therapy are Geant4, EGSnrc, PENELOPE, MCNPX and Fluka, upon which other tools such as Gate (Sarrut et al., 2014), Topas or custom tools may be built. Most of them support various physics models (in Geant4 these are called physics-lists), each with different assumptions or based on different data, and with different tunable parameters. Various groups use such packages and each have their preferences. Detector designers will appreciate the ease of use of building geometry, while clinical users that verify treatment plans might care more about the presence of user friendly interfaces. Simulation speed, always a concern, is different between the packages, and can often be tuned by choosing certain physics models or limiting the physics to the part that one is interested in. The latter is often controlled with physics or production cuts: rather than removing parts of physics outright (for instance: in photon simulations usually only electromagnetic interactions are needed for correct estimates of dose), a threshold determines whether or not a particle is created or the energy that would have gone into that particle is simply registered as 'deposited'. People interested in dose calculation, might chose to save time by not generating and propagating secondaries for instance. When a so-called rare process is the interaction of interest, variance reduction methods can be employed: it reduces the variance for any given simulation with respect to the regular or so-called analog simulation. As a simulation runs, the output converges, which is to say: its variance reduces. A common measure to ensure sufficient convergence is to require that the uncertainty σ associated to the quantity of interest is not more than 2% of the signal. A variance reduction technique then translates into reaching the threshold faster, and therefore a gain with respect to analog MC. For photons, such methods are abundant, helped by their relatively simple physics.

Even though the goal of every MC package is to converge on the same physics, each package may produce different results, and also the choice of physics(list) and productions cuts will influence results. While such packages are continuously updated, adjusted, validated and compared (Dedes et al., 2014; Pinto et al., 2016; Robert et al., 2013), the user has, certainly in clinical applications,

the duty to know about these differences and know which packages or physics are suited for which application. Most packages have recommendations based on the purpose of the simulations.

1.5 Summary

At this time particle therapy is a well established modality for treating solid tumors that is seeing a rapidly expanding number of facilities worldwide. The theoretical benefits over X-ray treatment have been largely confirmed for pediatric cancer, and are currently being assessed for other cancers, as evidenced by an increasing interest. The biological effects of particles impinging on patients tissues would benefit from further study. There are currently no clinically established ways of observing the particles directly or indirectly during treatment, which stresses the importance of quality control and a solid understanding of the behavior of protons in patient tissues.

Chapter 2

Uncertainties

A challenge in EBRT in general and particle therapy in particular is developing our understanding of how radiation affects tissues, how radiation may be applied in the right places, and not in the wrong ones, and which side-effects are unavoidable or how they may be minimized. Robust and successful treatment and patient survival are the goals, but must be broken down into the multitude of factors influencing it if we are to improve the way we treat cancer. Figure 2.1 is a rough summary of the uncertainties encountered in *proton* therapy and categorizes their severity in descending order. The order of causes is different for different particles; for carbon therapy the uncertainty due to RBE would be higher for instance. In this section I will deconstruct the EBRT workflow and investigate the uncertainties in each part, and I will mention current and future strategies to handle those uncertainties.

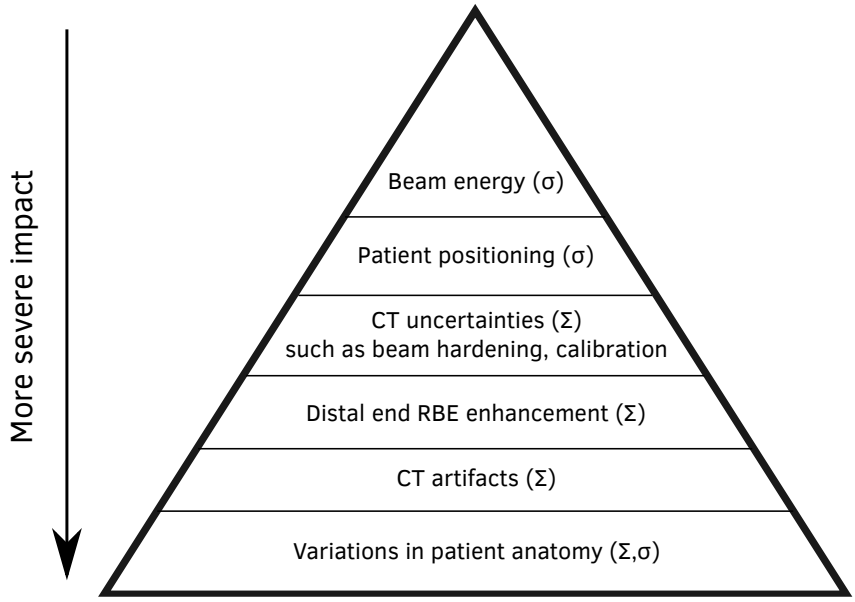


Figure 2.1: The Bermuda Triangle of Uncertainties, as presented by Tony Lomax at the Hadrontherapy School at CNAO, 2014. Σ refers to systematic uncertainties and σ to statistical effects.

2.1 Clinical workflow EBRT

Different types of tumors have different types of symptoms, and those will not be discussed here. Once symptoms are reason enough for a doctor to suspect the presence of a tumor, a computed to-

mography (CT) scan is one of the default diagnoses performed. If it is determined EBRT will be (one of) the treatment modalities, a planning CT is made. Based on this, clinicians use software to produce a treatment plan (TP) which is then applied to the patient. Depending on the progression of the treatment (side-effects may develop, other observables may indicate a deviation from the TP) the treatment may be altered, although this generally is an exception. At certain stages of the treatment, intermediate images may be made, usually Cone Beam CT (CBCT) images. Such images, and other diagnostics such as blood tests and physical exams, may lead to a change or interruption of treatment, although the latter is exceptional. Such modified treatments based on observations during treatment are categorized as Adaptive Radiotherapy (ART).

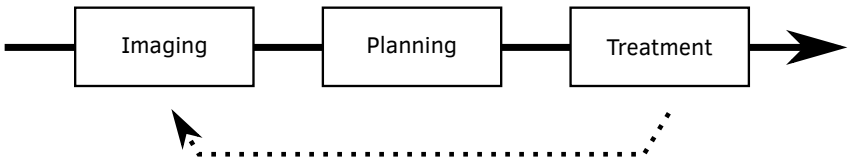


Figure 2.2: Clinical workflow in EBRT. In some cases, when clear and significant morphological changes in the patient have been established, a new image is made, upon which a new treatment plan is prepared and executed in the remaining treatment fraction.

The workflow for EBRT from a bird's eye view is show in fig. 2.2. For each component, various factors introduce uncertainties, the major ones will be elaborated upon in the remainder of this chapter.

2.1.1 Imaging

Various modalities exist for intra-body probing. Magnetic Resonance Imaging (MRI), Ultra Sound imaging (US), radioactive tracers visible using Positron Emission Tomography (PET) or Single Photon Emission Computed Tomography (SPECT) imaging exist. However, the oldest method is radiography, refined into Computed Tomography (CT), and has the added benefit that it probes using the same particles (photons) as almost all EBRT treatments (X-ray). Clinical practice is therefore, without exception, the creation of a high resolution high accuracy CT image, on which Treatment Planning Software (TPS) will compute the best treatment for the criteria entered by a clinician. A CT scanner, also referred to as computerized axial tomography (CAT) scanners, shoot a bundle of photons, usually in the 100 keV range, through a patient and behind the patient measure the transmission, from which absorption (or attenuation) can be obtained through Beer-Lamberts law:

$$\mathrm{I} = \mathrm{I}_0 \mathrm{e}^{-\mu \mathrm{x}} \tag{2.1}$$

where I is the transmission, I_0 the emission, μ the mean linear attenuation coefficient (over the beam energy spectrum) of the target and x the traversed path in the target. A side effect of X-rays in matter is beam hardening: low energy photons are absorbed by the patient, effectively increasing the mean beam energy as the beam traverses matter. This can be corrected by pre-hardening the beam with metal filters or with post-processing on the measured data, each with their own added uncertainty.

Since μ is material-dependent, one single acquisition only shows the cumulative effect of the various materials that happened to have been in the photons path. Remember that obtaining a three dimensional image of the patient, thus of the different tissues and materials is the goal. A CT scan is then a sequence of radiographs, under various angles θ about the patient, at various positions along the patient. The radiograph data is usually organized so that at each position z along the patient, for each θ a 1D line radiograph, over the other spatial dimension of the camera t remains, facilitating a slice-wise or helicoidal reconstruction (see fig. 2.3). For each slice, at position z, the 1D radiographs can be composed into a 2D image of axis θ and t. These images are called *sinograms*, and they

are the starting point for most reconstruction algorithms. The end goal is 2D maps of attenuation coefficients, at each z a transverse slice in coordinates x and y.

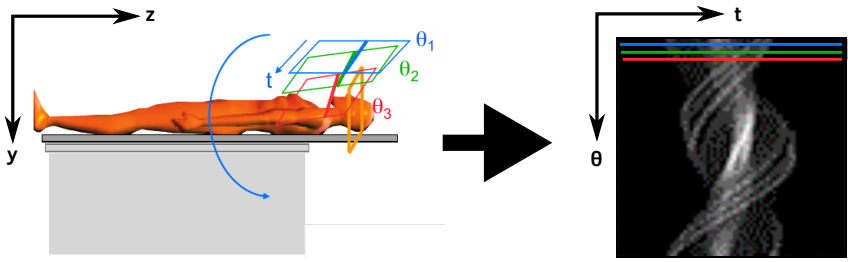


Figure 2.3: Data from a series of CT acquisitions is reordered into sinograms. Illustration adapted from material by Irène Buvat.

The sinograms have values at each pixel, and various reconstruction methods map this back into values in the (real) x and y plane. Some of these methods include Iterative Reconstruction Algorithm, Fan-Beam Reconstruction and perhaps most popular the (Filtered) Back Projection Algorithm (FBP). Since the number of angles is limited, the number of back projections is as well, and the image may appear streaked, and the filter in FBP takes care of that. Once each slice is reconstructed, they may be put together for a 3D image.

Since the voxel sizes are fixed, μ in eq. 2.1 may be obtained for each voxel, and thereby the material present at that location. μ conveys information of both density and material type, and for convenience these values are rescaled to Hounsfield units:

$$HU_x = \frac{\mu_x - \mu_{water}}{\mu_{water} - \mu_{air}} \times 1000 \tag{2.2}$$

The Hounsfield Unit (HU) or CT number is derived from the linear attenuation coefficient μ of photons in matter. The HU is defined in eq. 2.2, such that $\mu_{water} = 0$ and $\mu_{air} = -1000$, bringing most clinical materials into a [-1000,1000] window (dense bones or metallic implants may have a significantly higher HU-index). It is a convenient scale for CT imagery, because the HU is, modulus the rescaling, describing the direct observable μ, a material property that is also used when photon treatment planning is performed. For particle therapy the relevant quantity is not μ but the stopping power. Since images on which treatment planning is being done are CTs anyway, these HU indices must be converted, to particle stopping power.

2.1.2 Planning

Now that patient data is obtained, a treatment can be devised. Clinicians start by drawing contours on the CT image: delineations of particular biological structures such as the tumor, certain organs, in particular the radio-sensitive organs (AORs) and the patient outline (the surface of the skin). Usually, contouring is done slice by slice, and a 3D surface is interpolated. Contouring may be aided by software that auto-contours easily detectable features such as the patient outline. A final review, with (small) alterations if necessary, is performed manually. Occasionally data from other sources such as PET or MRI scans are overlaid to aid the contouring process. An important note here is that in particular for the tumor margins are added. The visible tumor contour is called the Gross Tumor Volume, or GTV. Because imaging has resolution limitations and in particular contrast limitations, the tumor may be in fact larger than what can be seen. A margin of up to a centimeter wide is added, creating the Clinical Target Volume (CTV), which also helps ensure that cells that diffused away from the tumor are treated. Then, to deal with technical limitations of the treatment delivery, patient positioning and motion, and uncertainties in planning, a second margin is added to arrive at the PTV or Planning Target Volume.

Once the contours are in place, a prescription is defined according to current protocols and clinicians' input. For some tissues a minimum or maximum mean dose is defined. Certain organs, so called parallel organs, can tolerate a higher dose as long as this high dose is confined to a certain fraction of the organ so that the remainder survives and can (partially or fully) compensate for the damages caused in the high dose part. Conversely, serial organs do not exhibit such intra-organ redundancy and if such organs are radio-sensitive, they must be kept under their limit over their whole volume. In case of severe radio-sensitivity, of which nerves are the canonical example, the organ may be classified as an AOR.

Certain field or beam directions may be preferred, and for particle treatment the presence of a gantry or number of fixed beams affect the possibilities for the TP. Another factor is the number of fractions: the number of days the patient will return for treatment. A typical total dose to the PTV is 60 Gy, and often this is spread over 30 days so that each day a 2 Gy fraction is delivered. These constraints serve as input for a dose optimization procedure: an algorithm that will find a particular set of fields with field contours, and energy layers in the case of particle treatments, that best match the prescription. Such TPs are analyzed by clinicians using dose contours overlaid on the CT data or Dose Volume Histograms (DVHs) such as in fig. 2.4. Of course, the optimization procedure itself is a source of uncertainty, there are better and worse, and faster and slower programs.

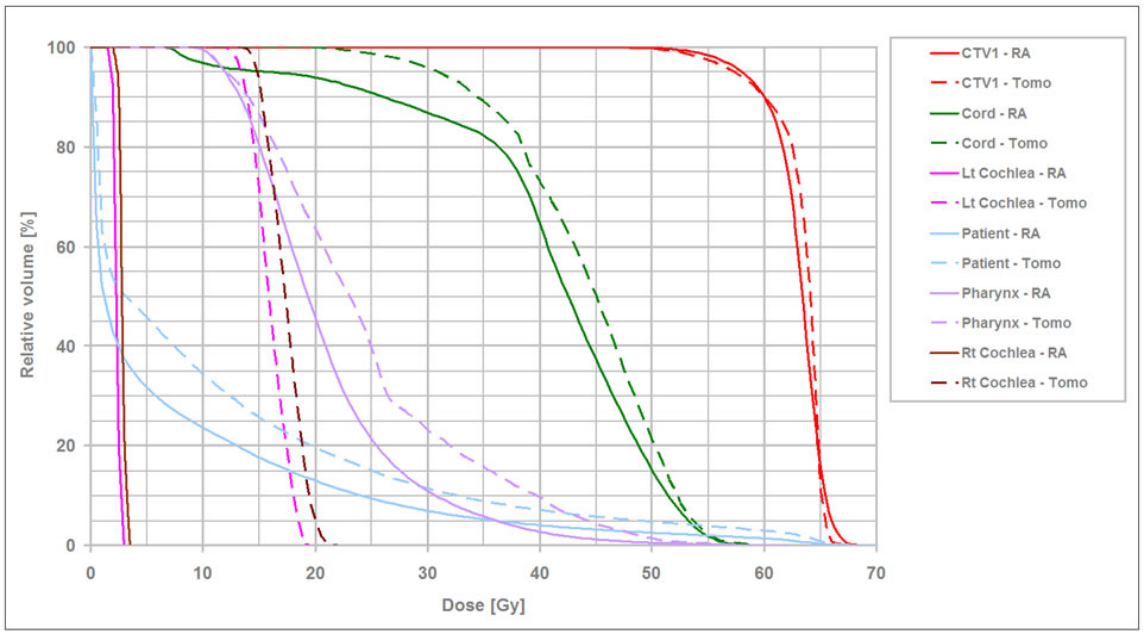


Figure 2.4: Dose Volume Histograms show for selected contours the fraction of the volume covered as function of the dose. The PTV should reach at minimum its prescribed dose in nearly all of its volume, while for organs at risk such as the Cochlea, the maximum allowed dose must not be exceeded in all of its volume. Other organs can tolerate a bit more, and in such a diagram it can be seen how much (extra) dose these organs receive. Image CC-BY licensed by Wikipedia Authors.

Besides the CT image, contours and prescriptions, the TPS requires an additional bit of information: the beam description. Be it an X-ray machine or a proton synchro-cyclotron with a scanning nozzle: the precise possibilities and tolerances of the system must be known to the TPS for it to generate the best possible TP. For particle treatments, it is usually the nozzle and available energies and their tolerances that are known, including scanning speeds and energy switching time. Not only must this model be available to the TPS, they must be well established and monitored as time goes by. The beam delivery system requires regular and well specified Quality Assurance (QA) testing: not only are HU-units tested with a waterbox, the particle's energy and positions must be measured, as well as the

tolerances. The main job of a clinical physicist is to monitor and chart these properties, so that the system conforms to the beam description that the TPS has, ensuring as few differences as possible between treatment plan and treatment delivery.

2.1.3 Treatment

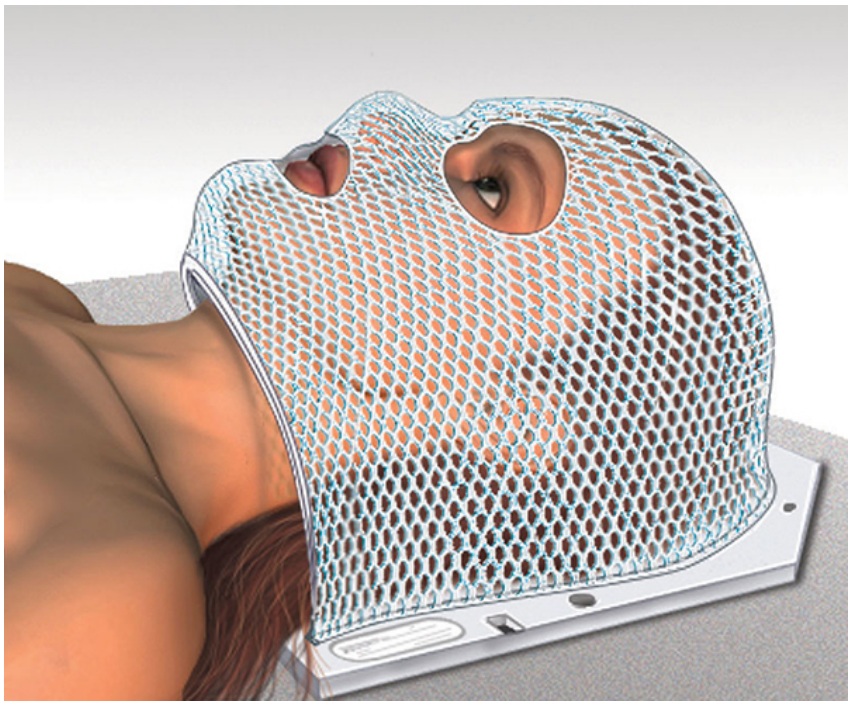


Figure 2.5: Radiation Therapy Head Mask. Patients getting radiation to the head may need a mask. The mask helps keep the head from moving so that the patient is in the exact same position for each treatment. Image released in the public domain by the United States government at `cancer.gov`.

When a patient is treated, the setup, which includes the beam *and* the patient, must conform to the models as were present in the TP. Clinical physicists make sure the machine is aligned, within tolerance, and so on, and are responsible for the setup of the patient. In particle therapy, molds are often used to fix the body and in particular the body part affected by the treatment in position. For a head and neck treatment, a mask as seen in fig. 2.5 is common. While not very comfortable, with particle therapy, precise alignment has become much more important due to the sharper dose gradients compared to X-ray treatments. Typical tolerances are 3 mm translations and 3 degree rotations. The TPS takes such tolerances into account, of course at the cost of a lack of accuracy of the treatment. For some regions of the body, temporal changes greatly affect the type of margins built-in to the TP. 4D planning is an area in development, where motion tracking or compensation is incorporated in the TP and helps reduce these margins. For instance: a lung tumor may be imaged and more welldefined when the patient has exhaled, but of course when inhaling the tumor may be centimeters away from the imaged position. The simplest form of 4D planning is gating: simply switching the beam on and off synchronous to the breathing motion of the patient. While conceptually simple, interplay with the time dependence of beam scanning is far from trivial.

2.2 Commissioning, Quality Assurance

As is clear now, the act of treating patients relies on a large set of complex machines operating according to known or defined specifications. The TPS has a beam description, reads the patient planning CT and associated contours and prescription, and produces a list of lateral beam positions as function of energy as output. But also the position of the patient is controlled during treatment. When a new particle therapy center is built, the commissioning phase is a full systems check of all systems involved in executing the treatment. The commissioning may include:

- commissioning of delivery system
- verification of dose distributions
- mechanical alignment (gantry, couch, snout, ...)
- imaging equipment
- TPS
- collection of beam data
- verification of calculation algorithm
- simulation equipment
- Hounsfield to stopping power calibration (CT)
- milling machine (to create patient specific boluses)
- apertures and range compensators
- immobilization devices (such as the head mask)
- beam perturbation

Once in operation, regular checks are performed to ensure systems remain within their defined tolerances. This is called Quality Assurance (QA). Through continuous research and development tolerances are improved every iteration of a (sub) system, but they remain inherently present. The way to deal with tolerances, or uncertainty, is to add various margins to the contours in the image data. A larger volume than required is irradiated, to ensure any movement, misalignment, variation or patient morphological change does not lead to the GTV being outside of the irradiated volume. This comes at the cost of irradiating healthy tissues that surround the GTV (fig. 2.6).

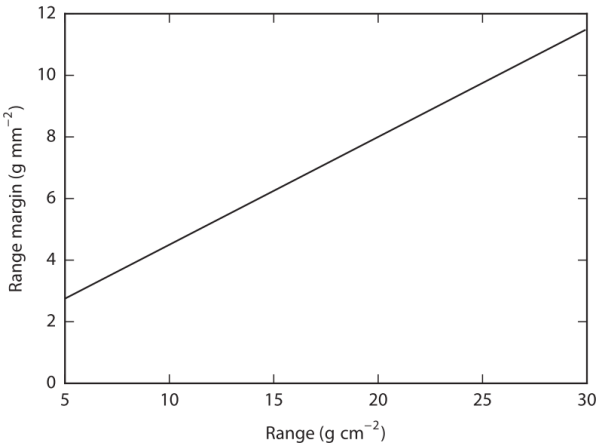


Figure 2.6: The specified uncertainty on a proton pencil beam range at the Massachusetts General Hospital in 2015 was a 3.5% + 1mm as a function of the range in water. The uncertainty is dealt with by adding margins to tumor contours resulting in the PTV. Commissioning establishes this curve, QA ensures its correctness in subsequent treatments. Source: Verburg (2015).

2.3 Physics and Biology

As said in the introduction to this chapter, EBRT is about putting the right dose in the right place. *Dose* has been introduced in the introductory chapter as a way to quantify damage, and in this section the uncertainties that arise from the way dose is quantified will be discussed.

2.3.1 RBE

Since particle therapy inherits many of its clinical practices from conventional radiotherapy, some out of necessity, some out of habit, it seems fair to say that particle treatments are decidedly photon oriented in their construction. Planning image data is always obtained with a CT, and the concept of RBE relates dose induced with non-photonic modalities (effective or biological dose, eq. 1.6 on page 11) to dose induced by photons (physical dose, eq. 1.5 on page 11). In fact, even the physical dose photon dose equality is being questioned Liang et al. (2017). Briefly recalling the RBE section of the previous chapter, dose is a measure of damage, and based on the transfer of energy (LET) from the primary particle to the patients tissues. Using cell survival studies, the damage relative to the photon dose is calibrated by adding a factor for the RBE.

The computation of the RBE is performed by comparing the cell survival for a given LET, which in the case of proton radiation means the proton energy, with a reference, X-rays. If 1.0 Gy was required to kill 80% of the cells using X-rays, and 0.8 Gy using 140 MeV protons, then the RBE for 140 MeV protons, for that cell type, is $\frac{1}{0.8} = 1.25$. The RBE for proton therapy is set at a constant 1.1 in almost all clinics worldwide, and this 1.1 is the average RBE in the mid-SOBP region of various studies, with various in-vivo measurements, with various proton energies (Paganetti et al., 2002). There are two problems with this. Based on the same data presented in those studies, we know the variation is systematic: different cells have different radiosensitivities, and averaging their RBEs does not make sense from a biological point of view. Secondly, it is widely accepted that proton energies have very differents LETs (fig. 2.8). That means that the RBE for x MeV protons is not related to the RBE for y MeV protons. Averaging makes no sense from a physics point of view. Note that the proton loses energy until it is stopped, which means the LET depends on the position in the patient. Moreover, since the SOBP is made up of various primary proton energies, any region except the most distal will undergo dose deposition by protons of a range of energies, therefore a range of LETs. The original point of inventing the LET was to remove the statistical fluctuations of the discrete energy transfers that a particle traversing matter experiences. LET fundamentally varies with the energy and therefore range of the particle. As far back as Rossi (1959) assuming LET constancy was known to be incorrect. The evidence for RBE variation notwithstanding, for protons the variation of LET with energy has been considered weak enough to ignore it (Coutrakon et al., 1997) and the RBE remains at a constant 1.1.

For carbon, the change in LET along the path is too much to ignore (see fig. 2.7), and here models more complex than RBE=1.1 are in clinical use to compute the biological dose. As mentioned in the previous chapter, two main families of models may be distinguished: various version of the Local Effect Model (LEM) in European particle therapy centers and semi-empirical models based on the Zaider–Rossi approach in Japanese centers, sometimes called the NIRS approach (Fossati et al., 2012). More recently, the mMKM (modified microdosimetric kinetic model) and NanOx (NANodosimetry and OXydative stress) models were developed (Cunha et al., 2017), seeking to further increase the predictive power and correct estimation of the actual biological damage. In Molinelli et al. (2016) NIRS and LEM are compared, and since the predicted biological doses are different and both models are clinically applied, we must conclude that at this time there is no agreement on what the biological dose for carbons is. An updated version of LEM predicts a significantly different RBE at the very end of the Bragg Peak for protons, resulting in an effective range shift of a few millimeters (see fig. 3.16 on page 53), which is why people have been calling for an update to the proton RBE as

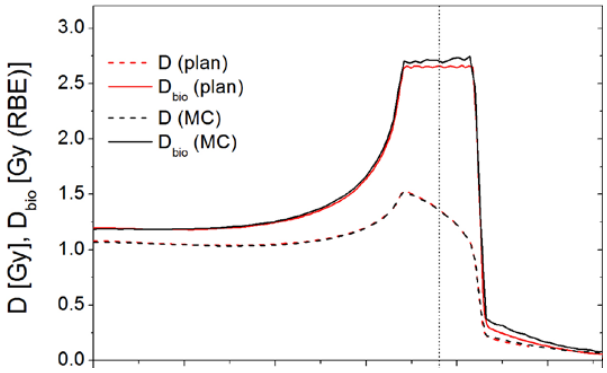


Figure 2.7: Physical (dashed) and biological dose (solid) of a carbon beam. Source: Inaniwa and Kanematsu (2014).

well (Paganetti et al., 2002). Figure 2.8 aggregates many RBE studies in one graph. It is clear that the RBE is a complex function of many parameters, such as radiation quality, dose, dose rate, biological endpoint, cell type and cell cycle stage, and that one constant value may be too much of a simplification. After all, more accuracy and precision means that margins can be eventually lowered and improve the quality of treatment.

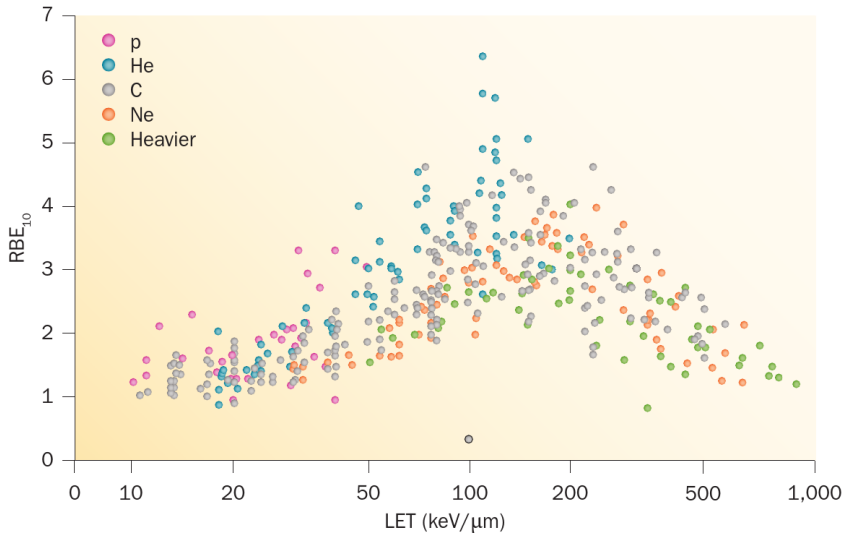


Figure 2.8: RBE compared with LET from published experiments on in vitro cell lines. Different colours indicate different ions, from protons to heavy ions. Source: Loeffler and Durante (2013)

2.3.2 Conversion to Stopping Power

When it comes to Hounsfield to stopping power conversion, every particle center builds a conversion curve as part of commissioning. The Hounsfield unit is centered on water, thus the stopping power in water is the elementary ingredient in the commissioning of a new center. Depth-dose profiles for all spots and all beam energy levels are meticulously measured with a peak-finder, and fed into the TPS' model of a pencil beam, the dose kernel. Then, for certain materials a Water Equivalent Path Length (WEPL) is measured by inserting slabs of various thickness of the material in the waterbox. The shift in BP is then a measure of the WEPL, that can be translated into a Relative Stopping Power (RSP) or Stopping Power Ratio (SPR). In figure 2.9 such curves are shown, and using these the path

length of particles can be computed in all voxels of a CT image, and so a correct treatment produced. In clinical practice this procedure is thus quite experimental.

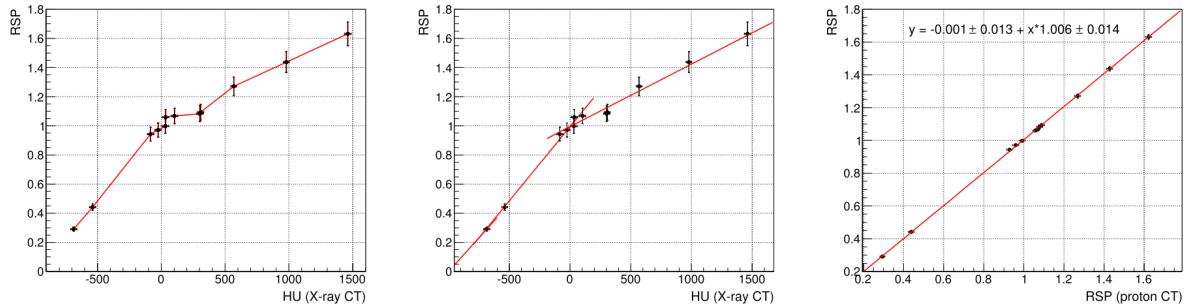


Figure 2.9: RSP-HU calibration curves for X-ray and proton CT scanner simulations. For the X-ray CT, linear interpolation (left) and a two linear functions fit (middle) are shown. For comparison, on the right a proton CT conversion curve where a single linear function suffices. The choice of interpolation influences the precise conversion of soft tissues the most. Source: Arbor et al. (2015)

The use of the dose kernel, where a relatively simple analytic model about beams is calibrated with measurements, together with the idea that different tissues introduce nothing more than a simple shift, is a direct carbon copy of X-ray procedures. To the physicist's eye it may be a bit painful to see the liberties with reality being taken, and while they may seem appropriate for X-ray treatment, to some it might seem a miracle particle treatments already work as well as they do. Chapter 26 of (Mayles et al., 2008) discusses the various dose kernels in use, the so-called analytic models, while chapter 28 of the same book summarizes work being done on the use of Monte Carlo (MC) tools in predicting the dose and its distribution. MC simulation may be performed to validate analytic models used in TPSes (Paganetti, 2012; Verburg et al., 2015). From a physics perspective MC might seem the most correct way to study ionization and dose deposition, because it models the fundamental physics down to each individual interaction. However, as of today this is still too time consuming to execute for each daily clinical practice.

Analytical methods to predict stopping powers exist and assume that, since electronic stopping dominates the particle's deceleration in all but the last millimeter, the relative electron density (RED) is usually taken as equal to the RSP, certainly in TPS. However, this is an oversimplification as a correct definition would involve a factor involving the logarithm of the ionization potential I, see eq. 9 in (Yang et al., 2010), which is almost but not quite the same for different materials. The factor comes from Bethe's equation in 1.3, so an analytical estimation of the RSP can be produced using the electron density and ionization potential of a material, as was presented in (Schneider et al., 1996). Clinical practice tends to favor conservative margins because many uncertainties are not well known, while good modeling and better measurement may reduce uncertainties and thereby margins.

General CT improvements are being developed, such as dual energy CT. In dual energy CT the radiograph is acquired at two (and possibly even more) photon energies, leading to two attenuation coefficients per voxel. Since μ is a convolution of density and material, and will be different for different photon energies in different materials, having this second data point leads to more confidence that the correct material is inferred (Hudobivnik et al., 2016; Vilches-Freixas et al., 2016).

One aspect of X-ray radiographs has not been discussed yet, and that is the quite terrible contrast with soft tissues of the human body. Moreover, the Hounsfield to stopping power calibration is most uncertain here. Dual energy can help, but the MRI modality provides excellent contrast of soft tissues. Work is being done to improve the CT with data from an MRI image, possibly eventually replacing CT altogether (Lagendijk et al., 2016). Moreover, with adapted equipment, MRI imagery can be acquired during X-ray treatments, providing an opportunity for real-time beam re-positioning.

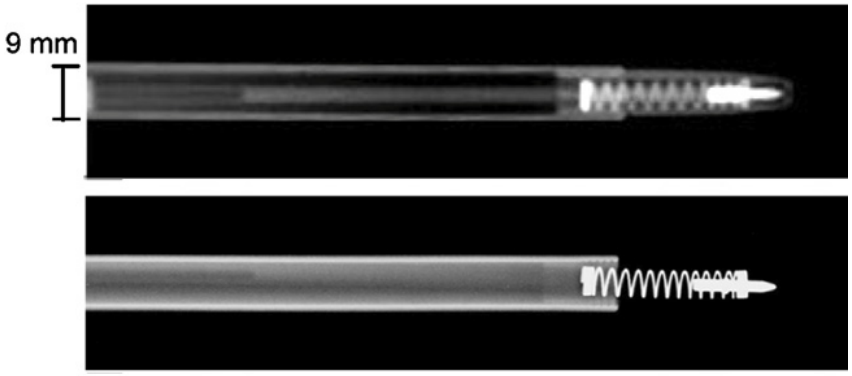


Figure 2.10: The top image shows a proton beam image and the bottom image shows a standard X-ray scan. Note the missing features in the X-ray image and the difference in sharpness. Because of the lateral scattering, the edges in the proton images are not as sharp as with the X-ray. Image courtesy of Ryu et al. (2008).

Protons exhibit a greater soft-tissue contrast compared to photons, as seen in fig. 2.10. Proton CT was originally proposed because of this (Hanson et al., 1981; Koehler, 1968; Wilson, 1946); after all, most of the human body is made up of soft tissue. Because a proton beam is required and such facilities have only become available outside of research contexts in the last one or two decades, the idea did not gain much traction until then. Since the conversion between the modalities can be avoided by making radiographs with the same particle as the treatment, proton therapy lead to a return of interest in proton CT (Amaldi and Braccini, 2011; Benedikt and Wrulich, 2011; Rietzel et al., 2007; Rossi, 2011; Schardt and Elsässer, 2010). With proton CT, on top of the enhanced soft-tissue contrast, the stopping power is directly observed, instead of μ. After detection comes reconstruction, and colleagues have been working on reconstruction methods tuned for proton CT (Arbor et al., 2015). In Quiñones et al. (2016), proton energy-loss CT and proton count attenuation CT were compared, as well as the effects of various scatter corrections investigated, which is a significant difference with X-ray CT. Proton CT performs worse in the transverse plane and better in the longitudinal direction, so combining photon and proton data into a single CT is currently being investigated (Krah et al., 2016).

2.4 Treatment delivery and planning

Now that we have seen what dose is, we must know where to deposit it. It is a good time to note that with respect to X-rays particle treatment planning has an extra degree of freedom: the particle's energy, which governs its depth. The type of beam determines much of which type of planning is possible. Passive beams require a very different type of planning, and has as output not a list of spots with their positions and energies but only a list of energies and a patient specific bolus. The industry seems to have decided that the higher (commissioning) cost of active beams is worth the clinical benefit of less scatter and neutron dose due to fewer absorbing materials in the beam. So, here I will discuss treatment planning for scanned beams, which is entirely software-based.

Depending on the gantry and number of beams, various types of treatment optimization exist:

Single Field

- Single Field Optimization (SFO) or Single Field Uniform Dose (SFUD) planning.
  Per position in the transverse plane, the particle energies corresponding to the proximal and distal edges of the tumor are computed. Per position, between these energies, a SOBP is created resulting in a uniform dose distribution that aligns with the dimensions of the PTV.

Multiple Fields

- Multi Field Optimization (MFO)
  Each beam delivers a certain fraction of the total dose and each is optimized as in SFO. For instance, in a two-field plan each beam could be an SFO delivering a uniform 1 Gy dose across the PTV.
- Intensity Modulated Proton Therapy (IMPT)
  In IMPT each beam spot delivers an independently computed dose. No uniformity within a beam is enforced, and the ratio between the beams is not enforced either. All depends on the choices made by the optimization algorithm. IMPT is considered to provide the most conformal treatment plans to date (Durante and Paganetti, 2016; Fossati et al., 2012), at the cost of being less robust. Sometimes this is called unbalanced optimization.

In broad terms, robustness is improved when variations are suppressed. The latter can be achieved by knowing the variations better (e.g. better imaging, better motion tracking, better understanding of the RBE) or by reducing the interplay between multiple fields by reducing their number, accepting less conformal planning by planning with wider margins. Also the choice for the spot number introduces a consequent effect on robustness and precision. If you want uniform dose, and your means of doing so is to place dots, with arbitrary precision and size, and the delivery time of a single spot is constant, it follows that smaller dots equals more uniform and conformal dose distribution, and bigger dots faster treatment and a higher robustness to variations. Faster treatment means more patients treated in a day, which can be translated to a lower cost of therapy. In addition, patients are usually stabilized, which is not comfortable for longer periods of time. So, a choice like this depends on the clinical tolerances for a particular patient. As will be demonstrated in a later chapter, for pediatric cases the extra precision of more and smaller spots is worth the increase of treatment time. To enable smaller spots, nozzle designs are improved so that a finer beam position grid is possible. In addition, non-isocentric planning is investigated, where the patient is moved close to the nozzle (Grevillot et al., 2015).

In multifield treatments, uncertainties can have unexpected interplay effect. In opposing beam scenarios, in-axis set-up errors do not matter, while perpendicular misalignment does. Perpendicular beams will be sensitive to error in either direction. This means that there is a choice in treatment planning: trade-off between robustness and conformality (McGowan et al., 2015). Clinical practice at HIT in Heidelberg is per-field homogeneous dose (single field optimization), so that if there is an error, the second field can not be irradiated while still having a (more or less) known uniform dose distribution from which a new plan may resume. At CNAO prostate is treated the same way. CNAO treats all other cases with IMPT, and when using protons, a single field is preferred whenever possible. European and US centers tend to use homogeneous multifield planning while in Japan inhomogeneous fields are the norm. It appears there is no consensus on the best trade-off.

A treatment is broken up into sessions or fractions, because it was observed that this improves the ability of healthy tissue to recover (Mayles et al., 2008, Chapter 9). There are practical and biological reasons for this. Delivering 60 Gy at once would require a patient to lay still, fixed to the treatment table, for at least half an hour. Moreover, healthy and cancerous tissues have different radio-sensitivities. Tumors' main characteristic is that they are deregulation tissue, which brings one advantage to the fore: they are more radio-sensitive because their deregulation makes recovery from damage less likely. Fractionation is then a way to give healthy tissue time to recover, while tumors do not recover (as much) in the time between fractions. For different types of tumors, different fractionation regimes are customary, and they are related to the α and β in the survival studies of the previous chapter, and sometimes α/β is referred to as the fractionation factor (Mayles et al., 2008, Chapter 55.2). Hypofractionation, fewer higher dose fractions, is performed for lower α / β ratios. In certain cases, even single fraction treatments have been performed. Karube et al. (2016) presents the results of single fraction treatments for the elderly, whom may be inconvenienced more than usual

by frequent visits to the clinic. The life expectancy of older patients means the effects of complications in healthy tissue may never be felt. Closely related to fractionation is the dose rate: how fast the treatment is delivered. The 'saturation' effect that α and β imply, are valid on shorter timescales as well. Different dose rates can be translated to different RBEs, depending on the tissue (Wozny et al., 2016).

Some other innovations in clinical treatment delivery are the use of heat (Yarmolenko et al., 2011) and non-uniform dose planning. Sometimes called dose painting, certain regions of the PTV may receive a boost, which is to say an increased dose to ensure destruction of possible stem cells responsible for seeding the tumor (Shi et al., 2014). Mixed modality boosted treatments are also considered, where a normal X-ray treatment is supplemented by a proton boost (Hopfgartner, 2008). Radiosensitive nanoparticles that amplify the ionization are a topic of much interest (Haume et al., 2016).

2.5 Adaptive Treatment

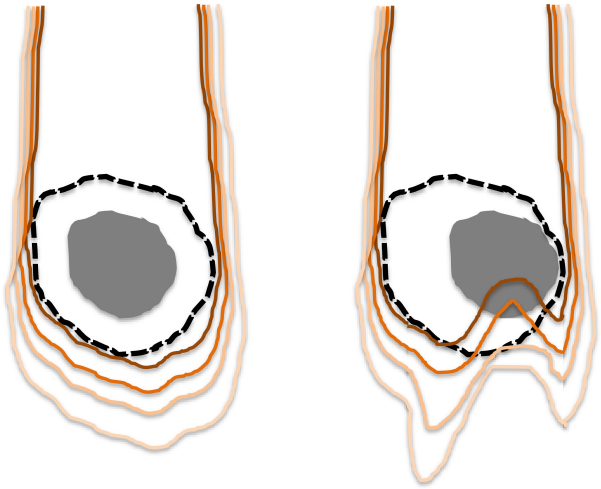


Figure 2.11: Schematic view of the effect of organ displacement. Iso-dose lines are drawn to demonstrate the effect. Margins are added to mitigate the effect at the cost of extra dose to healthy tissue. Spiky extrusions in iso-dose lines are sometimes referred to as dosefingers. Image courtesy of Paganetti (2016).

Even though many kinds of deviations are known, described, measured and taken into account, in regular treatments nothing is actually known: QA is how margins are established, not the measured position of the particular patient currently being treated. The dotted line in figure 2.2 on page 22 can, for a large majority of treatments, be left out. Relatively new are attempts to monitor certain uncertainties during treatment instead of fixing their state and adding margins. Looking back at figure 2.1 (page 21), patient change is the major component in the uncertainties of particle therapy, see figure 2.11.

Temporal changes can be distinguished based on their timescales: day-to-day changes are called inter-fractional change and changes during a treatment session is intra-fractional change. The various motions of the body are intra-fractional changes, while setup error and organ filling (bladder, bowels) change the treatment between fractions. Organ filling can be dealt with by requiring the patient not to eat a number of hours before the treatment, or to require the patient to urinate right before the treatment. Instead of forcing the state known to the TPS onto the patient, which may be difficult for certain patients, in Adaptive Radiotherapy (ART) a more flexible approach is taken. For instance, a tumor near the bladder may be dealt with by modifying the planning CT into various

stages of organ filling: the planning CT happened to be taken at 50% filling then generates CTs at 0, 25, 75 and 100% filled bladders. For each, the same prescription is used to generate a TP and every day a "plan of the day" is chosen based on the present filling seen on a CBCT made right before treatment. It makes the treatment more accurate and allows the patient to be more comfortable, at the cost of having to make a daily assessment of which version of the TP to irradiate. In more advanced ART, the CBCT itself may be used to obtain a "PTV of the day" (Burridge et al., 2006). Naturally, the software performing the CT-CBCT registration and image deformations to obtain the image at various fillings introduce their own uncertainties. Also, the treatment itself is not delivered instantly, and therefore the time effects may be different from region to region (fig. 2.12). In the concept of *theranostics*, the nanoparticles mentioned in an earlier section are used to both enhance localized dose delivery and motion tracking: the particles will emit electrons when irradiated (Haume et al., 2016; Retif et al., 2015). If the beamline has a combined treatment and imaging nozzle, this enables therapy and diagnostics simultaneously.

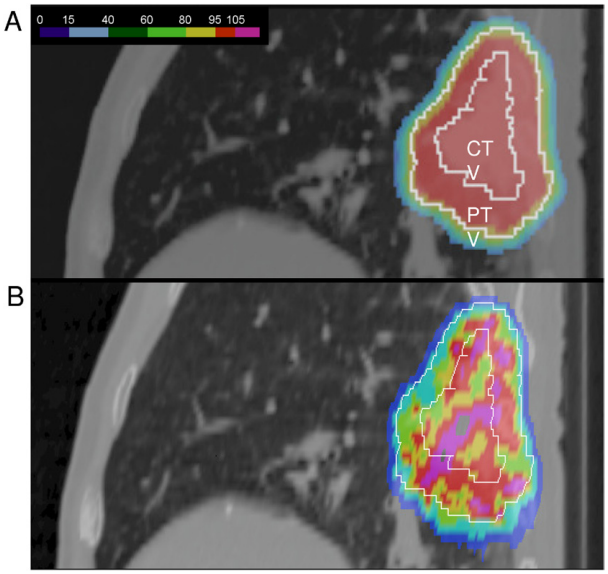


Figure 2.12: The interplay between pencil beam scanning and organ motion jeopardizes the dose distribution. (A) A lung tumor treated with beam scanning assuming no motion. A uniform dose is delivered to the clinical (CTV) and planning (PTV) target volumes. (B) the actual dose distribution caused by patient's breathing. Image courtesy of Prof Christoph Bert, University of Erlangen, Germany.

2.6 Detectors in particle therapy

Currently, the PET scanner is the only tool used for clinical particle treatment verification, in a limited number of particle therapy centers. The PET scanner was originally designed to image radionuclide absorption. β^+-emitting isotopes are coupled to a molecule that is absorbed in certain structures and would make the location of absorption show up in the PET scan. The emitted β^+ recombine with an electron, and emit two back-to-back photons with a signature 511 keV. Cameras must be able to detect the coincidence to be certain the photon originated from positron annihilation and therefore must be opposite each other. Having a full ring, or two opposing ring-parts, facilitates recording the photons, leading to the characteristic shape of PET scanners.

Particle therapy itself induces production of isotopes, so some activation of the patient's tissues are a natural byproduct of the treatment. It therefore makes sense to utilize this by scanning the

patient after treatment: a direct signal of the primaries' paths can be imaged. The activated target nuclei have a certain relaxation time, which means there is a time window between treatment and PET imaging: after one or two half-lives have passed, no useful signal may be expected anymore. The most common sources of β-emitters are ^{11}C, ^{15}O, ^{10}C, with half-lives of 20 min, 2 min and 19 s respectively (Kraft, 2000). In addition, biological processes displace the excited nuclei, for instance by blood flow or cellular metabolism. This 'washout' effect together with the half-lives limit the imaging window to between 5 and 20 minutes after treatment. Usually the PET scanner is in another room, or even another wing, which leads to a rather impractical workflow with limited results. If the patient cannot go the PET scanner, the PET scanner must go to the patient: in-room PET scanning is a way some particle treatment centers minimize the large time window between treatment and imaging. Integrated PET systems in the particle delivery gantry enable PET verification immediately after treatment, without requiring setup position and related errors and loss of time (Enghardt et al., 2004).

2.7 Summary and outlook

Currently dealing with uncertainties prevents the full potential of particle treatment from being realized. Exploiting the great precision of particle treatment is a balance between robustness and tissue sparing. The main challenge is that dose is not observable, and certainly not in real-time during clinical treatments. However, a particle treatment can be regarded as a fundamental physics experiment, where by good modeling of the interactions, the target, the beam, a detector, and secondaries may be used to deduce relevant quantities as the treatment has been or is being conducted. This way of gathering treatment feedback enabled by simulation is crucial for comparisons between expected and measured signals. Proton CT may help in removing conversion errors from the treatment planning workflow. Adaptive Radio Therapy seeks to combine measurements with improved treatment planning, and new methods such a theranostics seek to expand the increased precision by ensuring the dose is localized and tracked during treatment. Because particle treatments are less robust with respect to X-ray treatment, monitoring and eventually control are pertinent to reducing margins while improving the robustness at the same time.

An evolved notion of ART is IGRT, Image Guided Radio Therapy. IGRT takes it a step further: imaging, of some sort, could be normal procedure for any treatment. This requires imaging devices in the treatment room. Portal Imaging (van Elmpt et al., 2008) or EPID dosimetry (Rozendaal et al., 2015) are implementations where the X-ray absorption is measured during treatment, and thereby some information of where dose was deposited is reconstructed. For particle treatments, in-beam PET (Dendooven et al., 2015) is a way to image the activation of tissue during or immediately after treatment. While fragmentation is usually seen as a disadvantage, Kraft (2000) describes it in the case of carbon ions as a potential advantage: typically β^+ emitting ^{10}C and ^{11}C are produced that could be used for direct verification with a PET scanner. Online MRI guidance is seeing a lot of interest in the X-ray therapy community (Lagendijk et al., 2016), and some in silico studies demonstrate it is feasible in particle therapy as well (Hartman et al., 2015). Early stage work is being done on neutron-imaging, a natural byproduct of particle treatment and therefore no additional dose in imparted on the patient (Valle et al., 2016). Tracking secondary protons (Henriquet et al., 2012) in heavy ion therapy is under investigation. The aforementioned integration of PET in the delivery gantry led people to speculate that PET imaging *during* treatment might be feasible (Parodi et al., 2005), and a few years ago the first in-beam PET acquisition during a full beam was presented (Sportelli et al., 2013).

In the next chapter, another natural byproduct of particle treatment that may be imaged will be discussed in detail: Prompt Gamma rays.

Chapter 3

Prompt Gamma detection

3.1 Prompt Gamma physics

This chapter will detail the state of clinical Prompt Gamma (PG) applications. As mentioned in chapter 1, PGs are photons released in nuclear interactions between the incoming and secondary particles and nuclei in the target. Figure 3.1 shows how (physical) dose and PG production correlate, for a single spot in a simulated clinical proton treatment. The correlation is evident. Perhaps the most significant difference is the end of the signal: if we define the Fall-Off Position (FOP) as $\frac{\text{max - baseline}}{2}$ (see fig. 3.2), then the PG FOP is about half a centimeter before the dose FOP for this distal spot. Under influence of the primary energy and target material composition, the penetration depth and therefore the FOP will change. PGs are produced in nuclear interactions (fig. 3.3), but only if the incident particle has enough energy. At the distal end, the primary has only a few MeV left and that is not enough to both overcome the electric potential and produce fragments such as a PG, hence the difference in PG and dose FOP. Since the PG is produced in a nuclear interaction, the time delay between primary-target ion interaction and PG production is much below 1 ns, and originates directly from the particles path. PGs are thus a very good indicator of primary tracks and that is why they are a topic of investigation.

A general rule of thumb is that for 10 cm range in water, a proton has $\approx$ 10% chance to produce a PG and a carbon-ion about $\approx$ 40%. On clinical phantoms with clinical proton plans, I observe production factors between 3-10%. Figure 3.4 shows the PG production for 100k primary protons in Geant4, broken down by the multiplicity of the PG production number. About half of the nuclear interactions ("fHadronInelastic" in Geant4 parlance) produce no PG, a bit less produce 1 PG, and there is a tail of nuclear interactions producing more than one PG. PGs are produced with relatively high energies, up to about 8 MeV, significantly more energetic than the photons produced by most radionuclides used for PET and SPECT imaging. While recording high energy photons is a challenge, they provide one significant benefit: higher energies have lower probabilities to be absorbed or scattered in the patient. A less noisy signal with respect to PET or SPECT images can therefore be expected. Moteabbed et al. (2011) puts the transmission ratio between PET and PG at a factor 5 in favor of PG. Moreover, the production ratio is more favorable than PET, and with washout and delayed acquisition a difference between a factor of 60-80 can be expected.

An important consideration is that we currently assume that PGs are spatially emitted isotropically. Geant4 also adheres to this assumption. However, there exists evidence to the contrary (Sheldon and Van Patter, 1966; Verburg et al., 2012). Also, the target nucleus species has an influence on the PG energy. Certain materials tend to produce certain PG energies. For instance, the 4.44 MeV peak in the PG production energy spectrum shown in 3.5 is attributable to ^{12}C. Hence, recording the PG energy reveals something about the matter in which is was produced: 4.44 MeV PG would indicate a carbon-rich tissue.

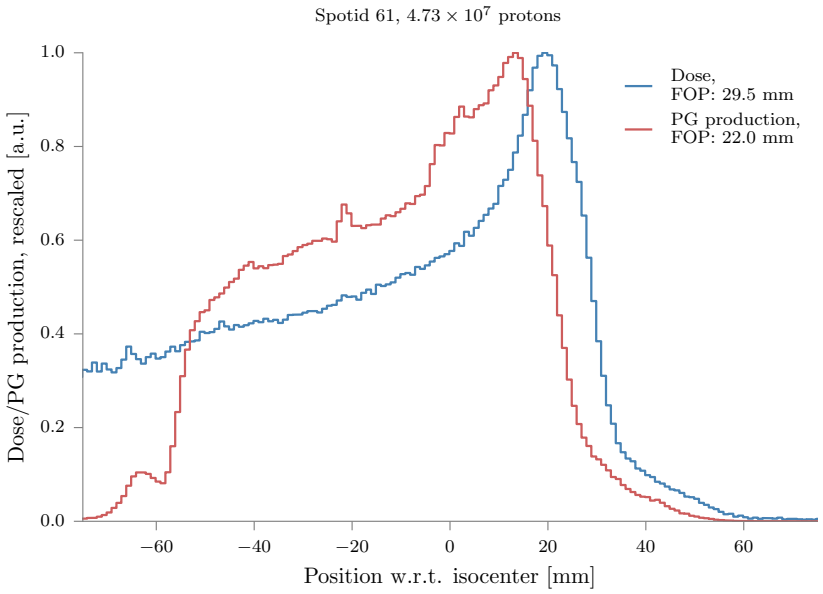


Figure 3.1: A clinical proton therapy simulation in Gate/Geant4 where 10^6 protons are launched and the resulting dose and PG distributions are plotted, rescaled to a maximum of 1.

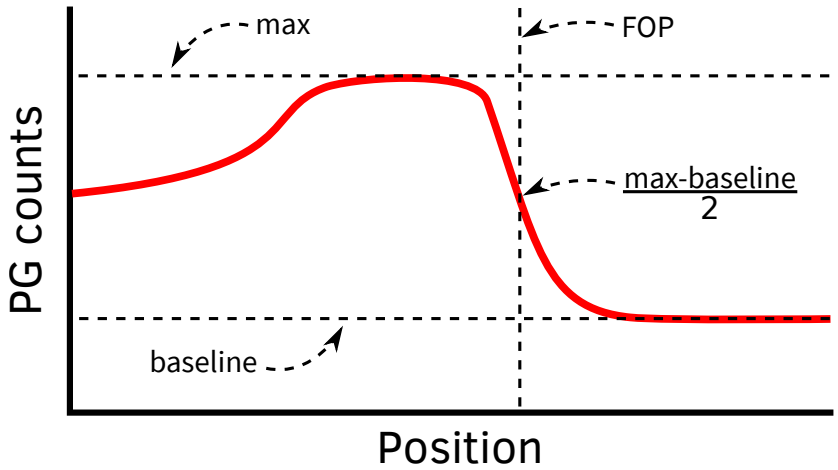


Figure 3.2: A hypothetical PG profile measurement along the beam line (red). Dotted lines indicate the maximum, baseline, and the fall-off position, or FOP, which is obtained by drawing a vertical line at the position where $\frac{max-baseline}{2}$ intersects the profile. Note that $\frac{max-baseline}{2}$ is an arbitrary choice.

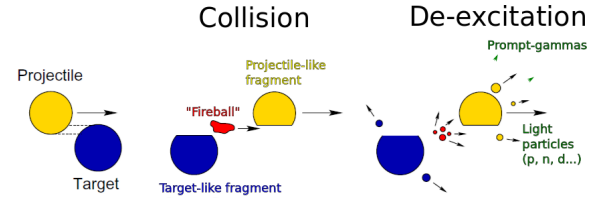


Figure 3.3: A schematic illustration of a nuclear interaction between an incident ion and an ion in the target. Here a partial head-on collision is sketched, where a fireball is sheared off both incoming and target ions. This fireball sometimes inherits a large fraction of the kinetic energy of the incident ion, causing it to have a longer range than the BP of the primary ions, resulting in the dose-tail. Both target, projectile and the fireball can relax through emissions of PGs.

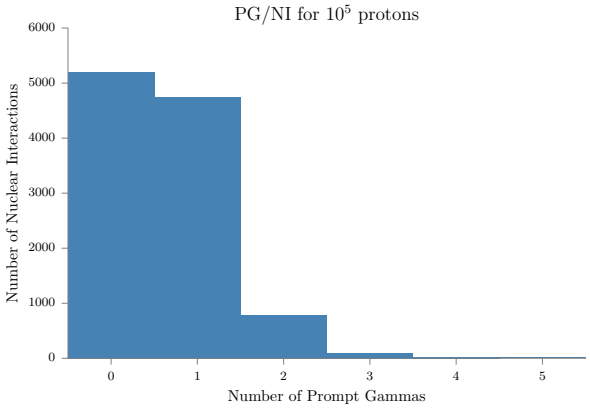


Figure 3.4: A histogram of the number of PGs emitted per nuclear reaction for a clinical proton therapy simulation in Gate/Geant4 where 10^5 protons are launched. The total number of nuclear interactions producing a PG is the sum of the data, and is about 6% for this clinical CT image plus plan.

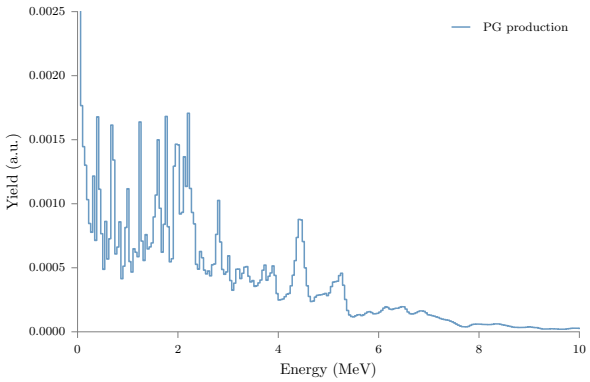


Figure 3.5: PG production energy spectrum of a clinical proton treatment simulation.

3.2 Gamma-ray detection principles

The second part of the term 'Prompt Gamma' is derived from the typical energies the PG has, in the 1-10 MeV range. Although the energy limits are not sharply defined, high energy photons are referred to as γ particles. In detection this distinction becomes important, because detecting higher energy photons is more difficult than lower energy photons (eV or visible light range). Intermediate energy photons ($\approx 100keV$) can be referred to as X-rays. For low energy photons, a variety of commercially available detectors exist. The two most common principles in these detectors are detection through excitation of an electron to the conduction-band in a semiconductor or by liberating an electron by way of the photoelectric effect. In the first category, the CCD (Charged Coupled Device) and CMOS (Complementary Metal Oxide Semiconductor) are two of the most common photon detectors, found in virtually every camera and phone today. The free electrons represent a charge which can be interpreted as a singal by a digitizing circuit. By binning the photosensitive area into pixels and reading out every pixel individually, an image can be recorded. In the second category, a photomultiplier tube (PMT) amplifies the initially liberated electron (from the cathode) by using electric fields and successive stages of dynodes to induce a cascade of electrons which ultimately are absorbed in an anode where the charge can be measured. The PMT cannot measure the energy of the incoming photon (because it usually releases only a single electron); the PMT only counts the number of photons.

Commercially available semiconductor-based imaging methods work best for photons with relatively low energies of up to a few electronvolts (eV). For gammas, such detectors are too thin for the photon to be fully absorbed and only their position might be measured. Since absorbing the energy is required in order to know the energy, semiconductor-based detectors would need to be very thick and therefore very expensive to guarantee absorption. Photons produced by radionuclides are typically in the 100 keV range, with the back-to-back photons produced after beta-decay at their signature 511 keV. PGs can go even higher as seen in fig. 3.5. For such high energy photons, a paired scintillator and PMT is the usual choice because they provide larger sensitive volumes than semiconductors for a given cost. The scintillator is a material opaque to photons at higher energies and transparent at lower energies. Scintillation means absorbing energy and re-emitting it as light. The scintillating material thus acts as a photon energy converter, where the number of outgoing lower energy photons is proportional to the energy deposited by the impinging high energy photon. This gamma energy proportionality to a quantity of light is best measured with a PMT: they are very good photon counters. Integrated the PMT output allows reconstruction of the gamma energy that impinged the scintillator.

For impinging high energy photons, high-Z materials or inorganic crystals are most effective. The photoelectric effect, Compton scattering, and pair production are the dominant interaction processes of X-rays and gamma-rays. The photoelectric cross section is proportional to Z^5, for pair production to Z^2, and Compton scattering goes approximately as Z, where Z is the atomic number. It is evident why high-Z materials are a common choice for gamma detection. Three properties can be measured about a photon: the incident direction, the energy and the time of arrival. In clinical context, retrieving any information about the treatment delivery is the goal, so these properties can be exploited in various ways to gain certain knowledge. From the direction we could build a picture of the emission profiles along the beam direction. By observing the fall-off, the range of the primary can be inferred (fig. 3.1). By recording the time difference between launching the primary and recording the PG, a range, similar to what is shown in fig. 3.1, can be inferred as well. By observing the energy, the material in which the primary was stopped may be inferred. These quantities can be used to measure similar or different quantities about the treatment, and therefore require different interpretations and clinical thought. In the remaining part of this chapter, the devices currently in prototyping phase measuring one of the three quantities will be discussed.

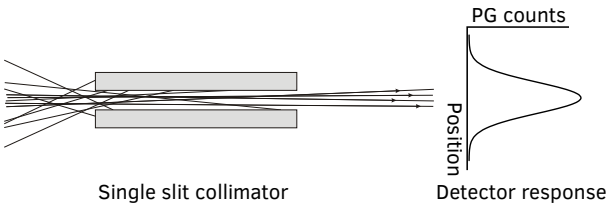


Figure 3.6: By placing photon absorbing materials in the right way, the incident angle is constrained to a (very) limited range. The fact that the angular acceptance is a finite range translates to a point spread function (PSF) in the detector. In the figure, the photons come in from the left, and on the right the PSF is sketched.

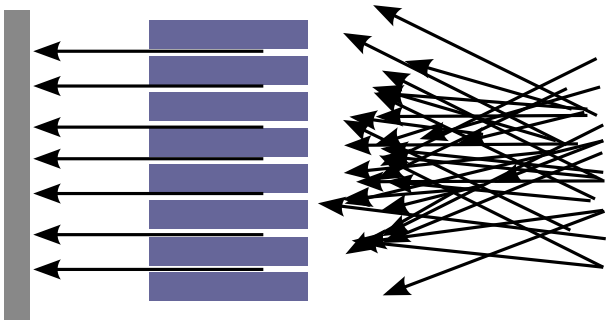


Figure 3.7: Parallel slits in an absorbing material provide multiple bins or pixels with the same constrained incident angle, allowing a 1D image of the source to be created. This type of collimeter is called the Söller collimator or multi (parallel) slit collimator. Image CC-BY-SA licensed by Wikipedia Authors.

3.3 Collimation

The incident *direction* of a PG can be established by limiting the field of view (FoV) of the photosensitive area, see figure 3.6. This is possible with high-Z materials, like with scintillators. The difference is that now a full absorption is needed, not a material that re-emits the energy as light. Metals such as lead and tungsten are a common choice. Depending on the thickness of the collimator and the width of the hole or slit, the incident angle is constrained to a limited range. The fact that the angular acceptance is a finite range and not a single value translates to a point spread function in the detector: particles emitted under an angle 'see' a narrower slit and therefore the collimator transmits fewer photons as the incident angle of the photon increases, the penumbra. The effect is described by a point spread function (or PSF, see right hand side of 3.6). A narrower slit means a smaller angular acceptance, which translates to a better precision on the incident path (smaller contribution by the penumbra), but a higher number of rejected (absorbed) photons and therefore a lower signal. Since sources of noise are usually constant, a lower signal means a worse signal-to-noise ratio. The number of transmitted photons per emitted photon may be called the collimation efficiency.

A single slit or hole would only give information about the production of one point or line. In PG verification, one of the things that is of clinical relevance is the proton range. In that case, as per figure 3.1, an image of the dimension along the beam must be recorded. That goal can be attained by either placing multiple slits along the beam, each collecting into a bin, or by using a relatively thin collimator with a single slit and project the angular dispersion onto a segmented detection surface.

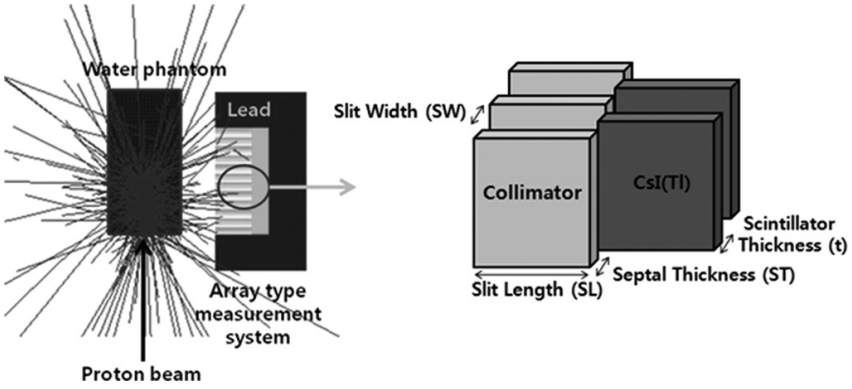


Figure 3.8: The MPS camera design validated by experiment at the IBA C230 synchrotron at National Cancer Center (NCC) in Korea, where the SW = 2 mm, ST = 2 mm, t = 3mm and SL = 150 mm.

3.3.1 Multislit Collimation

The most straightforward collimator geometry is the multislit collimator. Each slit is collected in a single bin, and by having these slits and bins along the beamline, a 1D PG emission profile is obtained. The perpendicular dimensions (the length of the slits) can be used to increase the collected signal, at the cost of losing the angular information in that direction.

Korea

A multi parallel slit (MPS) design (3.8) was presented in Min et al. (2012), were the setup was tested in silico using MCNPX. A parametrized PG source was created based on $2 \cdot 10^8$ primary protons, which facilitated obtaining four camera parameters: scintillator thickness (t), slit width (SW), collimator thickness (ST), and slit length (or depth, SL). Other parameters such as the scintillator length (50 mm) and the height of the scintillator (30 mm) were chosen based on 'intuition'. Surprisingly, the material of the collimator is not provided. AAn equation was derived for the optimal thickness of the scintillator based on the ratio between PG and background counts. The background was found by averaging the gamma counts within the 2–4 cm range distal to the PG peak. The source for the optimization stage was a 80, 150, and 220 MeV proton beam, while the target was a waterbox sized 20x20x40 cm. For one configuration, where the collimator has a 4 mm pitch, 2 mm slits (SW) and 2 mm collimator thickness (ST), scintillator thickness t = 3 mm (t) and slit length SL = 150 mm, an experimental validation was performed at the IBA C230 synchrotron at National Cancer Center (NCC) in Korea. Here, the CsI(Tl) scintillator (Institute of Scintillating Material, Ukraine) was coupled directly to a photodiode (S3588-08, Hamamatsu Photonics K.K., Japan).

The optimal dimensions of the scintillators and the collimation system were finally determined to be 3, 2, 2, and 150 mm for the t, SW, ST, and SL, respectively. A PG to background ratio of approximately 0.20 is obtained with the optimized parameters. Finally, fall-off estimation was performed with a sigmoid-fit. Unfortunately, no results for the fall-off precision are presented.

A patent filed later by competitors (Freud et al., 2015) describes an optimization of a MPS collimator based on the collimator presented here, resulting in different values for the parameters, which in turn resulted in a different signal to background ratio. In Min et al. (2012), parameters were optimized individually, while in Freud et al. (2015) it was attempted to take into account the interplay of parameter choice as well. This indicates a significant complexity when it comes to detection optimization, where parameters should not be chosen independently.

Case	Precision fall-off	Precision entrance	Precision profile length	FOV	Slit	Slab	Thickness	Distance axis-collimator	Distance axis-detector
1	0.59(0.06)	0.66(0.09)	0.88(0.11)	23.6	5.4	2.6	180.2	303.7	485.3
2	Same as Case 1								
3	0.70(0.08)	0.87(0.12)	1.12(0.14)	13.1	3.0	2.1	190.9	322.3	516.5

Table 3.1: Table 6 taken from Pinto et al. (2014), containing the collimator geometry and performance of the optimization carried out with $5 \cdot 10^8$ protons. Units are in mm unless specified otherwise.

CLaRyS

A MPS collimated design is currently under construction at the Institute of Nuclear Physics of Lyon, as part of the CLaRyS project (Contrôle en Ligne de l'hAdronthérapie par RaYonnements Secondaires), a cooperation between various French laboratories: IPNL in Lyon, LPC in Clermont-Ferrand, LPSC in Grenoble, CREATIS in Lyon, LIRIS in Lyon, CPPM in Marseille and the Centre Antoine Lacassagne in Nice, which hosts a medical proton beam. The detector consists of repurposed and refurbished scintillator-PMT units originating from a Phillips PET scanner. The scintillator is a streaked block of bismuth germanate oxide (BGO) providing 8 by 8 pseudo pixels. There are a few preferred choices for scintillation materials (most notably Cesium-doped materials such as LaBr3:Ce and LYSO:Ce), but BGO was determined to be comparable to both, and better on account of it having no intrinsic radioactivity Richard et al. (2012). Each scintillator block is read out by four PMTs which allows a reconstruction of the impact position via centroid calculation. Although not done at this time, sub-slit position measurements may improve the results at a later time. Although (Krimmer, 2015, Section 2.2) describes the BGO blocks in the context of a Compton camera, the blocks are in fact the same ones as will be used for the MPS camera. Read-out electronics for the blocks will also be the same. In addition, time of flight (ToF) measurement is enabled by the inclusion of a hodoscope, which timestamps incident protons. Based on the observation that the background is largely a constant noise, while the PG signal arrives within nanoseconds after a proton pulse, about a 60% of the background is removed by setting a ToF acceptance window on the detector output. The PG energy window is set at 1 - 8 MeV.

The design of the Tungsten collimator was largely determined by the study laid out in Pinto et al. (2014). Here, for the first time, an in silico (Geant4) optimization study was carried out by randomly varying the collimator parameters (slit width, slab width, slab thickness, distance to source) and testing its performance on a few key indicators. This resulted in three configurations performing best for three indicators: best precision on the fall-off position (Case 1), best precision on the profile length (distance to FoP, Case 2), and best precision on the fall-off position with a FoV lower than 15 mm (FoV here refers to the sum of the slit width + twice the penumbra, Case 3). The results of the optimization are presented in table 3.1. These results were obtained with a homogeneous target and a pencil beam counting $5 \cdot 10^8$ protons. The fall-off is retrieved by taking a high statistics profile as reference, and fit spot profiles by a non-uniform rational B-spline (NURBS) function in the region of the profile fall-off.

In summary, the collimator geometry optimized for best fall-off precision attains a precision of 0.59 ± 0.06 mm in a homogeneous target, using $5 \cdot 10^8$ protons with the aforementioned ToF selection. The mono-energetic 160MeV pencil beam was centered in a cylindrical PMMA phantom with a 75 mm radius and 200 mm in length.

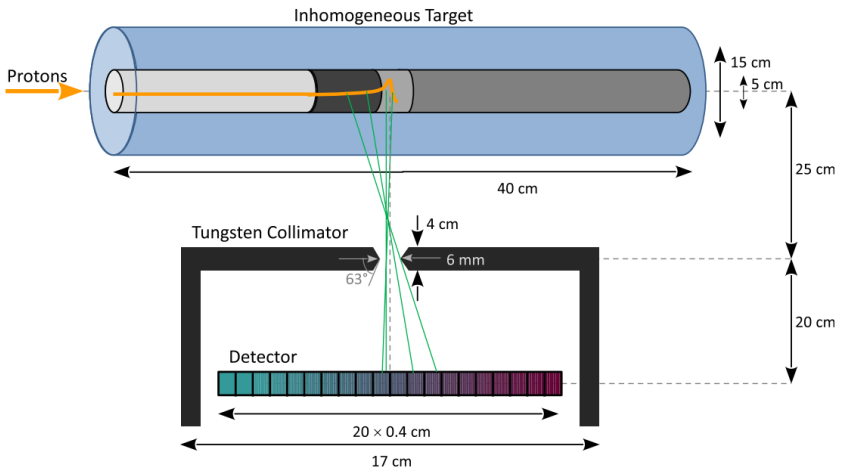


Figure 3.9: The IBA knife-edge camera as shown in Priegnitz et al. (2015). The collimator material is Tungsten. The phantom is an outer PMMA shell in which various inserts ('mask', air, lung, adipose, muscle, brain and bone) can be placed in various order, to study their effect on the range and the PG profile.

3.3.2 Knife-Edge

In the parallel slit configuration, each slit is supposed to provide a projection of the line-segment directly in front. Each slits line segment is placed sequentially to obtain a full profile. The penumbra of each slit is minimized by the relatively deep slits, although increasing the size and weight of such designs. The advantage is that the detection response is constant over the range of the profile, and the camera field of view is determined by the number of slits. Another approach is to take a single slit collimator and shape the slits edges. The effective slit width as function of the incident angle is kept constant, up to the angles of the slits edges. An image of a much larger line segment than a single slit in a MPS design can then be seen. This enables a thinner collimator, at the cost of a position dependent detection response: the line segments further away from the slit (in the transverse direction) are further away from the collimator and target. Because the resulting slit resembles a knifes edge, the collimator design is called the Knife-Edge Slit (KES).

IBA

Fully characterized in Perali et al. (2014), and the performance studied in detail by Priegnitz et al. (2015), the IBA prototype is a KES design and the first and only camera at the time of writing with published clinical results (Richter et al., 2016). The collimator is a 4 cm thick Tungsten block, with a 6 mm slit with 63 degree knife edges (fig. 3.9). The scintillator is LYSO (provided by Saint Gobain) and was selected for its density, its light yield and for its short decay time of 41 ns. LYSO suffers from the presence of the beta decaying isotope ^{176}Lu, which has an upper limit on the decay energy of 1.193 MeV. By setting an energy acceptance window on the detector output of 3 - 6 MeV, these events should be eliminated. The height was chosen to be 30mm, which facilitates a 1:1 coupling to the readout which is a SiPM array size of 31.5 mm. The LYSO and readout are segmented every 4 mm along the beam profile direction. The ratio between the source to collimator and collimator to detector distances implies a scaling factor of 0.8, which means that the total length of the detector, 8 cm, corresponds to an ability to measure 10 cm of the PG emissions at the source. It should be noted that if the camera is not moved during a clinical treatment, each spot would have a slightly different magnification factor, due to the different transverse spot positions and the different source

to collimator position that implies. No solution for this implication is provided. Since moving the camera towards or away from the patient synchronous to each spot is an impracticable solution, translating the transverse distance to the camera to the correct magnification factor is a solution that may be considered.

To retrieve the fall-off, Priegnitz et al. (2015) describes a procedure that entails taking the first and second order derivatives of the obtained profile:

1. Normalization and background subtraction: all profiles are scaled so that the mean of the three most proximal data points is one and the mean of the three most distal points is zero.
2. Distal edge definition: it is defined as the most distal part of the profile, where data points are larger than a certain threshold (here, we chose 0.1 to suppress statistical background fluctuations), the profile is decreasing (first derivative < 0), and only one inflection point occurs (second derivative is zero only once). What is not described, is what happens with data that does not pass this criterion.
3. Normalization for range estimation: defined distal edges are normalized to 1 and 0.
4. Range deviation estimation with distal slope matching algorithm, where the match is performed by profiles obtained with a high statistics measurement of simulation.

Next, the effects on the range and PG profile of replacing various inserts in the phantom and introducing shift due to a change in beam energy are studied. For each 'shift configuration' (be it a replaced material or a changed beam energy) a physical measurement with $2.6 \cdot 10^{11}$ protons is performed at the Proton Therapy Center Czech in Prague, Czech Republic, and in silico estimations at 10^7, 10^8 and 10^9 are performed. The in silico estimations are provided by the analytic variance reduction method for PG estimation laid out in Sterpin et al. (2015). As described there, at the time the algorithm did not deal with lateral inhomogeneities, which is the likely reason for the study to include a 1D phantom. Each in silico PG estimation is executed a thousands times and the fall-off obtained with the procedure described above histogrammed. The mean and standard deviations of these histograms provide the fall-off estimation precision. Then, based on the ground truth measured in the physical experiment, receiver-operating characteristic (ROC) curves are created and provide an estimate of whether or not the shift is correctly obtained as function of the number of primaries launched, and for each change in material or energy shift.

The best fall-off precision (standard deviation on the FOP measurements) reached is about 2 mm for targets that are homogeneous in the field of view of the camera, with 10^9 primaries. Both conditions may be expected to be rare in clinical conditions. Perhaps more interesting is the implication of (Priegnitz et al., 2015, fig 8): the shifts, beam energies and inhomogeneities interplay in interesting ways, which might be exploited to learn more than only about fall-off positions.

Some observations by the same group published in Smeets et al. (2012) are of interest:

- PG profiles after 4 - 5 MeV energy selection have the best signal to background ratio,
- PG profiles don't correlate with deposited dose with an upper energy threshold of 1 MeV, which complements the observations regarding the intrinsic activity of LYSO,
- Lower detection energies show less sharp Bragg Peak, complementing statements regarding scattered PGs in Min et al. (2012).

Delft

A KES design was combined with ToF in Lopes et al. (2015), which features a very similar collimator as the IBA prototype, where the collimation slab is Tungsten and 4 cm thick with a 6 mm slit (3.10). The angle of the knife edge is not given. The source to collimator and collimator to detector distances are identical: 15 cm. The work builds on the results the group published in Biegun et al. (2012) by demonstrating measurements with the fully realized prototype obtained in the German Proton

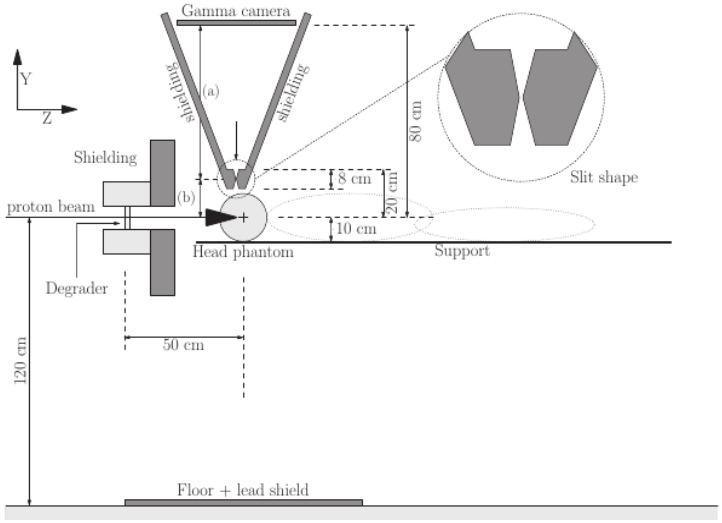


Figure 3.10: The Delft knife-edge plus time of flight selection design. Source: Lopes et al. (2015).

Therapy Centre Essen (WPE), with a 160 MeV proton beam and the camera aligned with the expected fall-off position. Two scintillator materials are compared (LYSO:Ce and BGO), and a detailed analysis of the data acquisition electronics is provided, in particular how the ToF and PG energy measurement is optimized. It is stated that the BGO scintillator shows better sensitivity for high (>6 MeV) PGs. The effectiveness of ToF at removing noise is underlined. The ToF-energy matrices prove to be a good way to visualize which part of the distribution correlates to PG events and which not, and these results as well as the signal to noise ratios per energy bin are provided for each scintillator material. The results are a bit different, so optimal choices for ToF and energy spectrum threshold appear to be a function of the scintillator material.

Sigmoid curve fitting was used to obtain a fall-off position from profiles measured with $6.5 \cdot 10^9$ protons shot into a PMMA cylindric phantom with a 15 cm diameter and a 20 cm length. A reduced statistics measurement was obtained by random sampling of the $6.5 \cdot 10^9$ measurement, resulting in an estimated $6.5 \cdot 10^8$ acquisition. Via bootstrapping, a 100 measurements per setting are simulated, making the estimation of the standard deviation possible. The inflection point of the distal slope was chosen to be the fall-off position, similar to what was seen earlier. The phantom was shifted in its entirety over 1, 2, 3 and 4 mm and the inflection points plotted. A standard deviation was obtained Results are given for BGO, LYSO:Ce and LYSO:Ce + ToF, and per number of primaries. In some cases, with $6.5 \cdot 10^9$ primaries, a precision on the FOP of better than 1 mm is reached, however, especially with BGO, small shifts are not seen. With LYSO:Ce, small shifts are systematically underestimated.

3.3.3 Pillar Grid Collimator

The Pillar Grid collimator described in Kim et al. (2016) is probably the most ambitious design to date. The goal is nothing less than 3D dose reconstruction by way of PG computed tomography. The study was carried out in silico (MCNPX for the proton propagation, Gate for the PG propagation). The optimal parameters (figure 3.11) were found, where a trade-off was found for maximal measures for the half-maximum (FWHMs) and best signal-to-noise ratios (SNRs). Also the effect on the FWHM and SNR of various cuts on the PG energy was studied. Finally, reconstruction was carried out with three methods: back projection (BP), filtered back projection (FBP) and maximum likelihood expectation maximization (MLEM). Fall-off positions are obtained with a sinusoidal fit, and compared between the three reconstruction methods, resulting in values within 2.5 mm of each other. The setup was an 80 MeV proton beam shot into a waterbox. The authors admit that to save on simulation time, a

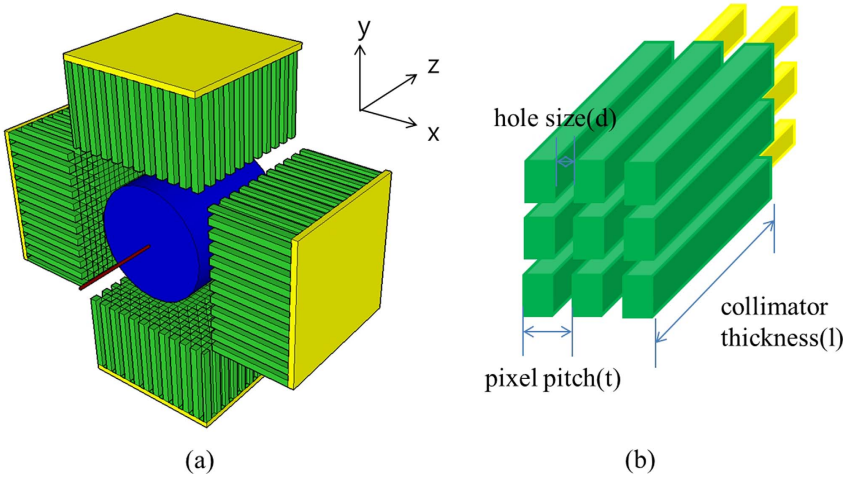


Figure 3.11: An ambitious design proposal: pillar grid collimation with the aim of enabling 3D dose reconstruction through PG imaging. Optimal parameters were found at l = 7 cm, t = 5.5 mm and d = 3 mm. Source: Kim et al. (2016).

limited number of primaries had to be used and the results should be evaluated in 'a relative manner' and the quality of the reconstructed images would be improved by increasing the simulation time, although they do not expect a change in precision on the BP. An efficiency of 0.93% is given, where efficiency is defined as the number of valid interactions used for the image reconstruction divided by the number of incident protons. A correlation between dose and PG fall-off is found, but not subjected to a statistical analysis.

3.3.4 Comparison study

Of course, understanding the differences between different designs such as MPS and KES is easiest in a direct comparison. In Lin et al. (2016) an in silico comparison (using a relatively old version of Geant4/Gate) is performed where a 160 MeV pencil beam of 10^8 protons was centered in a cylindrical PMMA phantom with a 75 mm radius and 200 mm in length; the exact same target as in Pinto et al. (2014). Each simulation was run 5 times to obtain a standard deviation. Apart from the type of collimator, all parameters were kept equal so that any difference in results may be linked to nothing else except the collimator design. A LYSO scintillator was used, a KES design similar to what was described in the IBA section and a MPS design similar to the Korean design, see figure 3.12. Fall-off positions were retrieved with the method presented in Smeets et al. (2012), and PG detection was binned in 1 mm bins. It should be noted that the detectors were optimized with different procedures and then modified to have the same detector-to-collimator and collimator-to-source distances. From Roehlinghoff (2014) we know that the two types have similar performance with similar parameters.

The authors note that the neutron-to-photon ratio is higher in the MPS design, causing a degradation in signal and in the conclusions the authors note that ToF selection would have helped. The fall-off width (or fall-off region) is wider for the KES design, but according to the authors the fall-off slope may lead to less accurate range detection, which is contradicting a finding in Roellinghoff et al. (2014). However, a better precision of the fall-off position (FOP) is found with KES design nonetheless: KES σ_{FOP} = 0.49mm and MPS σ_{FOP} = 0.85mm. This is attributed to the better peak to background contrast in the KES profiles. Then, an energy window giving the best BP correlation is found at 2-7 MeV for the MPS collimator 3-7 MeV for the KES. With these cuts, the accuracy improves, and the precision is better, and now in favor of the MPS, perhaps on account of the wider energy cut and the correspondingly higher count rate. What is interesting is Table 2, where the various parameters of

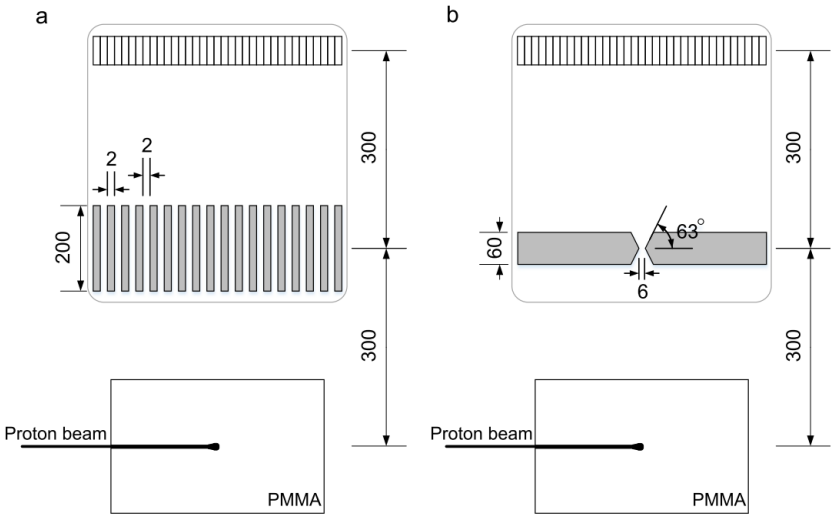


Figure 3.12: The two collimated detector designs compared in Lin et al. (2016). Detector a is the MPS model; detector b the KES. The dimensions of the KES collimator were based on the IBA KES prototype, while the MPS dimensions reflect the Korean MPS. Beam-collimator and collimator-detector distances were equalized.

the fit procedure are given in terms of a mean and standard deviation, for three primary energies (50, 100 and 160 MeV). Precision is generally in favor of the KES, while at 50 MeV the accuracy of the MPS is best. The effect of the location-dependent solid angle in the KES design is noted. In table 3, the effect of a larger phantom is studied, and the attenuation affects the baseline and peak magnitudes. An effective 10 cm extra path through the PMMA phantom translates to a reduction in peak contrast of 30% with the MPS collimator and 40% with the KES collimator. The last analysis concerned shifting the phantom between 1 and 20 mm, which resulted in standard deviations on the shift estimates of 3.82 and 2.46 mm for MPS and KES collimators respectively. When the authors remove the background signal, to simulate a perfect ToF selection, standard deviation of 1.87 mm and 1.76 mm are obtained. All precisions are based on simulations with 10^8 protons and 5 realizations for statistics, the latter of which is reason for caution with respect to the standard deviations. Also, the collimators were optimized independently by different authors, so general conclusions with respect to MPS versus KES designs are difficult to draw.

3.4 Compton Camera

Compton cameras feature scatter planes as a method of focus. One or more planes of a certain material will scatter the incoming photon, through Compton interactions where part of the energy is transferred to an electron in the plane. What makes this physical effect useful in a camera is that the photon scatter angle and photon energy loss are related, according to the following equation:

$$\cos\theta = 1 + m_e c^2\left(\frac{1}{E} - \frac{1}{E'}\right) \tag{3.1}$$

where E is the incident photon energy, E' the photon energy after scatter, θ the angle of scatter, and $m_e c^2$ the energy at rest of an electron (511 keV). If the scattered photon is then (fully) absorbed in the scintillator, similar to those in collimated designs, the full remaining energy is equal to the energy after scatter, E'. The energy released in the scatter plane can be measured, and the initial photon energy E is then equal to the sum of the energies measured in the scatter plane and absorber. In both

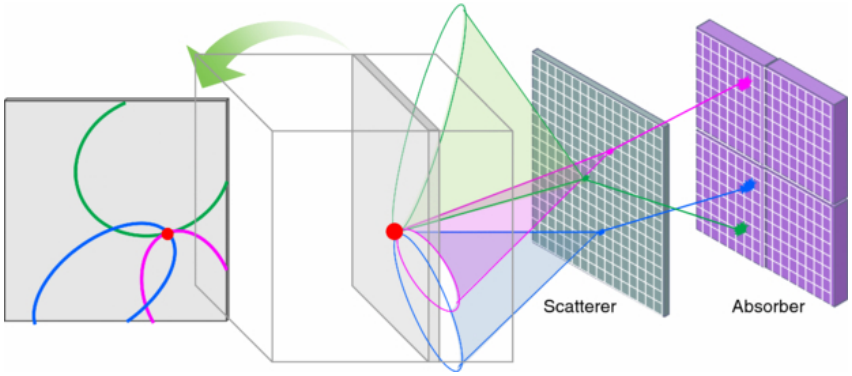


Figure 3.13: A scatter plane and an absorber allow (the cosine of) the incident angle to be reconstructed. Various measurements results in various cones, which intersect at the position of the source. Image courtesy of Kim et al. (2013).

scatter plane and absorber the position is measured, after which the angle of the incoming photon can be reconstructed. A few measurements would results in a few so-called Compton cones, after which the point(s) where the cones overlap is where the source was (fig. 3.13). Dedicated reconstruction algorithms, most based on iterative methods, infer the mean ion track of the treatment portion that is considered (a spot or an energy layer for instance). In the context of particle therapy, having a hodoscope track the ions direction could provide a constraint for a cone-by-cone reconstruction.

The main challenge is filtering out noise due to random coincidences. Like in PET, Compton cameras detect coincidences in different detectors due to unrelated events as well, which cannot be distinguished from true coincidences. That means there is a saturation effect at high rates. PGs are produced by the primary particle, and they are part of bunches in the accelerator. So depending on the accelerator, and specifically its micro time structure, we can have the expected PG rate at the Compton camera. At clinical beam intensities, Ortega et al. (2015) states 12.5 protons impact the patient per nanosecond, which is well under their timing resolution. New accelerator designs such as the IBA S2C2 will have even denser micro time structures, leading to higher expected PG rates and higher numbers of false coincidences. One possible solution is to start the treatment with a reduced rate, and after sufficient signal is recorded and no significant difference is measured, the treatment could continue.

Since the incident photon energy is not known, perfect absorption is a requirement for the computation of the correct angle of the Compton cone. A way to obtain the correct incoming angle without perfect absorption is by having multiple scatter planes, as was shown in Richard et al. (2011). A simulated double scattering Compton camera was demonstrated to have a "true efficiency", the number of true reconstructed events over the number of (isotropically) emitted PGs, of $0.7 \cdot 10^{-5}$. Increasing the number of scatter layers means increasing the detection efficiency.

MACACO (Valencia)

One of the two complete prototypes is MACACO from Valencia. The prototype has three layers, each a mix of an absorber and scatterer. Each layer is composed of a slice of $LaBr_3$ attached to a SiPM (silicon photomultiplier) readout. The absorber thickness is 5 mm for the first two layers, and 10 mm for the layer furthest from the target. The transverse size of each layer is 27.2 by 26.8 mm for the first layer and 32 by 36 mm for the other two. The first images of a radioactive source acquired with the fully realized prototype were presented last year (Llosá et al., 2016), and the first tests with a 150 MeV proton beam in a PMMA target in Solevi et al. (2016). In in silico experiments (using Gate), the efficiency of an PG undergoing a double interaction in the first and second detector layers

ranges from $6.9 \cdot 10^{-6}$ for lower PG energies and $2.5 \cdot 10^{-5}$ for higher PG energies. With energy cuts these number drop, but the spatial resolution improves. In the acquisition with a proton beam, cuts around the 4.44 MeV PG energy spectrum spike were used to study a ±5 mm shift; a 7 mm shift was measured based on 18000 coincidences for the -5 mm shift and 28000 for the +5 mm shift. Only the incident protons *rate* is given: $9 \cdot 10^7$ protons s^{-1}.

Polaris-J (United States)

The second complete prototype is the Polaris-J Compton camera, which was subjected to proton beam tests in Polf et al. (2015). The camera has four detection planes, each identical, made up of a Cadmium Zinc Telluride scintillator of 1.5 cm thick an a total transverse cross section of 4 by 4 cm. The detector configured in double coincidence mode, so between two and four layers must have a simultaneous hit to be recorded as an event. The main aim was to establish whether or not the camera could deal with clinical beam intensities, so this was tested with a single pencil beam of 150 MeV ($8.2 \cdot 10^6$ protons/second) and a layer of 5×5 114 MeV pencil beams ($6.4 \cdot 10^7$ protons/sec), both shot into a $20 \times 20 \times 32$ cm waterbox. A "stochastic origin ensemble" PG image reconstruction code is used, published earlier and tested on Monte Carlo generated data.

The 150 MeV beam PG yield was acquired after reducing the beam current to a coincidence rate of about 4 kHz in the detector (which explains the sub-clinical rate of $8.2 \cdot 10^6$ particles per second mentioned above), while for 114 MeV 4.5kHz was determined to be optimal. The authors note this agrees with the efficiency of $\approx 5 \cdot 10^{-4}$ reported by the previous group in Ortega et al. (2015). With each beam, the waterbox was shifted 0, 3, and 5 mm, and between $3.85 \cdot 10^5$ and $1.24 \cdot 10^6$ coincidences obtained per acquisition. The 2D and 1D PG images are analyzed shown, but no quantitative statements with respect to the shifts are given. An estimated minimum number of protons required to compute an image with a sufficiently acceptable SNR ratio is computed to be $7.2 \cdot 10^8$, which is later presented as a required factor of 25 improvement in PG detection rate.

Another interesting finding is that the 60% level of the distal PG falloff peak corresponds well with the depth of the 90% level of the Bragg Peak, supporting other observations that in the final millimeters of the primary next to no PGs are produced. Further investigation into the exact relationship is suggested.

CLaRyS

The absorbed layer being built for the CLaRyS MPS camera is also being built to be the (distal) absorber layer for a Compton camera designed simultaneously alongside the MPS camera. By sharing the BGO absorber, associated readout, and hodoscope, comparison between Compton and MPS designs will be interesting to observe.

In the design stage of the camera, an in silico (Geant4) study of the optimal geometry is performed (Roellinghoff et al., 2011). The scatter planes consist of 2 mm thick silicon strip detectors and a LYSO absorber. The number of scatter planes was varied between 2-34, with them equally spaced and 9 cm between first and last plane, and the first plane 10 cm from the source. 10 scatter layers was chosen to be the optimum between cost and efficiency. The distance between the last scatterplane and the absorber was varied between 5-50 cm. 40 cm was chosen as the distance between first and last layers, which seemed a good trade-off. A good spatial resolution and ToF favors greater camera length (great distance between photon interactions), while detection efficiency favors a wide FoV, which is reached by placing the detection planes closer together. A different absorber material (LaBr3:Ce) is thought to provide increased spatial resolution. An interesting calculation follows in the conclusion: a carbon TP is considered that delivers 1 Gy to a 120 cm^3 tumor, in 39 slices. That takes about $7 \cdot 10^8$ incident carbon ions. The Bragg peaks of consecutive slices are in this plan separated by 3 mm. A PG emission rate of 10^{-3} photons per mm and per incident ion (carbon, between

95 and 300 MeV/u) is expected before the Bragg Peak. Due to peak summing, the most distal energy slice will be delivered with about 10% of all incident ions, so about $7 \cdot 10^4$ gammas per mm are expected for the distal layer. With the setup under consideration, reaching a detection efficiency of about $2.5 \cdot 10^{-4}$, this means about 2 gammas per mm for the distal energy slice of this 1 Gy fraction. A spatial resolution of 8.3 mm FWHM is reached, falling short of the 3 mm difference between energy layers.

The current design involves double sided silicon strip detectors ($90 \times 90 \times 2$ mm^3, 2×64 strips) as scatter layers and the readout electronics for both scatter and absorber layers are currently being produced.

3.5 Prompt Gamma Timing Camera

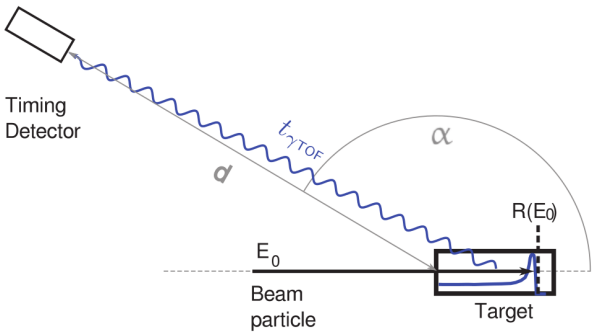


Figure 3.14: The Prompt Gamma timing camera schematic design as proposed in Golnik et al. (2014).

Apart from position, time of arrival (or time of flight ToF) provides information about the beam-patient interaction. PG timing (Golnik et al., 2014) cameras measure ion range by converting the ToF to a distance (fig. 3.14). The primary crossing 5-20 cm of tissue translates to a transit time of 1-2 ns. By positioning the camera under a small angle with the beam line, the ToF is almost doubled, which for a given camera timing resolution improves the ToF estimate. The paper explores the details of ToF cameras by analyzing, among others, the impact of timing resolution uncertainty, clinical beam currents and effective counting rates, energy cuts and absolute count rates based on the target material. Correlation between range and ToF spectra are found, but further study is required.

A later paper from the same group provides the first measurements, with a clinical proton beam and an analytic target: a cylindric tube of PMMA in which cylindric inserts can be placed (Hueso-González et al., 2016). It is confirmed that different materials have a different ToF-spectrum. The bunch time spread is a limiting factor: depending on proton energy it is found to be between 350 ps (230 MeV primary protons) and 2 ns (100 MeV), whereas transit time of the primary through the target, and consequently the width of the ToF spectra is between 3 and 4 ns. Based on replacing a PMMA insert with air or bone inserts, it is hinted that a single detector is able to recover 5 mm range errors with 10^8 protons.

3.6 Spectral PG Camera

PG spectroscopy is being developed at Massachusetts General Hospital (Verburg and Seco, 2014) (fig. 3.15), where the target tissue and proton energy is deduced from the PG energy spectrum. Certain peaks in the PG spectrum correlate strongly to excited states of target nuclei, and their intensity depends also on the energy of the primary. By measuring a few spectral lines in the spectrum, and computing their ratios, a sort of fingerprint is obtained. By tabulating this data based on the beam

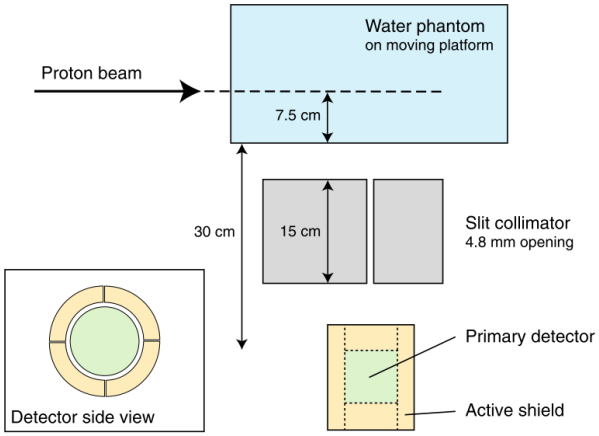


Figure 3.15: The schematic design for the spectral PG camera as proposed in Verburg and Seco (2014).

and CT data beforehand, the fingerprint can be looked up in the tabulated data and so the material and residual primary energy is known. From this, the position along the beam axis in the patient can be inferred, and with the residual primary energy the fall-off position can be estimated. Detailed knowledge of PG production cross-sections is required and much effort is dedicated to measuring and modeling them. After millimetric precision was presented, the authors expect clinical testing to begin soon (Verburg and Bortfeld, 2016).

An interesting detail is that these authors do not assume PG emission isotropy. The PG production cross-sections are computed for specifically a 90 degree angle with the proton beam.

3.7 Monte Carlo

Particle therapy is a much more complex than X-ray therapy, and adding detectors on top of that results in a large system with many components that interplay in even more ways. It is then not a surprise that so much work is still in the planning or prototyping phase. Fortunately, the quality of physics simulation (MC) tools such as Geant4, EGSnrc, PENELOPE, MCNPX and Fluka is improving so that they can be used in medical studies. This includes MC-aided detector design. For instance, when designing a collimated PG camera, design choices must be made and motivated. To name a few: collimator size, thickness, number of slits, slit width, knife edge angle, scintillator size and material, collimator to source distance, collimator to detector distance. For each of these, we know there is an influence on the expected signal, in terms of yield, precision and so on. What is usually not known is the precise influence, or the interplay between various parameters. Testing for these effects in actual setups is laborious and time-consuming. Beam time is expensive and in limited supply, while computing time is a fraction of the cost and increasingly ubiquitous. In silico optimization studies such as Lin et al. (2016) are therefore a huge time-saver and enable the investigation of parameters and their interplay which would not be possible in real experiments because there is no time for it. It should be noted that as far as I have seen, Pinto et al. (2014) is the only optimization study where the parameters were *not* studied independently, but in a multitude of combinations by employing a brute force approach: randomize all parameters and test all combinations. I think such studies will be increasingly a common part of any design study where physical parameters are expected or have a known effect on the signal. Moreover, for medical research such as described here the end goal is a proof that in clinical conditions the detector improves treatment or provides new pertinent information to the clinicians. That must eventually be shown on real patient irradiations, another experimental condition that is in short supply and constrained by safety and health factors.

Being able to read patient data, use treatment plans, irradiate the plan and observe the effect

for the thousands of combinations of possible parameter choices is an attractive thought. Of course, this requires not only that MC tools interoperate with clinical data, but also that their physics is tuned and validated for medical purposes. Historically, many physics MC tools were built for high energy physics (HEP) experiments, so that detection can be compared to the effect of physics as we think it is. The MC tool used in the original work in the following chapters of this thesis is Gate, which wraps Geant4, a product of the CERN collaboration. Gate provides an extended macro-interface and various components relevant to medical research, as opposed to the HEP-experiments focused components present in Geant4. Gate adds no physics itself, it relies completely on the Geant4 and its physics–lists for the physics. It does add variance reduction techniques, speeding up the convergence towards certain outputs. A physicslist in Geant4 is a collection of particular implementations of physics models. The same physics may be described various models, each optimized for different targets. For instance, a model can be optimized for precision or speed, or a particular energy range. For charged particles, the main options are Bertini, Binary Cascade, and QMD. For electromagnetic particles (electrons, photons) there is Standard, Livermore and Penelope. Other models are geared towards a particular subject, such as DNA, where (from the perspective of HEP) very low energies are involved, and we see the start of modelling chemical interactions. Quite recently, the Geant4 collaboration made a physicslist specific for hadron therapy available. Most physicslists do not include modelling for the magnetic field, so some lists provide variants that do include it. The details of these lists are outside the scope of this thesis, but one thing is not: to be aware that it is a choice that must be made and that the user must be aware of the level of validation the physicslist has received for his particular quantity of interest. For instance, the Hadrontherapy physicslist may be optimized for dose precision, but that does not mean it provides accurate PG estimates. Each purpose must be validated by experiment, and for PG we know Geant4 physicslists overestimate the yield. In recent years, the overestimate has grown less, but work is ongoing (Dedes et al., 2014; Pinto et al., 2016; Robert et al., 2013). PG background estimation – that is, all the signal a detector may pickup that is *not* a PG – has not seen any validation, which is at this time a major missing component when using the tool for PG detector validation.

3.8 Summary and outlook

We have seen the state of the art of most PG verification methods, collimated designs in particular. The first clinically used PG detector was a KES design and collimated designs certainly seems to be leading the field, probably because of its conceptual simplicity. Compton cameras are seeing a lot of interest too, although the only other design reportedly close to clinical testing is the spectral PG camera.

The field is focused on range verification; that is to say to exploit the correlation between range of the primary in the patient, the consequent distal dose fall-off (with protons no dose past the BP) and the PG profile fall-off. Optimal use of particle therapy implies optimal use of this sharp dose gradient, which is most interesting at the outer edges of a tumor, and even moreso on a tumor-AOR interface. Because of range uncertainties, such planning is currently avoided, forfeiting the advantage particle therapy could have. Observing the range, or in fact the fall-off position, might help reduce the range uncertainties and at the very least provides feedback on the dose of the day.

The precise correlation between primary range and PG fall-off is still a topic of research (Moteabbed et al., 2011). In addition, España and Paganetti (2010) state that duo to the stoichiometric HU to RSP calibrations, a sub 0.5 mm proton range variation is expected. They also say material calibration leads to much larger variation in PET computation, presumably because the material composition has a larger effect on positron production that on proton stopping power. It stands to reason that the effect for *PG* might be similarly larger than 0.5 mm. Clarification is needed.

Moreover, uncertainties in RBE may complicate the precise relationship. As alluded to in the

introduction, the RBE may be a function of the energy of the primary. There is reason to believe that the RBE is higher in particular around the BP. If correct and sufficiently proven, this leads to an effective *biological range extension* (fig. 3.16).

Imaging the PGs and in particular their range seems a sensible approach for particle treatment verification. However, the precise correlation between deposited dose and PG fall-off requires careful measurement of PG production cross-sections for various materials, which is a tedious, long and complex affair. Fortunately, even without dose-PG correlation, the difference between measured and expected PG signals can be used to provide clinical feedback. Moreover, as will be explored later, other features than the PG fall-off position may be useful for clinical feedback.

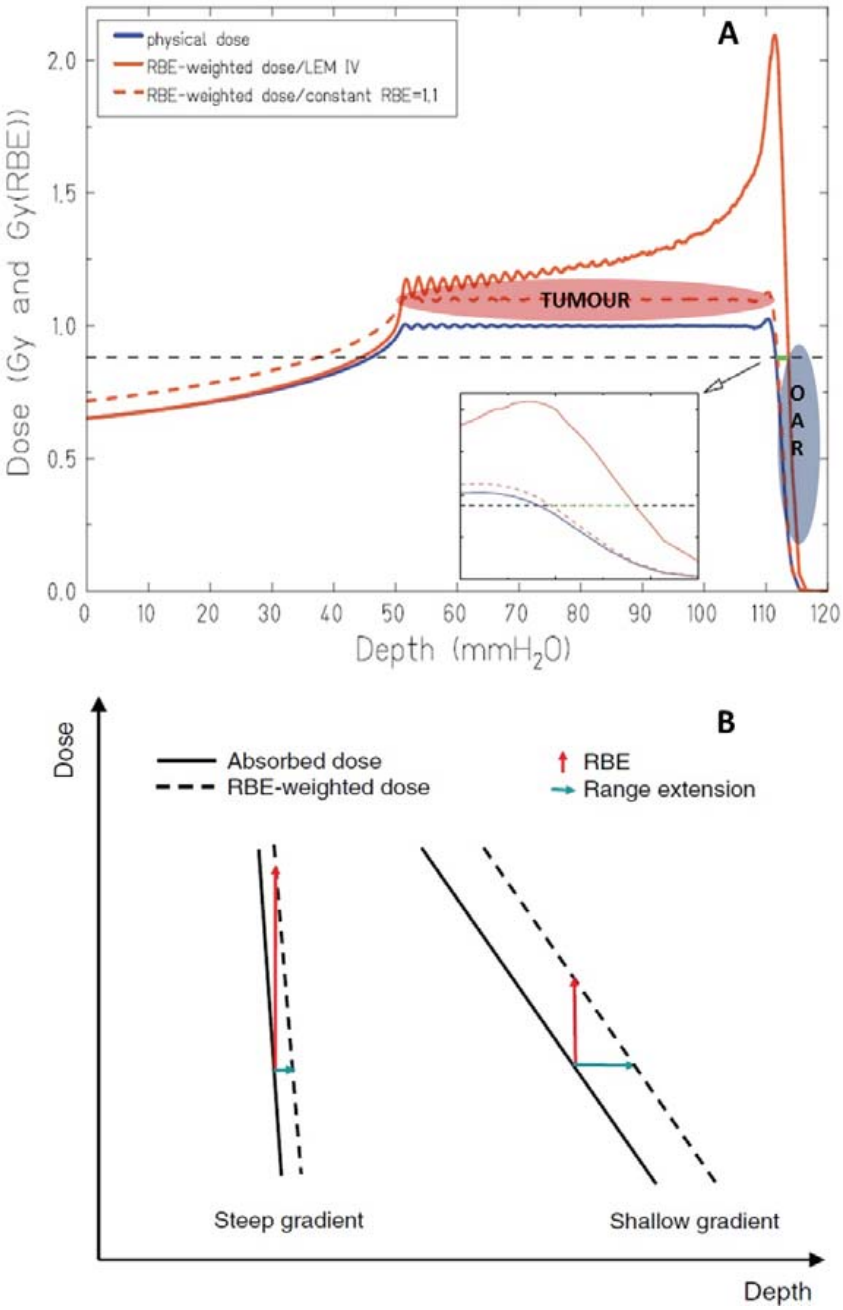


Figure 3.16: Variable RBE values produce a increase of the local RBE value towards the distal range (top). If such a model is assumed for the dose computation, we see that there is an effective range extension (bottom). Obviously this has serious clinical repercussions. In addition, many PG cameras study the range with similar precision as the shift due to the choice of RBE models, so it seems important to know well which predicts the biological dose best. Source: Tommasino and Durante (2015).

Chapter 4

Accelerated Prompt Gamma estimation for clinical Proton Therapy simulations

In this chapter the effort on the accelerated simulation of PG estimates is presented. The experimental difficulties with low yields discussed in the previous chapter pose a problem for simulations as well: converging on an acceptable estimate of a rare process requires much computation and therefore long wall-times, even with the aid of cluster computing. Details on these difficulties are also discussed in the paper, which was published in 2016 in *Physics in Medicine and Biology* (Huisman et al., 2016). After the paper, some notes will be discussed, adding some detail and implications about the implementation. In addition, an alternative recently published PG estimation method will be briefly discussed.

Accelerated Prompt Gamma estimation for clinical Proton Therapy simulations

Brent F. B. Huisman[1,2], J. M. Létang[1], É. Testa[2], and D. Sarrut[1]

[1] CREATIS, Université de Lyon; CNRS UMR5220; INSERM U1206; INSA-Lyon; Université Lyon 1; Centre Léon Bérard, Lyon, France
[2] IPNL, Université de Lyon; CNRS/IN2P3 UMR5822; Université Lyon 1 Lyon, France
E-mail: brent.huisman@creatis.insa-lyon.fr

September 27, 2016

Abstract

Purpose: There is interest in the particle therapy community to use prompt gammas (PG), a natural byproduct of particle treatment, for range verification and eventually dose control. However, PG production is a rare process and therefore estimating PGs exiting a patient during a proton treatment plan executed by a Monte Carlo (MC) simulation converges slowly. Recently, different approaches to accelerating the estimation of PG yield have been presented. Sterpin et al. (2015) described a fast analytic method which is still sensitive to heterogeneities. El Kanawati et al. (2015) described a variance reduction method (pgTLE) that accelerates the PG estimation by precomputing PG production probabilities as a function of energy and target materials, but has as drawback that the proposed method is limited to analytical phantoms.

Materials and Methods: We present a two-stage variance reduction method, named voxelized pgTLE (vpgTLE) that extends pgTLE to voxelized volumes. As a preliminary step, PG production probabilities are precomputed once and stored in a database. In stage one, we simulate the interactions between the treatment plan and the patient CT with low statistic MC to obtain the spatial and spectral distribution of the PGs. As primary particles are propagated throughout the patient CT, the PG yields are computed in each voxel from the initial database, as function of the current energy of the primary, the material in the voxel and the step length. The result is a voxelized image of PG yield, normalized to a single primary. The second stage uses this intermediate PG image as a source to generate and propagate the number of PGs throughout the rest of the scene geometry, e.g. into a detection device, corresponding to the number of primaries desired.

Results: We achieved a gain of around 10^3 for both a geometrical heterogeneous phantom and a complete patient CT treatment plan with respect to analog MC, at a convergence level of 2% relative uncertainty in the 90% yield region. The method agrees with reference analog MC simulations to within 10^{-4}, with negligible bias. Gains per voxel range from 10^2 to 10^4.

Conclusion: The presented generic PG yield estimator is drop-in usable with any geometry and beam configuration. We showed a gain of three orders of magnitude compared to analog MC. With a large number of voxels and materials, memory consumption may be a concern and we discuss the consequences and possible trade-offs. The method is available as part of Gate 7.2.

1 Introduction

The well-defined range of particles in matter is the main reason they are used in cancer treatment today. Unfortunately we are not able to take full advantage of this property, because of uncertainties in patient positioning, uncertainties on the proton range due to unknown displacements or deformations of organs, ill-defined lung, bowel or bladder filling, and the inherent uncertainty in the Houndsfield unit to particle stopping power conversion. Often, medical practice is to plan conservatively, namely adding margins around the tumor, greatly reducing the potential benefits of particle treatment (Knopf and Lomax, 2013). Some form of in-vivo, online monitoring is generally considered to be a way out of this predicament. Online monitoring would make measurements of uncertainties such as mentioned above possible, and thereby permit more precise planning which could take maximum advantage of the steep Bragg Peak (BP) fall-off and reduce damage to tissues surrounding the tumor.

One way to perform monitoring is to use prompt gammas (PGs), a natural byproduct in particle treatments. PGs offer the potential of monitoring treatment at the spot level (Roellinghoff et al., 2014; Smeets et al., 2012) and are therefore of prime interest (Golnik et al., 2014; Gueth et al., 2013; Janssen et al., 2014; Moteabbed et al., 2011). These particles are produced in the inelastic nuclear collisions between the incident particle and the medium (tissue) it is traveling through, and they correlate very well with dose deposition. This has been experimentally demonstrated for protons by Min et al. (2006) and for carbon by Testa et al. (2008). The latter also showed that adding Time of Flight (ToF) measurement enables the discrimination between neutron-induced gammas and PGs, greatly cleaning up the signal, which was confirmed by Biegun et al. (2012); Lopes et al. (2015); Pinto et al. (2015); Roellinghoff et al. (2014); Smeets et al. (2012); Verburg and Seco (2014). The authors conclude that online range monitoring is possible, once a detector with sufficiently large solid angle can be realized. In Fall 2015, at OncoRay in Dresden, Germany, a knife-edge PG camera (Perali et al., 2014; Richter et al., 2016) was put into clinical operation. In addition to geometrical collimation, work is being done on designs that exploit Compton scattering to record the PG signal (Kurosawa et al., 2012; Roellinghoff et al., 2011)

Another approach is to use the fact that in certain nuclear reactions PGs with specific energies are produced (Verburg and Seco, 2014; Verburg et al., 2013). The composition of the produced PG spectrum is dependent on material and proton energy. Certain PG energies tend to be produced close to the BP; that is to say, the associated cross sections are largest when the primary is near the end of its range. Analysis of these spectral lines at given positions may provide sufficient information to deduce both the material composition and BP position. Verburg et al. exploit this fact to reconstruct the spatial location of the BP falloff position with a single collimated detector placed at the end of the primary range. Then there is an attempt (Golnik et al., 2014) to use the travel time of the primary to reconstruct and control the PG range. Using a camera with high timing resolution and a short, pulsed beam, the ToF between the proton leaving the beam nozzle and the PG arriving in the camera can be recorded, and by putting the camera close to the nozzle looking back at the patient, we know that a larger ToF corresponds to more distal interactions, and a smaller ToF to proximal interactions.

Imaging paradigms such as PG detection are validated against experiments, and often also with Monte Carlo (MC) simulations (Golnik et al., 2014; Gueth et al., 2013; Janssen et al., 2014; Moteabbed et al., 2011; Robert et al., 2013). Conventional MC methods propagate particles according to a set of physical processes through materials. The tracking of a particle is broken up into steps, where at each point, a weighted list of all possible next steps is built and one option is selected by a random number. For rarely occurring processes, convergence to the model of the truth to within acceptable statistical error can be slow. PG emission in particle therapy, when viewed on a voxel-by-voxel basis, is also a rare and slowly converging process (Gueth et al., 2013; Pinto et al., 2015; Sterpin et al., 2015). This has important implications: firstly for detector designers, secondly for those who simulate, and thirdly for those interested in comparing the two online in, say, clinical conditions. In the first case, detectors are optimized to minimize signal loss (see fig. 1) and advanced reconstruction can be employed to maximize the use of information in the signal. Gueth et al. (2013) has demonstrated a method that works around the low PG yield by using machine learning to correlate predefined patient translations (setup errors) to PG output signals, which reduces the time to produce an estimated translation based on the detected PGs. Since convergence to the model of the truth requires long simulation runtimes, we may compensate with variance reduction methods or cutoffs in the time, space, or spectral domains (e.g. fixed runtime, larger voxel size, larger spectral bins). One such variance reduction method is MC splitting, where the moment a rare process occurs not one, but multiple possible futures for that process are computed. The weighted total of these futures are then stored, and so the convergence accelerated.

Recently, a variance reduction method (pgTLE) was described (El Kanawati et al., 2015). The PG estimation is accelerated by precomputing PG production probabilities as a function of energy and target materials, with as drawback that it only works for analytical phantoms. Here, we present a two-stage method, voxelized pgTLE (vpgTLE), that extends pgTLE to voxelized volumes, making it useful for clinical simulations. Next we describe the method and give some simulation results in various configurations. At this time, only protons as primaries have been considered and validated. We finish with a discussion of background noise estimation, other variance reduction methods and points for improvement.

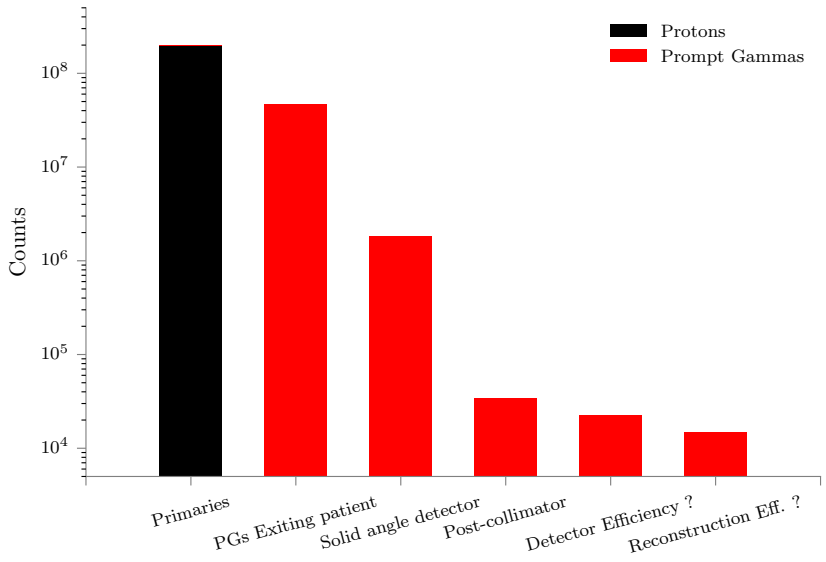


Figure 1: Approximate progressive loss of PG signal per spot. The particles were recorded in a simulation with a clinical head and neck CT and associated clinical plan, with a 30-by-30 cm collimator with 43% fill factor, at 40 cm from the patient, with estimated efficiencies (66%) for detection and reconstruction. Note this is not the configuration investigated in the rest of this paper.

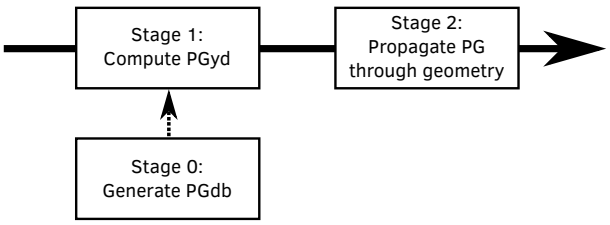


Figure 2: Flow chart of the vpgTLE method. Stage 0 is an initial PG database (PGdb) and is computed once. Every subsequent simulation is broken up into Stage 1 and Stage 2. Stage 1 generates a CT specific PG yield distribution (PGyd) using a limited number of primaries. In Stage 2 the PGyd is used to generate and propagate a representative set of PGs throughout the geometry.

2 Materials and Methods

2.1 A voxelized pgTLE

A full *voxelized Prompt-Gamma Track Length Estimator* (vpgTLE) simulation is broken up into two stages (fig. 2). The process presumes the existence of a prepared database (PGdb) which is an estimate of the effective linear PG production coefficient modulo the density, per element ($\frac{\mathbf{\Gamma}_Z}{\rho_Z}$, see eq. 1), per PG energy bin per primary energy bin. The PGdb is computed once for a list of common elements, and can then be reused. In stage one, a PG yield distribution image (PGyd) is created, specific to a particular phantom (or CT-image) and primary source (e.g. a treatment plan, a single spot). This PGyd image stores, per voxel per PG energy bin, the yield per primary. Stage two uses the PGyd and the assumption of isotropic PG emittance to generate and propagate the PGs throughout the rest of the geometry that the user has defined (e.g. the PG detector).

2.1.1 Stage 0: PG Database

First, as explained in El Kanawati et al. (2015), we precalculate a database of PG production probability per PG energy, as a function of element and proton energy. Eq. 1 describes the computed quantity: the PG spectrum $\mathbf{\Gamma}$ as a function of proton energy E for a set of elements Z. We take the ratio of the number of produced gammas $\boldsymbol{N}_\gamma$ over the number of inelastic collisions N_{inel}, scaled by the linear element attenuation coefficient κ_{inel}, related to the proton inelastic nuclear process. We normalize with the density ρ of element Z. Bold symbols represent vectors as function of PG energy. We compute the PGdb for protons up to 200MeV in 250 bins.

$$\frac{\mathbf{\Gamma}_Z(E)}{\rho_Z} = \frac{\boldsymbol{N}_\gamma(Z,E)}{N_{inel}(Z,E)} \frac{\kappa_{inel}(Z,E)}{\rho_Z} \tag{1}$$

Currently, PG emission models and cross sections implemented in various MC codes (Geant4, Fluka, MCNPX) are still evolving as differences have been observed between experimental PG data and simulations (Pinto et al., 2015), and between MC codes (Pinto et al., 2016; Robert et al., 2013; Verburg et al., 2012) and research is ongoing to improve accuracy. Hence, if cross sections or implementations of models are updated, the database must be recomputed. In particular, when comparing between simulations, such as in this study between vpgTLE and a reference analog MC, the same physics list must be used.

2.1.2 Stage 1: Compute phantom-specific PG yield distribution

By performing a low statistic MC, an estimate of the proton track lengths per voxel per proton energy bin is obtained, hence "Track Length Estimation". Note that "track" corresponds to "step" in Geant4 terminology. Together with the effective linear PG production coefficient, the PGyd image is obtained.

Before the simulation starts, we query the CT image for a list of present materials and build $\mathbf{\Gamma}_m$ for each material m encountered. We compute the PGdb in terms of elements, rather than materials, to give the user the option to add new materials or modify the composition of existing ones without recomputing the PGdb. In eq. 2, we sum $\mathbf{\Gamma}_Z$ over all k constituent elements Z in material m_v, weighted by density fraction ω_k of element Z in m_v, and scaled by density ρ_{m_v}.

$$\mathbf{\Gamma}_m(E) = \rho_{m_v} \sum_{k=1}^{k_{m_v}} \omega_k \frac{\mathbf{\Gamma}_{Z_k}(E)}{\rho_{Z_k}} \tag{2}$$

Then, a limited number of particles is propagated from the source of the treatment plan into the target, typically a patient CT image. We define a 4D scoring matrix in the CT (the PGyd), which may have a different size and pixel spacing from the CT. Protons are propagated with an unmodified analog MC tracking engine. Per step, per voxel v in the PGyd, alongside executing the unmodified analog MC processes, we compute and add the product of the step length L_g and $\mathbf{\Gamma}_{m_v}$ (linearly dependent on proton inelastic cross section, see eq. 1), with m_v the material at voxel v and g the proton energy bin, according to eq. 3. Put into words, we compute the PG yield probability energy spectrum at every step, and add it to the yield of voxel v of the current step.

$$\widehat{\boldsymbol{S}}_i(v) = \mathbf{\Gamma}_{m_v}(E_g) L_g(E_g, v) \tag{3}$$

At the end of the simulation, we have accumulated the yield spectra per voxel v. The computed output is weighted by the number of primaries to obtain the final PG production probabilities per voxel per PG energy

bin. The PGyd written to disk is therefore *per primary*, and the sum of all the probabilities is the probability that a single primary particle will emit a PG.

In order to obtain the variance in this paper, we opted for the classic batch technique. Although El Kanawati et al. provide an analytical derivation for the variance, they assume no correlation between proton energy and track lengths. We observed that this assumption does not seem to hold, in that the result is quite different from the variance obtained with the batch technique. Derivations assuming partial and full correlation are possible, but we felt that the short run-times of the vpgTLE method, coupled with the simplicity of the batch technique, and the innate correctness of the variance obtained in such a way, were the best choice for understanding and presenting this method. Note that this assumes the initial database computed in stage 0 has converged sufficiently and does not contribute significantly to the variance.

2.1.3 Stage 2: Propagate PG through other geometry (detectors)

The PGyd computed in stage 1 is used as a source of PG emission probability per primary particle. All that is left to complete the simulation (e.g. record PGs in a detector) is to produce as many PGs weighted by the source image as necessary. If, say, the user is interested in the PG signal of a 2 Gy fraction, and a 2 Gy fraction is composed of 10^{11} protons, the PGyd can be requested to give the expected output for that number of protons. Each PG is created and then propagated through the patient and into the detector, according to regular analog MC mechanisms. However, depending on the number of PGs the user requires in their detector, a lower number of PGs may be requested, scaled up, and consequently runtime is lowered.

An important consideration is that we currently assume that PGs are spatially emitted isotropically. Geant4 also adheres to this assumption. This conveniently relieves us from recording any spatial information. However, there exists evidence to the contrary (Sheldon and Van Patter, 1966; Verburg et al., 2012). Once any possible anisotropy is introduced in the MC physics models, it can be added to our code. Recording the anisotropy factor during stage 1 may be an intuitive solution (Henyey and Greenstein, 1941).

2.2 Validation procedure

Each simulation is executed with both analog Monte Carlo scoring and with the vpgTLE method. The analog MC serves as reference. To obtain an estimate of the statistical uncertainty, we employ the batch technique and run each type of simulation 10 times. When studying the bias and the relative uncertainty within selected subregions of the phantoms, σ is computed on the projection considered. That is to say, σ represents the standard deviation on the mean yield over the 10 simulations. For the study on efficiency and convergence of relative uncertainty, the variance is computed per voxel. Taking the median of (a subregion of) this 4D 'image of variance' (i.e. the median of the variance) provides a stronger test. We take the median rather than the mean because the variance distribution tends to a log-normal distribution. For skewed distributions such as this the median is a better measure of central tendency.

2.2.1 Test cases

Two test cases are presented. The purpose of Test case 1 is to verify that the transition to voxels has been done correctly. The phantom proposed by Parodi et al. (2005) and used by El Kanawati et al. (2015) is converted into a voxelized representation, see figure 3. In Test case 2 (fig. 4), a clinical head and neck image with corresponding proton therapy treatment plan is examined and is intended as a demonstration of the possibilities of the vpgTLE method. Precise beam properties may be found in table 1. Compared to El Kanawati et al. (2015), the number of analog primaries used for the reference are increased from 10^7 to 10^9. This is required in order to obtain a sufficiently noiseless figure for the spatial projection along the beam axis. For the vpgTLE simulations, four simulations are executed with 10^3, 10^4, 10^5 and 10^6 primaries respectively. Next, we define certain windows of interest. Knowing that most PG detectors do not measure outside the 1-8 MeV energy range (Testa et al., 2013), or even narrower (Smeets et al., 2012), we limit our analysis to this energy window. In addition, PG yield outside the spatial region that represents 90% of the total yield in the reference simulation is discarded. For all analyses these two cuts are applied.

2.2.2 Bias

To establish the presence of bias, as function of spatial or spectral dimensions, the relative difference with respect to the reference is investigated. In addition, certain subregions are studied separately for Test case 1,

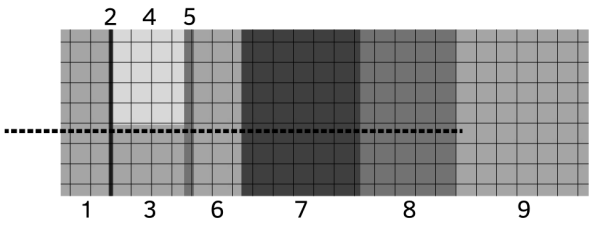


Figure 3: Top-down view of the Parodi phantom (Parodi et al., 2005), where the shading represents the material densities. Parts 1, 3, 6, 9 are PE; 7 is PMMA; 2 is Bone; 5, 8 are Muscle; 4 is Lung. The beam is illustrated with the dotted line coming in from the left. A voxelized version of this image is created, with a 2^3 mm^3 voxel size.

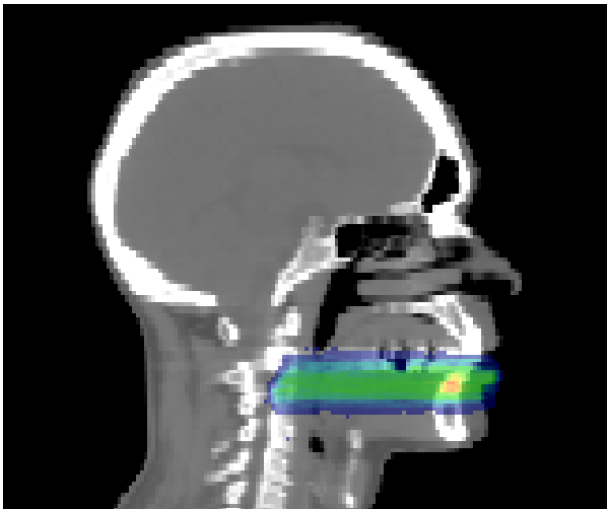


Figure 4: Sagittal view of the patient CT image, illustrated with the PG yield caused by the associated treatment plan. A beam from an original treatment plan has been rotated to align with the image axes, to both make projections easier and to increase the heterogeneity of the materials and thereby increase the challenge to vpgTLE. The applied radiation is the distal layer of one of the beams of the original plan so that the distal falloff is better defined in contrast to the spread out Bragg Peak (SOBP). The voxel size is again 2^3 mm^3.

because material-based regions can easily be isolated.

2.2.3 Efficiency, Gain and Convergence

An important quantity that characterizes a variance reduction method is the efficiency ϵ_k, and is computed by considering the time t required to reach a variance σ_k^2 per voxel k, see eq. 4. Comparing the ratio of efficiencies of two methods gives a quantified gain G (eq. 5, where TLE and A refer to vpgTLE and analog MC respectively). Using a measure of centrality (e.g. mean, median) on the gains per voxel G_k, a global measure for the efficacy of vpgTLE is obtained. A histogram of the gains within an image is presented to have an idea of the worst and best case performance of vpgTLE.

$$\epsilon_{TLE,k} = \frac{1}{t \times \sigma_{TLE,k}^2} \tag{4}$$

$$G = \frac{\epsilon_{TLE,k}}{\epsilon_{A,k}} \tag{5}$$

The proposed vpgTLE is a variance reduction method: it reduces the variance for any given PG simulation with respect to an analog simulation. As a simulation runs, the output converges, which is to say: its variance reduces. A common measure to ensure sufficient convergence is to require that the uncertainty σ associated to the quantity of interest is not more than 2% of the signal. A variance reduction technique such as vpgTLE translates into reaching the threshold faster, and therefore a gain with respect to analog MC. We compute the gain therefore both by taking the ratio of the runtime of the two methods at the 2% convergence level, as well as by computing the gain (eq. 5). With the same data, we estimate the total runtime required to produce a sufficiently converged image.

2.2.4 Influence of voxel size

An essential property of vpgTLE is the fact it records PG production probabilities all along the primary's path, instead of waiting for the MC engine to produce a PG. This means that vpgTLE requires much fewer propagated primaries to touch all voxels. This effect could be magnified further when smaller voxels are used. Recent developments (Marcatili et al., 2014) in super-resolution or multi-scale CT models involve smaller voxels in order to increase simulation accuracy for smaller or thinner organs such as the rectum, bladder, or spine. We demonstrate the increased benefits of vpgTLE on Test case 2.

2.2.5 Hardware, software, parameters

Gate 7.1 (Jan et al., 2004; Sarrut et al., 2014) with Geant 4.10.01 and the QGSP_BIC_HP_EMY physics list, commonly used for PG studies, are used in this analysis.

In Gate, scorers are defined as actors, which can be attached to volumes, all defined through the Gate macro language. The vpgTLE method was implemented as an actor, that keeps a PG spectrum for each voxel, the number and volume of which can also be controlled through the macro language. A helper actor, that outputs the analog MC in a similar manner as the PGyd was implemented to facilitate analysis. The output was validated against the GatePhaseSpaceActor, a thoroughly used and tested part of Gate. We used a gamma production cut of 3 keV in order to remove a high number of photons that cannot exit the phantom or patient geometry.

Timing information is obtained with the aid of the GateSimulationStatisticActor, executed on a Intel Core i7-3740QM CPU @ 2.70GHz, SpeedStep off, whilst using a single core. In summary, table 1 lists the main parameters used for the simulations.

3 Results

3.1 Test case 1

First we verified that vpgTLE yields are identical to the results produced with pgTLE, shown in El Kanawati et al. (2015). Then, we compared our method to the analog reference. Figure 5 depicts the yield on the first row, as a function of the depth (left column) and the PG energy (right column). The relative difference of the

Table 1: Summary of vpgTLE analysis parameters

	Test case 1	Test case 2
Beam	160 MeV, disc shaped cross-section, Gaussian spatial and angular distr. with σ of 3.5 mm, 2 mrad respectively	Distal layer (133.08 MeV) of clinical treatment plan (7 spots)
Phantom description	Parodi et al. (2005)	Clinical head and neck CT
Phantom voxel size	2mm	
PGyd voxel size	2 mm	1, 2, 5 mm
PGdb primaries/element	10^9	
PGdb max. proton energy	200 MeV	
Number of PG bins	250	
Number of proton bins	250	
vpgTLE primary-sets	$10^3 - 10^6$	
Analog primary-sets	$10^6 - 10^9$	
Reference	Analog primary-set 10^9	
Studied projections	Spectral and along beam	
Extra studies	Central voxel line yield	Influence of PGyd voxel size
Spatial window	90% yield region	
Spectral window	1-8 MeV	
Variance computation	10 batches per primary-set	
Step size	1 mm	

PG yield is shown, both integrated over the entire coronal plane (second row) or at the voxel line on the beam path (third row). Row one, a plot of the yield, shows a perfect overlap of vpgTLE with respect to the reference. We must look to the relative differences of the various vpgTLE outputs with respect to the reference in row 2 to observe any differences. The shaded areas represent 2σ error bands.

With 10^3 primary particles, the mean is noisiest, as expected. An overestimate beyond 170 mm is visible, which is about the location of the Bragg Peak. The average relative difference over the depth is 4.0×10^{-4} along the beam, which is a good performance, but due to the relative difference with the reference exceeding 1% in the distal region and the very wide error bands, we would argue 10^3 primaries is insufficient for a reliable prediction. The distal systematic shift has reduced when using 10^4 or more primaries. Two regions with bias remain: a consistent overestimate of about 0.5% at around 160 mm depth, and then, past the Bragg Peak, an erratic mean with wide uncertainty bands. The latter can be explained by nuclear events. Once a proton collides and is absorbed, it can no longer produce PGs. Towards the end, the precise number of remaining protons grows more uncertain, and just 10^3 primaries are not enough for a good estimate of the variance. Increasing the number of primaries reduces the uncertainty and improves the mean yield, but the effect remains. The average relative difference over the depth is of the order of 10^{-4} for all primary sets.

The spectral column on the right demonstrates vpgTLE is close to the analog reference over the whole spectrum, with a small fluctuation at the high end of the spectrum. The pattern present for all primary sets must be due to the PGdb, and is the only bias we observe. The average relative difference varies between -1.3×10^{-4} with 10^3 primaries and 6.0×10^{-5} for 10^6 primaries. This supports the hypothesis that we have converged to the bias introduced by the uncertainty of the PGdb. The wide error bands for 10^3 are again visible, with the error bands for 10^4 or more primaries staying within 1% of the mean over the entire range.

The bottom row, the line of voxels centered on the beam path, shows a more erratic behavior. One major difference is the proximal overestimate and distal underestimate with 10^3 primaries. With 10^5 primaries or less, the average relative difference is on the order of 10^{-3} or more, while 10^6 primaries results in 3.6×10^{-4}. The uncertainty is, naturally larger. The spectral view is stable and the depth view has an increased variance towards the end of the proton range.

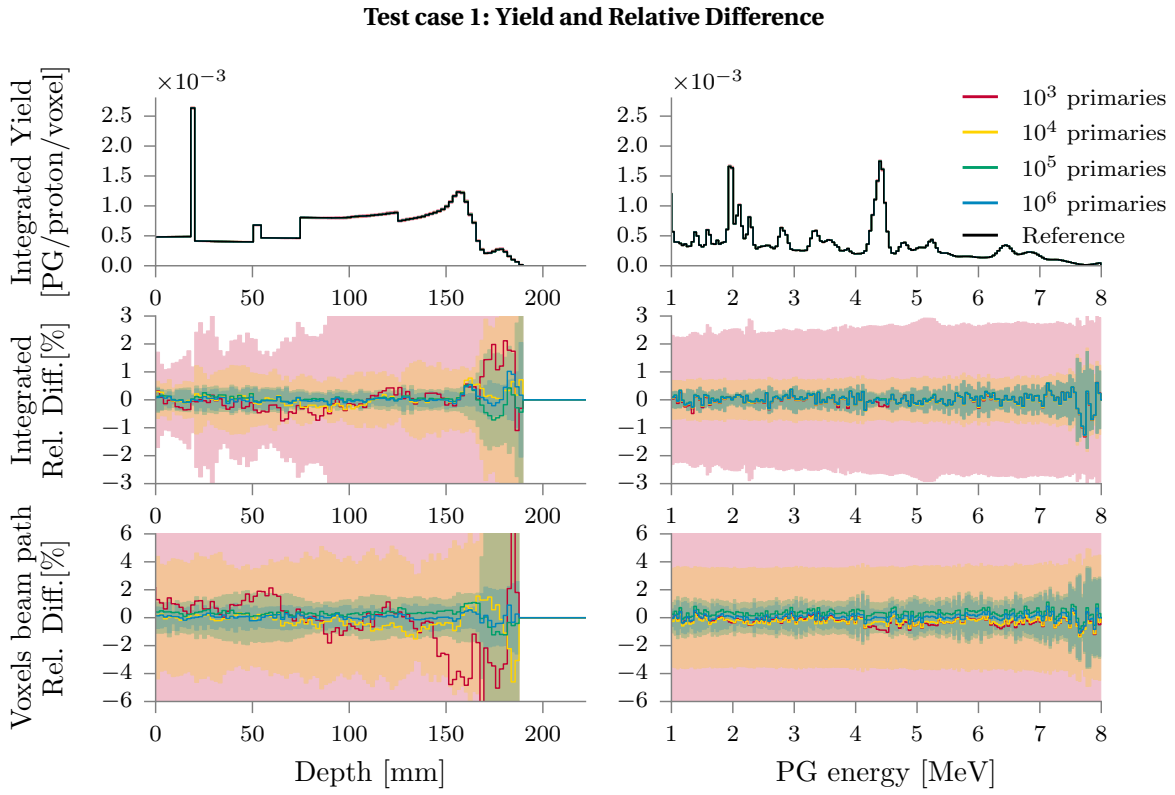


Figure 5: Test case 1: Row 1 shows the PG yields and row 2 the relative difference with respect to the reference ($\frac{vpgTLE-Reference}{Reference}$). The yield corresponds to the mean over 10 simulations. For both rows, note that the yield was integrated over all other dimensions. Row 3 shows the relative difference on the line of voxels on the center of the beam path, where we did not integrate over all other dimensions. The left column is a projection along the beam-axis, while the right column shows the spectral bins integrated over all voxels considered. The shaded areas represent 2σ error bands, where σ is the standard deviation over the mean of 10 simulations. Note that the covarying pattern in the relative difference is due to the noise of the analog signal, and does not represent any issue with the vpgTLE implementation.

Figure 6 shows side-by-side the convergence of the median relative uncertainty and a histogram of the gains computed with eq. 5. We see that a median convergence to within 2% is reached in about 3 minutes and about 68 hours with vpgTLE and analog MC respectively. At the 2% convergence level, the gain is 1.55×10^3. The histograms on the left show that the gains are stable with respect to the number of vpgTLE primaries. That means vpgTLE has no systematic problems. We can clearly see the skew of the distributions (note the logarithmic scale on the x-axis). The worst gain is a factor of 6.19×10^1, while the best voxel clocks in at 5.21×10^4, with a median of 1.40×10^3.

Test case 1: Gain distribution and Convergence

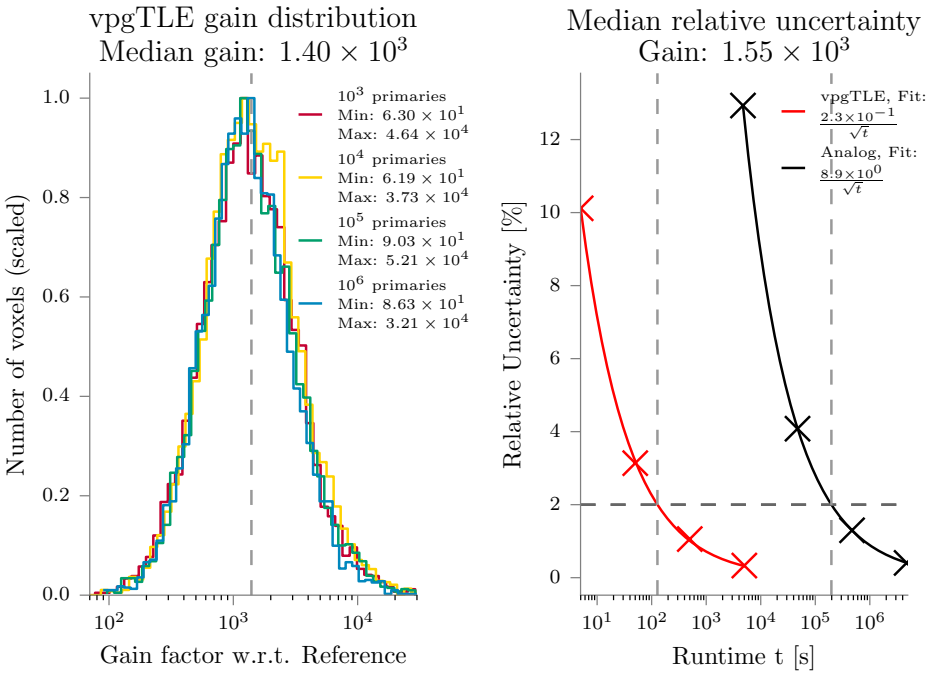


Figure 6: Test case 1. Left, the efficiency is computed, per voxel, according to eq. 4, for all vpgTLE primary-sets with respect to the reference. Right, the median relative uncertainty as a function of runtime t, for both the analog and vpgTLE methods. Each successive point is generated with 10^3 - 10^6 primaries for vpgTLE, and with 10^6 - 10^9 primaries for analog MC. As t increases, the relative uncertainty decreases as $\frac{C}{\sqrt{t}}$, where C is a fit factor. To compute the gain, we take the ratio of the runtimes at the 2% level, indicated by the dashed horizontal line, generally considered to be sufficiently converged.

3.2 Test case 2

Look to figure 7 for the yield and relative difference of vpgTLE with respect to the reference. The width of the 2σ-bands has increased with respect to Test case 1. Along the beam, left column of the figure, we see that 10^3 primaries produce an erratic line, while 10^5 and up are close to 0%. Past the distal end, we see significant divergence as in Test case 1. While the average bias is of similar magnitude as in Test case 1 (10^{-4}, except for 10^3 primaries), on the distal end the bias has not quite dissipated with 10^6 primaries. A likely explanation for the worse performance is the different beam with respect to Test case 1: now we look at the yield caused by a full energy layer composed of 7 spots, resulting in the primaries being spread over a larger volume and result therefore in lower statistics per voxel. We can again attribute part of the increase in error to the systematic error induced by the PGdb. The spectral view is as stable as in Test case 1, which is spread over the same 250 bins, and differs only by wider error bands.

Figure 8 shows the convergence and gain of vpgTLE for Test case 2. The gain is slightly less than in Test case 1. A sufficiently converged PGyd now requires little over 4 hours on a single core with vpgTLE and about 180 days with analog MC. Excluding the 10^3 primary-set because of its outliers, the worst gain is 2.70×10^1 and the best is 8.96×10^4, with a median gain of 9.98×10^2, and a gain of 1.03×10^3 at 2% relative uncertainty.

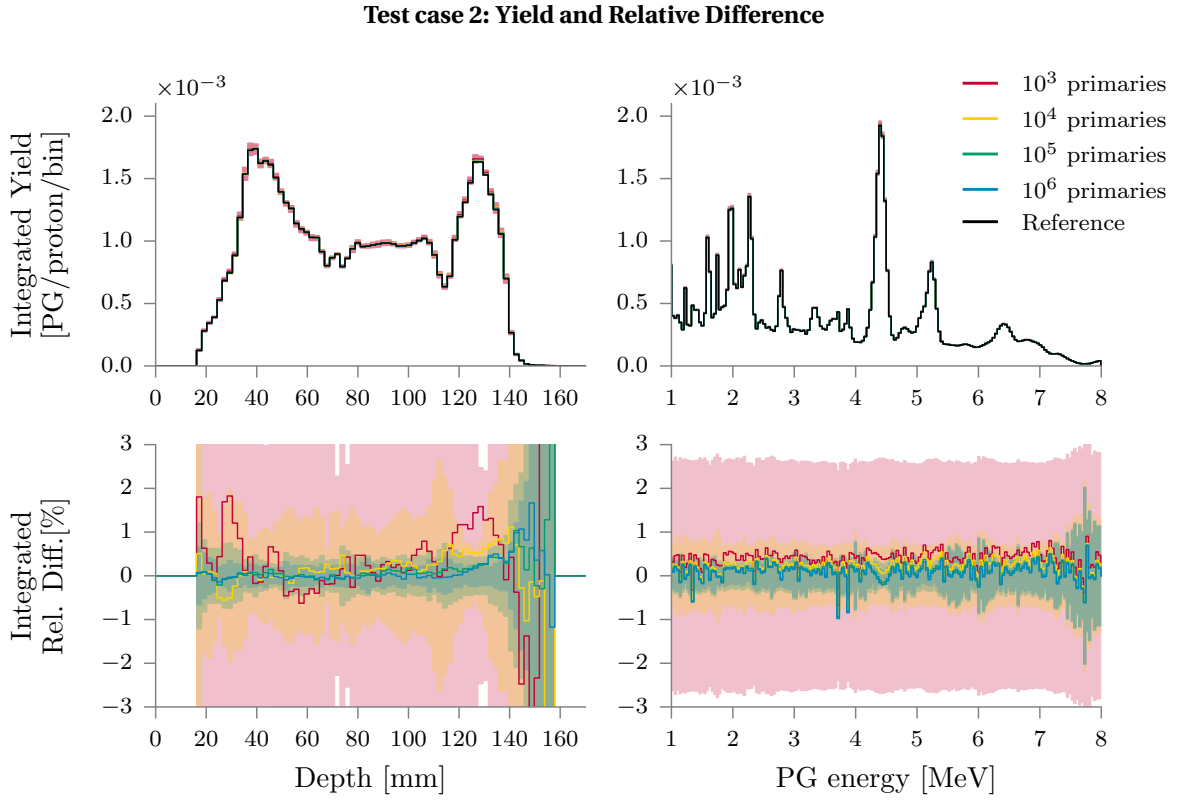


Figure 7: Test case 2. Row 1 shows the PG yields and row 2 the relative difference with respect to the reference ($\frac{vpgTLE-Reference}{Reference}$). For both rows, note that the yield was integrated over all other dimensions. The left column is a projection along the beam-axis, while the right column shows the spectral bins integrated over all voxels considered. The shaded areas represent 2σ error bands, where σ is the standard deviation over the mean of 10 simulations.

Test case 2: Gain distribution and Convergence

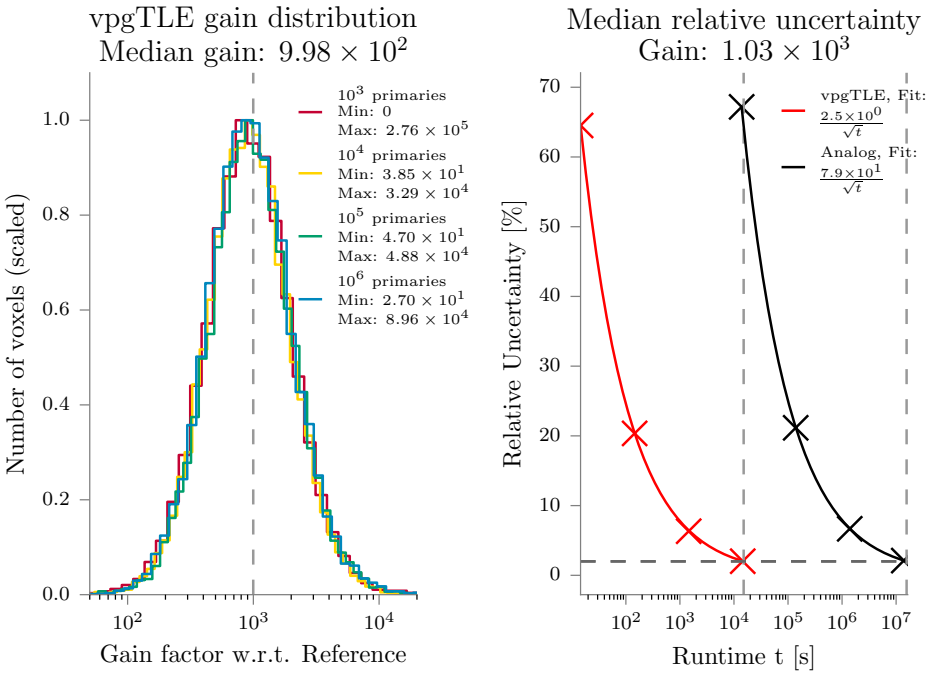


Figure 8: Test case 2. Left, the gain histogram is shown, for all vpgTLE primary-sets with respect to the reference. Right, the mean relative uncertainty is plotted as a function of runtime, for both the analog and vpgTLE methods. Each succesive point is generated with 10^3 - 10^6 primaries for vpgTLE, and with 10^6 - 10^9 primaries for analog MC. We take the ratio of the runtimes at the 2% level to obtain the gain.

The last result is the effect of changing the PGyd voxel size on the gain. We have observed in figures 6 and 8 that the gain distribution is independent of the number of primaries. Hence, in order to save time, we conduct this investigation with 10^7 primaries for analog MC and 10^4 for vpgTLE, which are at a similar level of convergence (see fig. 8). The gain ratio is computed as the ratio between these sets, per voxel size. Due to memory consumption considerations, the minimum voxel size was set at 1 mm^3, which results in an image of 1666 MB (833 MB on disk).

Figure 9 shows the gain histograms for each voxel size. It can be seen that the advantage of vpgTLE magnifies with respect analog MC as voxel size decreases. This is in agreement with the hypothesis that vpgTLE needs fewer primaries with respect to analog MC to estimate the PG yield. Users interested in super-resolution output can expect higher gains than reported for our two test cases.

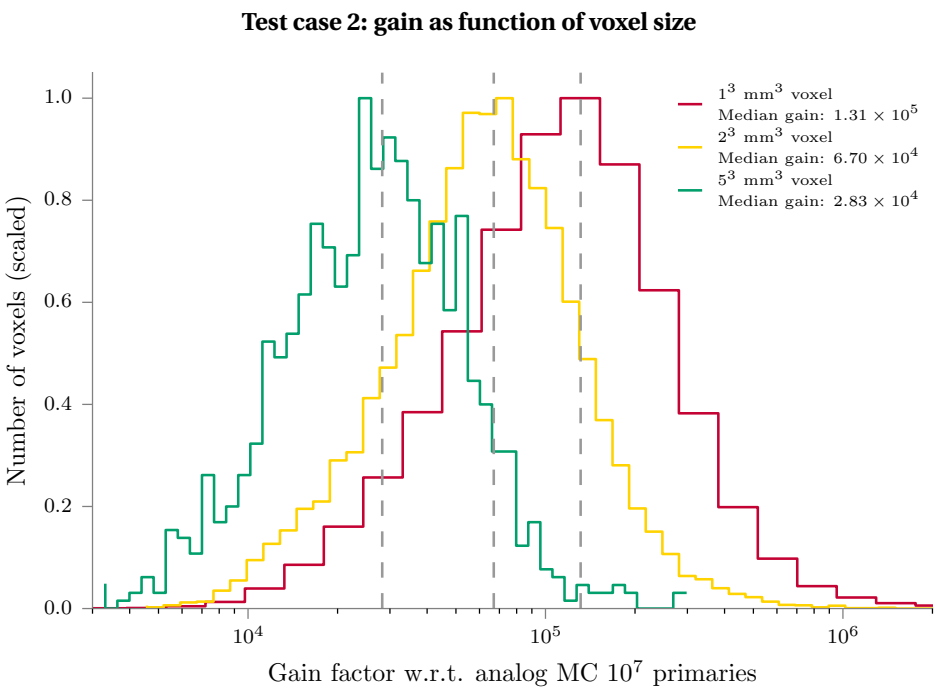


Figure 9: The distribution of gains of a vpgTLE simulation (10^4 primaries) w.r.t. an analog MC (10^7 primaries) are plotted. The distribution, and the medians, shift up as the voxel size decreases.

4 Discussion

4.1 Tradeoffs

The current implementation of vpgTLE stores each bin as a double (64 bit) in memory, and converts to float (32 bit) when writing to disk. The memory consumption is therefore linked to the number of PG bins, image and voxel size. By default the vpgTLE actor will copy the size and voxel size of the image it is attached to. As described before, for clinical CT images this will result in on-disk images of tens of GBs, and twice that in memory usage during the simulation. In this paper, we therefore shrank the PGyd volume to a region that envelops the PG production sites with some margin (see figure 10). This resulted in an on-disk image size of about 104 MB. With 1 mm^3 voxels, the image size increases eight-fold to 833 MB (double at runtime: 1666 MB). A PGyd with the size of the CT data used in Test case 2 with 1 mm^3 voxels would blow up to 120 GB on disk, 240 GB in memory at runtime.

Storing the intermediate PGyd is similar to the practice of storing intermediate phasespaces in complex accelerator simulations. A nice side-effect of having two stages is that if the user for instance wants to reposition a detector or compare different detectors altogether, only the PGs have to be re-propagated.

Before Stage 1, a number of important parameters are set: the number of primaries per element in the PGdb, the minimum, maximum and number of bins for both the primary and PG energy. Naturally, more

Test case 2: PGyd output region

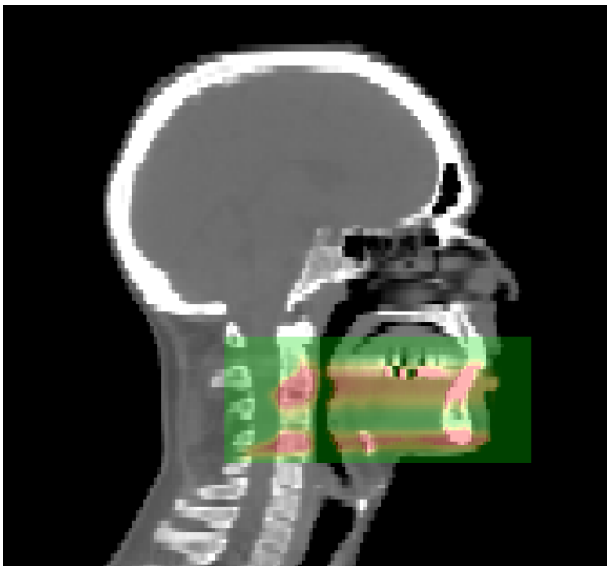


Figure 10: The region of the PGyd was set to a smaller region than the patient CT image, in order to reduce memory consumption. The region of the PGyd is visible as green, with the PG production visible as red bands.

primaries increases the quality of the estimate, while more bins spanning a wider range improve the precision, but slow down the convergence to an acceptable mean or median uncertainty. Assuming a maximum primary energy of 250 MeV, we need 250 bins for a precision of 1 MeV. The current implementation has linear binning, so assuming the location of the BP fall-off is of interest, where the primary energy is lowest, a 1 MeV bin translates to a proton range in water of about 24 μm, more than enough considering the typical 2^3 mm^3 voxel size. Note that computing the PGdb for more bins requires more particles to ensure proper bin filling. It took approximately 1000 days of CPU-time to compute the PGdb used in this study.

4.2 Comparison with other variance reduction techniques

A conventional approach to variance reduction for rare processes is interaction biasing (IB), where the probability of the interaction of interest is multiplied with factor α, and is compensated for by decreasing the weight of the continuing track (and secondaries) with factor $\frac{1}{\alpha}$. Parameter α is then chosen such that an interaction occurs once per interval of interest (say, once per voxel). Alternatively, interaction forcing (IF) forces an interaction per interval, and weights any subsequent interaction with the probability that former interaction occurred. As the incident particle may be killed in the process of interest (as is the case for PG production), some implementations (e.g. MCNPX, Geant4, EGSNRC) split the track into a collided and uncollided version to prevent the loss of statistics in the distal part of the track.

Another standard technique for rare processes is particle splitting: instead of producing a single new particle in the interaction of interest, N particles are produced, each with weight $\frac{1}{N}$. This method may be applied in addition to IB or IF. Storing the whole PG spectrum in Stage 1 and sampling it in Stage 2 could be viewed as delayed particle splitting. Not implementing the separation between Stages 1 and 2 would have been possible by directly generating a PG from the spectra computed in Stage 1 as the protons traverse the phantom.

When considering IB, the main disadvantage is the free parameter α, which should be chosen in such a way that sufficiently often a PG is produced. What sufficiently often means will depend on each individual simulation. Moreover, the PG prediction is still a probabilistic process: in some voxels a PG may be produced, in some may not. IF is much more similar to vpgTLE, in that it gives a deterministic prediction per interval. Having a larger number of calls to the physics processes is the main source of overhead for both IB and IF. For example, in an analog Geant4 simulation, we measure that 70% of computing time is spent in nuclear interactions. In a naive implementation of IB or IF, we run through the nuclear interactions an increased number of times, which indicates that the upper limit of the gain is $\frac{1}{70\%} \approx 40\%$. Both IB and IF may benefit from precomputed lookup tables to reduce the time spent performing the additional interactions. Lookup

tables are most practical if a single or limited number of outputs are sought, e.g. dose, PG.

Indeed, vpgTLE may be considered as a special optimized case of IF with the following differences:

- the Monte Carlo particle weights are not modified
- whole inelastic processes are precomputed instead of being called at each interaction
- the complete precomputed gamma energy spectrum is stored

The two last points can also be applied to IF implementations. Hence, PG yields per proton track remains essentially the same between IF and vpgTLE. The difference may essentially be viewed as conceptual, vpgTLE being inspired by the TLE approach originally put forward by Williamson (1987).

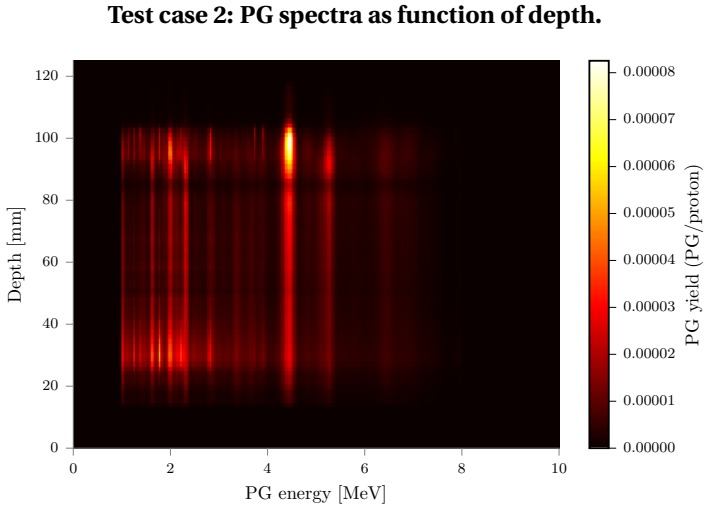


Figure 11: vpgTLE allows one to make full use of spatial and spectral information of PGs. As an illustration, this figure shows a heatmap of PG yields in Test case 2 as function of depth and PG energy, integrated over the transverse dimensions. The figure might incline a detector developer to tune a spectral camera to the 4.4 MeV peak.

As far as the authors of this paper could establish, the only other published variance reduction technique for PGs is described in Sterpin et al. (2015). This method is fully analytic, incorporating experimental or pretabulated PG emission data and a model that assumes the PG emission region can be modeled similar to the dose in pencil beams. The method raytraces the materials touched by a pencil beam spot, and computes the expected 1D PG profile by a weighted sum of pregenerated profiles per material, which takes 0.3 to 10 seconds. The authors admit that this approach does not deal very well with lateral inhomogeneity, a problem that vpgTLE does not have due to TLE methods in principle being assumption free. Another benefit of vpgTLE is the shape of its output: a 4D image where for each voxel a PG spectrum is recorded. This permits the incorporation of the method in many different kinds of PG detector simulations, not just detectors that measure the range. As a small example, vpgTLE can be used to investigate the origin of the global spectrum as function of depth along the beam path in figure 11. vpgTLE will work with any collimated or Compton camera design, PG spectroscopy, or any future detector that takes advantage of spatial or spectral components. Moreover, vpgTLE does not assume anything about the proton flux: it uses Stage 1 to estimate it. The difficulty estimating PG yields due to beam spots that produce different ranges due to lateral inhomogeneity is therefore avoided: vpgTLE is as sensitive to inhomogeneity as regular MC.

4.3 Background estimation

vpgTLE does not estimate other physical quantities than PGs produced in the target, which means no estimate of the background noise can be obtained with this method. For any detector seeking actual application, no analysis is complete without considering the signal-to-background ratio, and methods to improve it. The background consists of tertiary gammas, mainly produced by secondary neutrons undergoing nuclear interactions, in the target but also in elements of the beamline, other objects in the room such as the patient table,

and the detector itself (Pinto et al., 2014). Obtaining the background from simulation is problematic for two reasons:

1. No MC tool correctly estimates the background (see Pinto et al. (2014) section 3.1.2, Sterpin et al. (2015) section IV.A.4). Nuclear reaction models are continuously improved, but as far as we know not specifically for PG *background* at this time.
2. The computing time required for full room simulations is prohibitive, which is why room modeling is left out. Even without room components, the simulation runtime is still long using analog MC, as there are no variance reduction techniques for PG *background*.

Depending on the purpose of the simulation, various ways of dealing with the lack of reliable MC background prediction are considered. Sterpin et al. (2015), concerned with treatment prediction, deals with the background by adding an offset parameter to their fitting procedure of the PG profiles and accept the absence of background estimation in their fast PG method. A procedure like this could be applied for other devices. Simulations for camera optimization, effectively signal-to-background optimization for a multitude of configurations, also require a background estimate. Because the background may depend on camera parameters and obtaining a measured background for each configuration is not tractable, most groups limit the number of configurations. The multi-slit camera optimization of Pinto et al. (2014), which correctly does consider hundreds of configurations, was performed with a manual correction of a single background measurement.

Hence, the fact that in current practice the background is measured or modeled separately, puts vpgTLE at no particular disadvantage compared to other variance reduction methods or analog MC. However, if physics models could give an accurate estimate of the background, a generalized vpgTLE, most urgently of neutrons and neutron induced PGs, might be a possible avenue for variance reduction of the background estimate. The scorer could be attached to not only the target but other objects in the treatment room.

4.4 Future improvements

We did not take into account the systematic error (caused by the PGdb) in our analysis. The PGdb is computed by shooting protons of an energy at least as high as used in the treatment plan into a box of a certain material. This means that at low energies, the PGdb statistics are not as good as at higher energies, meaning that we have a systematic error around the BP region (where the proton energy is lowest). We might therefore consider to supplement our PGdb with the output of a second, low energy proton beam, simply to reduce the systematic error in the BP range. Since we compute the PGdb once, this has no effect on the efficiency of the vpgTLE method. It might also be possible to fill the database by querying Geant4 for the cross section at the respective bins. We did implement outputs for the analytic systematic and random error output as laid out in El Kanawati et al. (2015), but Stage 2 does not propagate these errors yet, which would be required for a quantitative analysis of the outputs of Stage 2. Moreover, we found that this analytic error estimate underestimates with respect to the batch method, because it assumes the independence of the track lengths and proton energies, a type of problem IF or IB techniques would not have. Therefore, at this time, the variance of vpgTLE can only be obtained by employ of the batch method.

A thorough analysis of the sparseness of the PGyd has not been conducted, but there is a likely opportunity for memory optimization here. Another option is to reduce the dimensionality of the image to that which the user requires. Users investigating their collimated camera may for example set the PG spectrum to a narrower window and a coarser binning, reducing memory consumption accordingly.

Precomputing an effective linear production coefficient, the main principle of vpgTLE, could be performed for other particles. Adding for instance an effective linear neutron production coefficient may supplement the vpgTLE output with a correct estimate of neutron-induced gamma noise in a PG detector, giving an indication of the background, not just the PG signal. However, such an addition must be investigated to obtain the real efficiency.

5 Conclusion

vpgTLE is a generic drop-in alternative for computing the expected PG output in voxelized geometries. The method has a fixed memory requirement (a 4D image) with a typical memory size of the order of a few hundred MB. The method reaches a global gain factor of 10^3 for a clinical CT image and treatment plan with respect to analog MC. A median convergence of 2% for the most distal energy layer is reached in approximately four hours

on a single core, at which point the output has stabilized to within 10^{-4} of an analog reference simulation, when the range or the spectrum is considered. The authors think the method is of interest to those developing and simulating PG detection devices, as well as clinicians studying complex clinical cases that require the precision and accuracy of MC level simulations not offered by analytic algorithms.

The vpgTLE method is open source and fully integrated in Gate. It is available from release 7.2 onwards.

6 Acknowledgements

This work was partly supported by Labex PRIMES ANR-11-LABX-0063, t-Gate ANR-14-CE23-0008, France Hadron ANR-11-INBS-0007 and LYric INCa-DGOS-4664. We thank Denis Dauvergne for his assistance with preparing the copy for this paper and Marc Verderi of the Geant4 collaboration for the discussions on variance reduction in Geant4.

References

Aleksandra K Biegun, Enrica Seravalli, Patrícia Cambraia Lopes, Ilaria Rinaldi, Marco Pinto, David C Oxley, Peter Dendooven, Frank Verhaegen, Katia Parodi, Paulo Crespo, and Dennis R Schaart. Time-of-flight neutron rejection to improve prompt gamma imaging for proton range verification: a simulation study. *Physics in medicine and biology*, 57(20):6429–44, oct 2012. ISSN 1361-6560. doi: 10.1088/0031-9155/57/20/6429. URL `http://www.ncbi.nlm.nih.gov/pubmed/22996154`.

W El Kanawati, J M Létang, D Dauvergne, M Pinto, D Sarrut, É Testa, and N Freud. Monte Carlo simulation of prompt γ-ray emission in proton therapy using a specific track length estimator. *Physics in Medicine and Biology*, 60(20):8067–8086, 2015. ISSN 0031-9155. doi: 10.1088/0031-9155/60/20/8067. URL `http://stacks.iop.org/0031-9155/60/i=20/a=8067?key=crossref.d38c03af2b13b06ae763a331d1d25dd2`.

Christian Golnik, Fernando Hueso-González, Andreas Müller, Peter Dendooven, Wolfgang Enghardt, Fine Fiedler, Thomas Kormoll, Katja Roemer, Johannes Petzoldt, Andreas Wagner, and Guntram Pausch. Range assessment in particle therapy based on prompt γ-ray timing measurements. *Physics in Medicine and Biology*, 59(18):5399–5422, sep 2014. ISSN 0031-9155. doi: 10.1088/0031-9155/59/18/5399. URL `http://stacks.iop.org/0031-9155/59/i=18/a=5399?key=crossref.5437fcd3059992135ec2113679c7dad6`.

P Gueth, D Dauvergne, N Freud, J M Létang, C Ray, E Testa, and D Sarrut. Machine learning-based patient specific prompt-gamma dose monitoring in proton therapy. *Physics in medicine and biology*, 58(13):4563–77, jul 2013. ISSN 1361-6560. doi: 10.1088/0031-9155/58/13/4563. URL `http://www.ncbi.nlm.nih.gov/pubmed/23771015`.

L. G. Henyey and J. L. Greenstein. Diffuse radiation in the Galaxy. *Astrophysical Journal*, 93:70–83, 1941. doi: 10.1086/144246.

S Jan, G Santin, D Strul, S Staelens, K Assié, D Autret, S Avner, R Barbier, M Bardiès, P M Bloomfield, D Brasse, V Breton, P Bruyndonckx, I Buvat, A F Chatziioannou, Y Choi, Y H Chung, C Comtat, D Donnarieix, L Ferrer, S J Glick, C J Groiselle, D Guez, P-F Honore, S Kerhoas-Cavata, A S Kirov, V Kohli, M Koole, M Krieguer, D J van der Laan, F Lamare, G Largeron, C Lartizien, D Lazaro, M C Maas, L Maigne, F Mayet, F Melot, C Merheb, E Pennacchio, J Perez, U Pietrzyk, F R Rannou, M Rey, D R Schaart, C R Schmidtlein, L Simon, T Y Song, J-M Vieira, D Visvikis, R Van de Walle, E Wieërs, and C Morel. GATE: a simulation toolkit for PET and SPECT. *Physics in Medicine and Biology*, 49(19):4543, 2004. URL `http://stacks.iop.org/0031-9155/49/i=19/a=007`.

F M F C Janssen, G Landry, P Cambraia Lopes, G Dedes, J Smeets, D R Schaart, K Parodi, and F Verhaegen. Factors influencing the accuracy of beam range estimation in proton therapy using prompt gamma emission. *Physics in medicine and biology*, 59(15):4427–41, aug 2014. ISSN 1361-6560. doi: 10.1088/0031-9155/59/15/4427. URL `http://www.ncbi.nlm.nih.gov/pubmed/25049223`.

Antje-Christin Knopf and Antony Lomax. In vivo proton range verification: a review. *Physics in medicine and biology*, 58(15):R131–60, aug 2013. ISSN 1361-6560. doi: 10.1088/0031-9155/58/15/R131. URL `http://www.ncbi.nlm.nih.gov/pubmed/23863203`.

Shunsuke Kurosawa, Hidetoshi Kubo, Kazuki Ueno, Shigeto Kabuki, Satoru Iwaki, Michiaki Takahashi, Kojiro Taniue, Naoki Higashi, Kentaro Miuchi, Toru Tanimori, Dogyun Kim, and Jongwon Kim. Prompt gamma detection for range verification in proton therapy. *Current Applied Physics*, 12(2):364–368, 2012. ISSN 15671739. doi: 10.1016/j.cap.2011.07.027. URL http://dx.doi.org/10.1016/j.cap.2011.07.027.

Patricia Cambraia Lopes, Enrico Clementel, Paulo Crespo, Sebastien Henrotin, Jan Huizenga, Guillaume Janssens, Katia Parodi, Damien Prieels, Frauke Roellinghoff, Julien Smeets, Frederic Stichelbaut, and Dennis R Schaart. Time-resolved imaging of prompt-gamma rays for proton range verification using a knife-edge slit camera based on digital photon counters. *Physics in Medicine and Biology*, 60(15):6063, 2015. ISSN 0031-9155. doi: 10.1088/0031-9155/60/15/6063. URL http://iopscience.iop.org/0031-9155/60/15/6063/article/.

S Marcatili, D Villoing, M P Garcia, and M Bardiès. Multi-scale hybrid models for radiopharmaceutical dosimetry with Geant4. *Physics in Medicine and Biology*, 59(24):7625–7641, 2014. ISSN 1361-6560. doi: 10.1088/0031-9155/59/24/7625. URL http://iopscience.iop.org/0031-9155/59/24/7625/article/.

Chul Hee Min, Chan Hyeong Kim, Min Young Youn, and Jong Won Kim. Prompt gamma measurements for locating the dose falloff region in the proton therapy. *Applied Physics Letters*, 89(18):38–41, 2006. ISSN 00036951. doi: 10.1063/1.2378561.

M Moteabbed, S España, and H Paganetti. Monte Carlo patient study on the comparison of prompt gamma and PET imaging for range verification in proton therapy. *Physics in medicine and biology*, 56(4):1063–82, feb 2011. ISSN 1361-6560. doi: 10.1088/0031-9155/56/4/012. URL http://www.ncbi.nlm.nih.gov/pubmed/21263174.

Katia Parodi, Falk Pönisch, and Wolfgang Enghardt. Experimental study on the feasibility of in-beam PET for accurate monitoring of proton therapy. *IEEE Transactions on Nuclear Science*, 52(C):778–786, 2005. ISSN 00189499. doi: 10.1109/TNS.2005.850950.

I Perali, a Celani, L Bombelli, C Fiorini, F Camera, E Clementel, S Henrotin, G Janssens, D Prieels, F Roellinghoff, J Smeets, F Stichelbaut, and F Vander Stappen. Prompt gamma imaging of proton pencil beams at clinical dose rate. *Physics in Medicine and Biology*, 59(19):5849–5871, oct 2014. ISSN 0031-9155. doi: 10.1088/0031-9155/59/19/5849. URL http://stacks.iop.org/0031-9155/59/i=19/a=5849?key=crossref.d1c598721e0b970b36f1c8a6ad1dd1a1.

M Pinto, D Dauvergne, N Freud, J Krimmer, J M Letang, C Ray, F Roellinghoff, and E Testa. Design optimisation of a TOF-based collimated camera prototype for online hadrontherapy monitoring. *Physics in medicine and biology*, 59(24):7653–7674, 2014. ISSN 1361-6560. doi: 10.1088/0031-9155/59/24/7653. URL http://www.ncbi.nlm.nih.gov/pubmed/25415207.

M Pinto, M Bajard, S Brons, M Chevallier, D Dauvergne, G Dedes, M De Rydt, N Freud, J Krimmer, C La Tessa, J M Létang, K Parodi, R Pleskač, D Prieels, C Ray, I Rinaldi, F Roellinghoff, D Schardt, E Testa, and M Testa. Absolute prompt-gamma yield measurements for ion beam therapy monitoring. *Physics in Medicine and Biology*, 60(2):565–594, 2015. ISSN 0031-9155. doi: 10.1088/0031-9155/60/2/565. URL http://stacks.iop.org/0031-9155/60/i=2/a=565?key=crossref.c4760640602341dfb0ce09b91884d4e4.

Marco Pinto, Denis Dauvergne, Nicolas Freud, Jochen Krimmer, Jean M. Létang, and Etienne Testa. Assessment of Geant4 Prompt-Gamma Emission Yields in the Context of Proton Therapy Monitoring. *Frontiers in Oncology*, 6(January):1–7, 2016. ISSN 2234-943X. doi: 10.3389/fonc.2016.00010. URL http://journal.frontiersin.org/Article/10.3389/fonc.2016.00010/abstract.

Christian Richter, Guntram Pausch, Steffen Barczyk, Marlen Priegnitz, Isabell Keitz, Julia Thiele, Julien Smeets, Francois Vander Stappen, Luca Bombelli, Carlo Fiorini, Lucian Hotoiu, Irene Perali, Damien Prieels, Wolfgang Enghardt, and Michael Baumann. First clinical application of a prompt gamma based in vivo proton range verification system. *Radiotherapy and Oncology*, 118(2):232–237, 2016. ISSN 18790887. doi: 10.1016/j.radonc.2016.01.004. URL http://linkinghub.elsevier.com/retrieve/pii/S0167814016000074.

C Robert, G Dedes, G Battistoni, T T Böhlen, I Buvat, F Cerutti, M P W Chin, a Ferrari, P Gueth, C Kurz, L Lestand, a Mairani, G Montarou, R Nicolini, P G Ortega, K Parodi, Y Prezado, P R Sala, D Sarrut, and E Testa. Distributions of secondary particles in proton and carbon-ion therapy: a comparison between GATE/Geant4

and FLUKA Monte Carlo codes. *Physics in medicine and biology*, 58(9):2879–99, may 2013. ISSN 1361-6560. doi: 10.1088/0031-9155/58/9/2879. URL http://www.ncbi.nlm.nih.gov/pubmed/23571094.

F. Roellinghoff, M. H. Richard, M. Chevallier, J. Constanzo, D. Dauvergne, N. Freud, P. Henriquet, F. Le Foulher, J. M. Létang, G. Montarou, C. Ray, E. Testa, M. Testa, and a. H. Walenta. Design of a Compton camera for 3D prompt-γ imaging during ion beam therapy. *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 648(SUPPL. 1):S20–S23, aug 2011. ISSN 01689002. doi: 10.1016/j.nima.2011.01.069. URL http://linkinghub.elsevier.com/retrieve/pii/S0168900211001471.

F Roellinghoff, a Benilov, D Dauvergne, G Dedes, N Freud, G Janssens, J Krimmer, J M Létang, M Pinto, D Prieels, C Ray, J Smeets, F Stichelbaut, and E Testa. Real-time proton beam range monitoring by means of prompt-gamma detection with a collimated camera. *Physics in medicine and biology*, 59(5):1327–38, 2014. ISSN 1361-6560. doi: 10.1088/0031-9155/59/5/1327. URL http://www.ncbi.nlm.nih.gov/pubmed/24556873.

David Sarrut, Manuel Bardiès, Nicolas Boussion, Nicolas Freud, Sébastien Jan, Jean-Michel Létang, George Loudos, Lydia Maigne, Sara Marcatili, Thibault Mauxion, Panagiotis Papadimitroulas, Yann Perrot, Uwe Pietrzyk, Charlotte Robert, Dennis R Schaart, Dimitris Visvikis, and Irène Buvat. A review of the use and potential of the GATE Monte Carlo simulation code for radiation therapy and dosimetry applications. *Medical Physics*, 41(6), 2014. doi: http://dx.doi.org/10.1118/1.4871617. URL http://scitation.aip.org/content/aapm/journal/medphys/41/6/10.1118/1.4871617.

Eric Sheldon and Douglas M Van Patter. Compound Inelastic Nucleon and Gamma-Ray Angular Distributions for Even- and Odd-Mass Nuclei. *Rev. Mod. Phys.*, 38(1):143–186, jan 1966. doi: 10.1103/RevModPhys.38.143. URL http://link.aps.org/doi/10.1103/RevModPhys.38.143.

J Smeets, F Roellinghoff, D Prieels, F Stichelbaut, A Benilov, P Busca, C Fiorini, R Peloso, M Basilavecchia, T Frizzi, J C Dehaes, and A Dubus. Prompt gamma imaging with a slit camera for real-time range control in proton therapy. *Physics in medicine and biology*, 57(11):3371–405, 2012. ISSN 1361-6560. doi: 10.1088/0031-9155/57/11/3371. URL http://www.ncbi.nlm.nih.gov/pubmed/22572603.

E Sterpin, G Janssens, J Smeets, François Vander Stappen, D Prieels, Marlen Priegnitz, Irene Perali, and S Vynckier. Analytical computation of prompt gamma ray emission and detection for proton range verification. *Physics in Medicine and Biology*, 60(12):4915–4946, 2015. ISSN 0031-9155. doi: 10.1088/0031-9155/60/12/4915. URL http://stacks.iop.org/0031-9155/60/i=12/a=4915?key=crossref.aabd8815e135401a22d165e343a7bac4.

E. Testa, M. Bajard, M. Chevallier, D. Dauvergne, F. Le Foulher, N. Freud, J. M. ??tang, J. C. Poizat, C. Ray, and M. Testa. Monitoring the Bragg peak location of 73 MeVu carbon ions by means of prompt ?? -ray measurements. *Applied Physics Letters*, 93(9):1–10, 2008. ISSN 00036951. doi: 10.1063/1.2975841. URL http://arxiv.org/abs/0809.0185http://dx.doi.org/10.1063/1.2975841.

Mauro Testa, Joost M Verburg, Mark Rose, Chul Hee Min, Shikui Tang, El Hassane Bentefour, Harald Paganetti, and Hsiao-Ming Lu. Proton radiography and proton computed tomography based on time-resolved dose measurements. *Physics in Medicine and Biology*, 58(22):8215–8233, nov 2013. ISSN 0031-9155. doi: 10.1088/0031-9155/58/22/8215. URL http://stacks.iop.org/0031-9155/58/i=22/a=8215?key=crossref.9766466f3cf307946adae72478a2726d.

Joost M Verburg and Joao Seco. Proton range verification through prompt gamma-ray spectroscopy. *Physics in Medicine and Biology*, 59(23):7089–7106, 2014. ISSN 0031-9155. doi: 10.1088/0031-9155/59/23/7089. URL http://stacks.iop.org/0031-9155/59/i=23/a=7089?key=crossref.2c380dcfa8360cb6e6defe62dd49f24a.

Joost M Verburg, Helen a Shih, and Joao Seco. Simulation of prompt gamma-ray emission during proton radiotherapy. *Physics in Medicine and Biology*, 57(17):5459–5472, sep 2012. ISSN 0031-9155. doi: 10.1088/0031-9155/57/17/5459. URL http://www.ncbi.nlm.nih.gov/pubmed/22864267.

Joost M Verburg, Kent Riley, Thomas Bortfeld, and Joao Seco. Energy- and time-resolved detection of prompt gamma-rays for proton range verification. *Physics in medicine and biology*, 58(20):L37–49, oct 2013. ISSN 1361-6560. doi: 10.1088/0031-9155/58/20/L37. URL http://www.ncbi.nlm.nih.gov/pubmed/24077338.

J F Williamson. Monte Carlo evaluation of kerma at a point for photon transport problems. *Medical physics*, 14(4):567–576, 1987. ISSN 00942405. doi: 10.1118/1.596069.

4.8 Notes on the article

A method that offers an accelerated estimation of PG yields in clinical proton therapy simulations has been presented. The method is useful for detector developers interested in system optimization: many permutations of different tunable parameters would need to be graded in terms of their performance in a MC simulation, as done by Pinto et al. (2014). As mentioned in the article, at this time vpgTLE does not estimate noise, which is absolutely a problem, but not a disadvantage to any other variance reduction methods, and likely not even to analog MC, as no validation of PG noise has been carried out so far for any simulation method. Which is not to say such investigations in the estimation of noise would not be interesting. Current studies tend to infer a simple analytical description of the noise based on measurement, and the question is to what extent this in insufficient.

In addition, vpgTLE might be used for clinical PG estimation for particular case studies. As described in the article, some alternative estimation methods exist. At PTCOG55 in Prague in May 2016, a poster presented work towards an analytic PG estimation integrated into treatment planning program. Analytical methods might need validation against MC, and vpgTLE may be used in that case. In addition, analytic estimation may be unable to reproduce correct PG estimates in complex CTs and plans, and here MC and vpgTLE could be used instead.

On reflection on this work, the analytic variance computation deserves further mention. Although implemented, it was found that the analytic variance gave a variance over a magnitude more optimistic than the batch method. The analytic variance was derived with the assumption that the scored tracklength per proton energy bin are independent between events. In El Kanawati et al. (2015) equation 4, this assumption is used to rewrite the explicit sum over n events as the average tracklength: the sum over proton energies and events have no crossterms. In appendix B, courtesy of Jean Michel Létang, a detailed note on the assumption of tracklength distribution is included, including a derivation of the variance with full correlation. These two assumptions represent two extreme cases (the correlation is likely neither zero nor one), and plots show that the statistical variance obtained with the batch method indeed lie between the two analytical extremes. The assumption of complete independence conveniently removes a nasty crossterm that would require us to score tracklengths per event or per 2D proton energy bin (see triangular sum in eq. 4 of appendix B). This was implementation is possible for the analytical phantoms that El Kanawati et al. (2015) considers, but are untenable in voxellized circumstances: memory usage would multiply with the number of proton bins. In our case, that would mean ≈ 2 GB $\times$ 250 = 500 GB of memory is required to run the simulation. Therefore, we can only compare the analytic statistical variance under the independence assumption and the batch method, which can be seen in fig. 4.1.

Another PG estimator was published very recently (Schumann et al., 2016). A 'filter kernel' was developed, that after convolution with the 1D dose profile, produced by either MC or an analytic dose engine part of a treatment planning package, produces a 1D PG profile. Moreover, an evolutionary method was developed to infer the dose based on a measured PG signal. In brief, a set of 1D dose profiles are generated, their associated PG estimates computed, and those profiles that match best the detected PG profile are allowed to survive. The next iteration is generated by applying a set of horizontal, vertical and local shifts to the 1D dose profiles and then again the new profiles are matched against the measurement. The paper states that about 1000 generations are required at the moment, which sounds possibly time consuming, but no performance numbers are presented. Only the 1D range seems to be considered, so no lateral prediction may be derived. In conclusion, while the methodology for inferring dose from the PG profile is intriguing, it seems that the applicability is much more limited than vpgTLE. Also, the apparent lack of sensitivity to lateral effects may limit the clinical use.

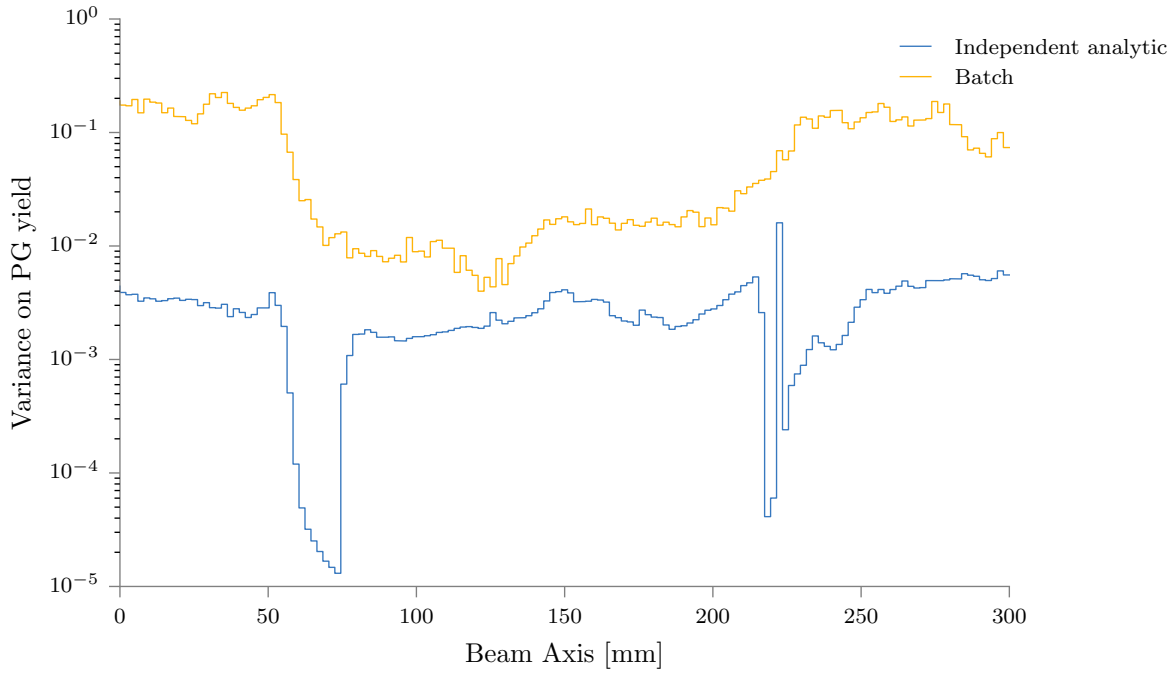


Figure 4.1: The statistical variance obtained in two ways is compared on a clinical CT with a treatment plan composed of opposing beams. Comparing the independent analytical statistical variance with the batch variance shows that the analytical gives an underestimate. It is expected, and shown for analytical phantoms, that the fully correlated analytical statistical variance would overestimate compared to the batch variance.

Chapter 5

Performance of Prompt Gamma fall-off detection in clinical simulations

In this chapter the first start-to-end clinical simulation of two optimized collimated PG cameras will be presented. Also, spot grouping methods based on the notion of iso-depth and iso-energy are compared. A small study of clinical spot weights was performed. Finally, a figure of merit is presented that provides an estimate of the probability of a measured FOP falling within or outside of the expected mean $\pm 2\times$ standard deviation. The chapter is intended to be submitted to Physics in Medicine and Biology. Before the submission, the spot grouping results will be expanded to include results for the iso-depth and iso-energy grouping based on all three spots studied.

Performance of Prompt Gamma fall-off detection in clinical simulations

Brent F. B. Huisman[1,2], É. Testa[2], and D. Sarrut[1]

[1] CREATIS, Université de Lyon; CNRS UMR5220; INSERM U1206; INSA-Lyon; Université Lyon 1; Centre Léon Bérard, Lyon, France
[2] IPNL, Université de Lyon; CNRS/IN2P3 UMR5822; Université Lyon 1 Lyon, France
E-mail: mail@brenthuisman.net

July 16, 2017

Abstract

Purpose: There is interest in the particle therapy community to use Prompt Gammas (PG) for treatment verification and eventually dose control. As the first detectors are deployed in clinical setting, we present a feasibility study of PG fall-off position (FOP) estimation using PG cameras with active beam delivery; the first where CT, treatment plan (TP) and beam description are all clinical data. Firstly, we investigate TP spot weight compositions, to establish what typical clinical spot weights are. Then, the statistics required for a consistent FOP estimate is investigated for two PG detectors. Finally, a spot-grouping method is proposed that combines better measurement statistics with fall-off preservation.

Materials and Methods: Our starting point is a study of spot weight distributions, based on TPs provided by various proton clinics in Europe. A head and neck case with a CT and replanning (RP)CT was selected for the PG detection study. A proton TP was constructed on the CT according to clinical protocols and irradiated in silico on both CT and RPCT. During the irradiation, two PG cameras implemented according to their published specifications, a multi-parallel-slit (MPS) and knife-edge (KES) camera, recorded the PG profile. A simple FOP estimation algorithm is presented, and executed on 20 CT and 20 RPCT realizations to obtain a shift distribution. A selected spot has its weight modulated, in order to study shift distributions as function of the spot weight, for each camera. Then, an iso-depth ion-range based spot-grouping method is presented and compared to iso-energy spot-grouping.

Results: Three spots were selected for an in depth study, and at the prescribed spot weights were found to produce results of insufficient precision, rendering usable clinical output on the spot level unlikely. When the spot weight is artificially increased to 10^9 primaries, the precision on the FOP reaches millimetric precision. On the FOP shift the MPS camera provides between 0.77 - 1.15 mm (1σ) precision for the three spots at 10^9 protons; the KES between 1.93 - 2.25 mm.

Conclusion:

While with spot measurements the measured shift was in agreement with the PG emissions shift, for both grouping methods this was not the case, unless the emissions was convolved with a point spread function to estimate the detected signal. The precision on the shift obtained with iso-depth grouping improved significantly with respect to iso-energy grouping for the KES camera (1σ = 3.82 and 2.87 mm, iso-energy and iso-depth resp.), but not significantly for the MPS (1σ = 1.17 and 1.05 mm, iso-energy and iso-depth resp.). It is shown that grouping spots does not necessarily negatively affect the precision compared to the artificially increased spot, which means some form of spot grouping can enable clinical use of these PG cameras, if the sum of the spot weights is at least 10^9 proton primaries. With all spots or spot groups the MPS has a better signal compared to the KES, thanks to a larger detection efficiency and a lower background level due to time of flight selection.

1 Introduction

The well-defined range of particles in matter is the main reason they are used in cancer treatment today. Unfortunately we are not able to take full advantage of this property, because of uncertainties in patient positioning, uncertainties on the proton range due to unknown displacements or deformations of organs, ill-defined lung,

bowel or bladder filling, weight loss, and the inherent uncertainty in the Hounsfield unit to particle stopping power conversion (Paganetti, 2013). Often, medical practice is to plan conservatively, namely adding margins around the tumor, greatly reducing the potential benefits of particle treatment (Knopf and Lomax, 2013). For example, irradiation fields with organs at risk (AOR) located just behind the Bragg Peak (BP) position are usually avoided, although they would better spare the AOR. Some form of in-vivo, online monitoring is generally considered to be a way out of this predicament. Online monitoring would make measurements of uncertainties such as mentioned above possible, and thereby permit more precise planning which could take maximum advantage of the steep BP fall-off and reduce damage to tissues surrounding the tumor.

Currently treatments can be verified by way of (offline) PET (Kraan, 2015). Integrating PET imaging in the beam gantry to reduce the effect of washout is beinga investigated (Sportelli et al., 2013). Another way to perform monitoring is to use prompt gammas (PGs), a natural byproduct in particle treatments. PGs offer the potential of monitoring treatment at the spot level (Roellinghoff et al., 2014; Smeets et al., 2012) and are therefore of prime interest (Golnik et al., 2014; Gueth et al., 2013; Janssen et al., 2014; Moteabbed et al., 2011). These particles are produced in the inelastic nuclear collisions between the incident particle and the medium (tissue) it is traveling through, and they correlate strongly with dose deposition. This has been experimentally demonstrated for protons by Min et al. (2006) and for carbon by Testa et al. (2008). A knife-edge PG camera (Perali et al., 2014; Richter et al., 2016) was put into clinical operation in fall 2015, at OncoRay in Dresden, Germany. The aim of such collimated cameras is to obtain a 1D profile of the PG production along the beam direction and obtain the position of the fall-off of the signal, which is strongly correlated to the BP position. Other approaches include detection of the target materials using a spectral PG camera (Verburg and Seco, 2014), using the primary particle's time of flight (ToF) in a timed PG measurement (Golnik et al., 2014) and Compton camera designs (Kurosawa et al., 2012; Llosá et al., 2016; Polf et al., 2015; Roellinghoff et al., 2011; Solevi et al., 2016; Thirolf et al., 2016).

Here we present an in silico study of the clinical relevance of fall-off position (FOP) estimation, based on two collimated PG camera designs: a multi-parallel-slit (MPS) and knife-edge (KES) design. First, we start with an analysis of the composition of clinical treatment plans (TPs). The main determinant for the available PG signal is the number of primary protons per quantity of interest (whole treatment, iso-energy layer, spot or otherwise). Depending on target type, location, and size the spot distribution varies. That is to say: the number of spots, their average number of protons and their energy distribution have consequences for PG detection. Then, a brief analysis of the data and some treatment planning effects are presented in context of these spot distributions. In the second part of this paper, performances of the two PG cameras are studied on a clinical case, fully implemented in Gate. For a patient with a cranial tumor, significant weight loss was observed and a replanning CT (RPCT) obtained. The planning and replanning CTs provide an excellent 'easy' case where PG camera should reliably detect the difference. We begin with a study of the minimum number of protons required for a reliable FOP retrieval. To conclude, we propose two spot-grouping methods that improve the reliability of fall-off retrieval.

2 Materials and Methods

2.1 Preliminary observations / Clinical Analysis

PG literature often quotes 10^8 protons as the reference number of primaries per spot in proton therapy. Precision estimates are therefore usually obtained for statistics of this order, although higher spot intensities are not uncommon, even in phantom studies (Priegnitz et al., 2016). Spot-by-spot (range) verification with one or two millimeter fall-off retrieval precision would be the eventual result, which is very appealing. We conducted a small study into the spot weight prescribed by current treatment planning software (TPS) to help focus the targets that PG cameras must be able to reach. Apart from detector performance and PG discrimination, the main determinant of the PG signal is, after all, the spot weight.

2.1.1 Treatment plan analysis

In Smeets et al. (2012) the composition of a particular TP is displayed: a scatter plot of the spot intensity as function of the proton energy. In greyscale the 1σ precision on the fall-off retrieval with the IBA Knife-Edge PG camera prototype, at that time, is colored in. Only for the spots with the highest intensities is the precision lower than 2 mm. Our Elekta Xio TPS produces similar maximum spot intensities. Our TPs are created for an IBA proton accelerator with scanning nozzle (Grevillot et al., 2011), which have some spots with 10^8 primaries

at their top ends. More modern accelerator designs are able to produce smaller spots, more closely packed, in order to increase dose conformity. Alternatively, moving the patient closer to the beam nozzle, non-isocentric planning, is a recent development Grevillot et al. (2015) enabling smaller spots. TPS uses the higher number of spots to produce more conformal TPs, at the 'cost' of a lower average spot-intensities. Our TPS produces plans with about 500 to 1000 spots per field, while newer accelerators and associated plans may have thousands of spots per field (see figure 3).

We analyzed a small set of clinical TPs to establish what clinical spot counts and proton counts per spot (spot weights) are, of which we present three 'archetypes' obtained from Uppsala University Hospital: a low precision case, a typical, medium precision case and a high precision case.

1. Low precision: Arteriovenous Malformation (AVM) case, figure 1.

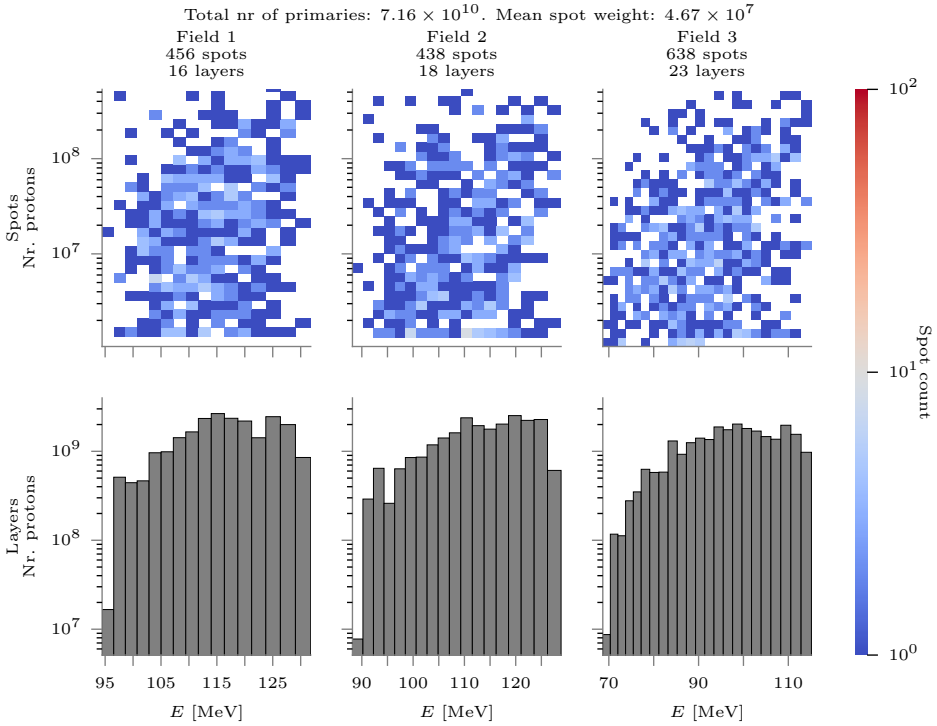


Figure 1: TP for a Arteriovenous Malformation (AVM) case, provided by Erik Almhagen and created in the Uppsala University Hospital, Sweden. On top: the spots are binned as function of energy and spot weight (number of protons), with high spot counts in red (hot spots) and low spot counts in blue (cool spots).

AVMs are not cancers, but tangles of blood vessels that divert arterial blood from the brain. The dose tolerance is high (up to 10%) in such cases. This high tolerance leads to fewer spots, with on average a higher proton count, which enables a quicker TP delivery and therefore higher patient throughput.

2. Typical precision: Meningioma case, figure 2.
   Cases such as these are common in Uppsala University Hospital. These medium tolerance plans involve a lower average proton count per spot with respect to the AVM case. Some spots reach the type of statistics that is commonly quoted in PG literature. This being a good representation for typical plans insofar we have been able to establish, it would be correct to uphold 10^8 protons as typical *for the maximum weight spots*. As presented in the figure, the mean spot-weight here is $1.08 \cdot 10^7$.
3. High precision: a large and complex pediatric brain and spine case, figure 3.
   Smaller spots, and therefore more spots, are used in plans where dose conformality is more pertinent. Much lower tolerances than usual are set because the patients are young and the side-effects must be minimized, which explains why the spot-count is over an order of magnitude more than the other TPs. This increase is accompanied with a decrease in per spot proton counts. The mean spot weight does not exceed 10^7, with the majority of spots below this number.

In conclusion, the number of spots can vary over more than an order of magnitude per plan, and therefore

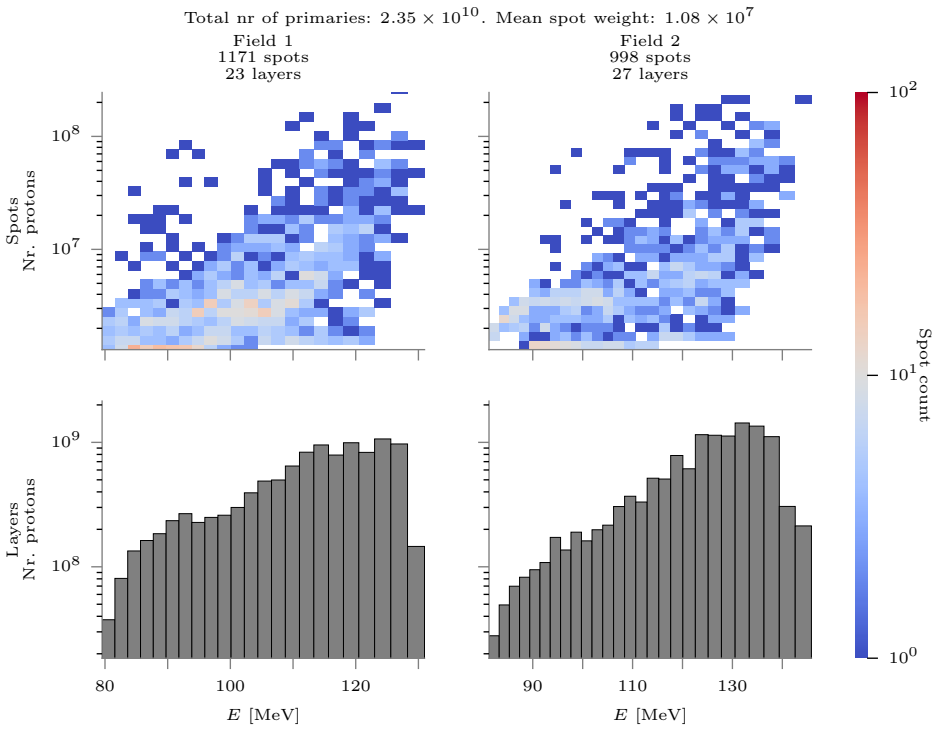


Figure 2: TP for a Meningioma case, provided by Erik Almhagen and created in the Uppsala University Hospital, Sweden.

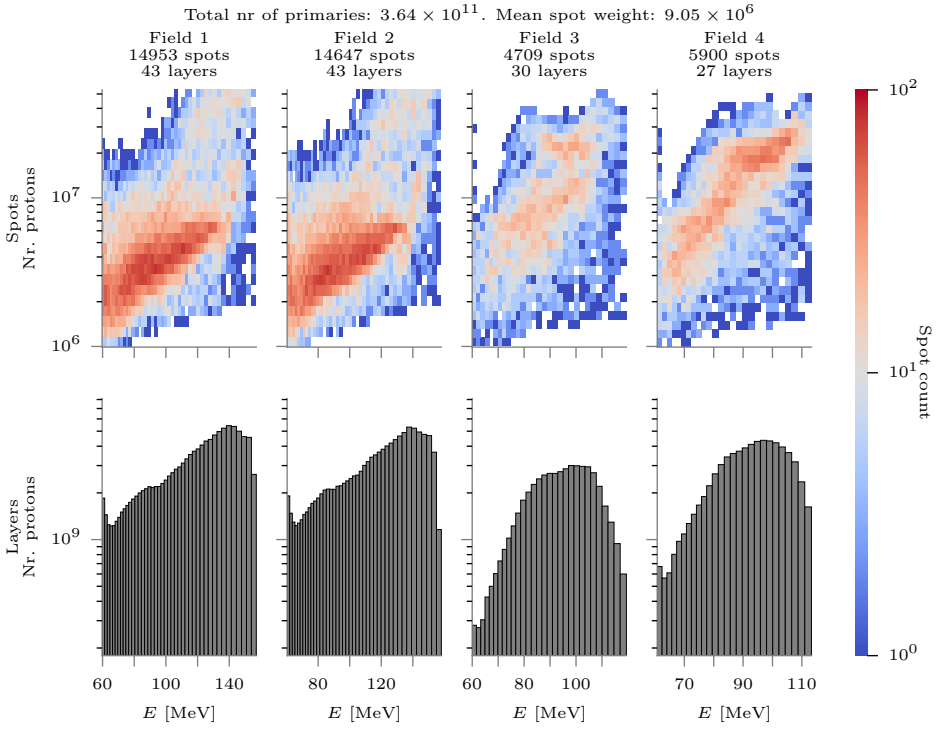


Figure 3: TP for a large and complex pediatric brain and spine case. The left two fields are for the brain, the right two for the spine.

spot intensities inversely vary over an order of magnitude as well. The trend most plans follow is roughly an oval spot distribution, tilted towards a positive correlation with energy. Translated to energy layers, the intensities roughly follow a forwardly skewed normal distribution. The negative correlation between the typical spot weight and plan robustness is an important observation, and presents a challenge: in robust plans where precision is required, and treatment verification seems most pertinent, PG cameras must be able to deal with lower spot weights than previously anticipated.

2.1.2 Iso-energy layers vs. iso-depth layers

An evolved version of the IBA Knife-Edge PG camera has been put into clinical testing at OncoRay, Dresden (Richter et al., 2016). Because at the moment of the test, this facility had passive beam shaping, PG imaging could only take place on the iso-energy layer level. Apart from the most distal and proximal layers, per-layer weights are close to or over over 10^9 primaries. This can be explained by the fact that tumor shapes are typically globulous, which means the most deeply penetrating energies only have to cover a small part of the transverse field (fig. 4). Depending on the entrance position, the tumor shape and any (large) density deviations in the beam paths, an iso-energy layer in general does not correspond to a iso-depth layer. However, a collimated PG camera measures the PG yield as function of position along the beam axis; it measures PG yields per iso-depth layer. Collected counts per iso-energy layer then implies range mixing.

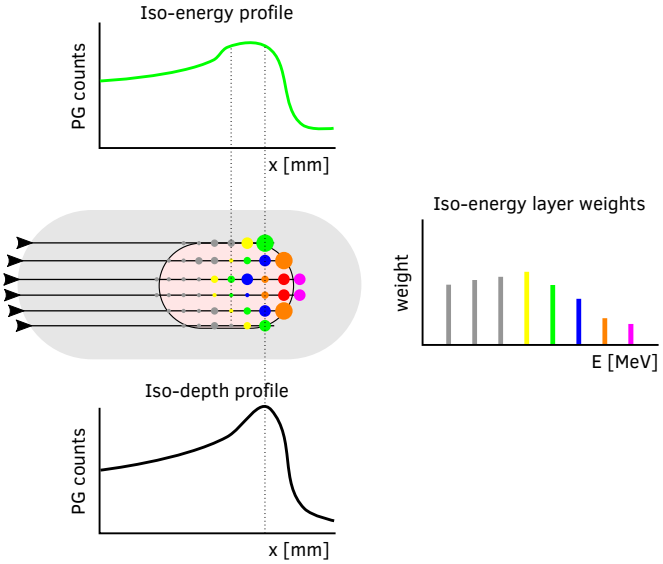


Figure 4: A schematic view of a patient (light gray), planned treatment volume (peach-pink), and a proton beam (black horizontal lines coming in from the left). Superimposed are various spots, where the size of the circles indicate spot weights. The spot colors indicate the iso-energy layer of which the spot is part. On the right, the weights (proton counts) per iso-energy layer are plotted, in similar fashion to fig. 2. On top, the PG profile is sketched for an iso-energy layer. On the bottom, the PG profile is sketched for iso-depth layer (spots of various color on the dotted line). In violet and red, the four dots on the distal end of the two most central beam lines, the effect of interpolation is illustrated: not always is every distal spot planned with a high weight.

In particle therapy and particle therapy imaging literature we often see it mentioned that a spread out Bragg Peak (SOBP) is achieved by giving the most distal iso-energy layer the highest weight and each successive iso-energy layer is of lower energy and of lower weight. This is correct on the spot-level, for homogeneous phantoms, and when interpolation is not required. Interpolation happens when the energy levels of the beam delivery system do not correspond with the distal tumor contour (Clasie et al., 2012), in which case the dose contour is approximated by spreading the distal dose of the nearest two possible energies. In general, for clinical cases, the correlation between energy and range is only true in a general sense, as supported by the vertical spread in fig. 2. The maximum weight spots are often seen in one of the most distal energy layers but not necessarily in the most distal one. Moreover, the weights of the distal layers are rarely the highest, so

the accepted wisdom of the distal iso-energy layer being the most intense and thus most pertinent to range verification, must be challenged.

2.2 PG camera modeling

As PG detectors we chose two prototypes that are tailored to FOP verification (illustrated in fig. 5):

- the CLaRyS multi-parallel-slit (MPS) camera, Case 1 (Pinto et al., 2014)
  The paper describes the optimization study of a multislit camera, where parameters such as collimator pitch, axis-to-collimator and axis-to-detector were varied, and their impacts evaluated on three criteria:
  - Case 1: best precision on the FOP
  - Case 2: best precision on the profile length
  - Case 3: best precision on the FOP with a FoV lower than 15 mm

  Each criterion resulted in a different optimal configuration. We are interested in the fall-off retrieval performance, which is why we select Case 1. As was done by Pinto et al. (2014), the lengths are chosen *up to* 300 mm, such that the length is an integer multiple of the pitch size, with for the collimator a collimator-leaf-width extra, to ensure each pixel has a leaf on both sides. With the 8 mm pitch and 2.6 mm collimator-leafs, this results in a scintillator volume of length 296 mm and collimator length 298.6 mm.
- the IBA knife-edge (KES) camera (Perali et al., 2014; Sterpin et al., 2015)
  The IBA knife-edge camera has seen the first clinical test of a PG camera. Richter et al. (2016) provides the first clinically obtained results. At this time, no other camera has been subjected to clinical tests, which is why we consider this prototype a benchmark.

An important detail is the fact that Pinto et al. (2014) carried out the optimization with a ToF selection window. Such a selection removes part of the background, based on the fact that the PG production is pulsed synchronous to the beam, while neutron-induced background is delayed. For the IBA C230 accelerator with a period of 10 ns, Pinto et al. (2014) chose a window of 4 ns around the PG maximum, based on experimental ToF spectra. This means that about 60% of the noise could be removed. For the KES prototype ToF is not used, leading to a higher background, as is evident when one compares the backgrounds as published in the two publications. A second difference is the energy selection window. The IBA group employ a 3-6 MeV window, whereas the CLaRyS collaboration produced their optimization with a 1-8 MeV window. We will compare each camera with their published properties, that is to say: a 1-8 MeV window and ToF for MPS and a 3-6 MeV window without ToF for KES.

2.3 Simulation

Imaging paradigms such as PG detection are validated against experiments, and often also with Monte Carlo (MC) simulations (Golnik et al., 2014; Gueth et al., 2013; Janssen et al., 2014; Moteabbed et al., 2011; Robert et al., 2013). For rarely occurring processes such as PG simulation, convergence to the model of the truth to within acceptable statistical error can be slow. This paper presents an in silico study of the feasibility of the clinical relevance of PG FOP estimation using collimated cameras (section 2.2), and uses the vpgTLE variance reduction method described in Huisman et al. (2016). In brief, a full *voxelized Prompt-Gamma Track Length Estimator* (vpgTLE) simulation is broken up into two stages. In stage one, a PG yield distribution image (PGyd) is created, specific to a particular phantom (or CT-image) and primary source (e.g. a treatment plan, a single spot). This PGyd image stores, per voxel per PG energy bin, the yield per primary. Stage two uses the PGyd and the assumption of isotropic PG emittance to generate and propagate the PGs throughout the rest of the geometry that the user has defined (e.g. the PG detector).

Gate 7.2 (Jan et al., 2004; Sarrut et al., 2014) with Geant 4.10.02 and the QGSP_BIC_HP_EMY physics list, commonly used for PG studies, are used in this analysis. Stage 1 is generated once (taking between 1 and 2 hours) and Stage 2 corresponds to a realization in the rest of this paper, and takes about 1-2 hours with 10^9 protons. (No protons are propagated in Stage 2, but the number of photons propagated is always proportional to the proton count, by way of the cumulative PG production factor of the PGyd.) These durations are obtained with a single core on an Intel(R) Core(TM) i7-3740QM, and could easily be parallelized.

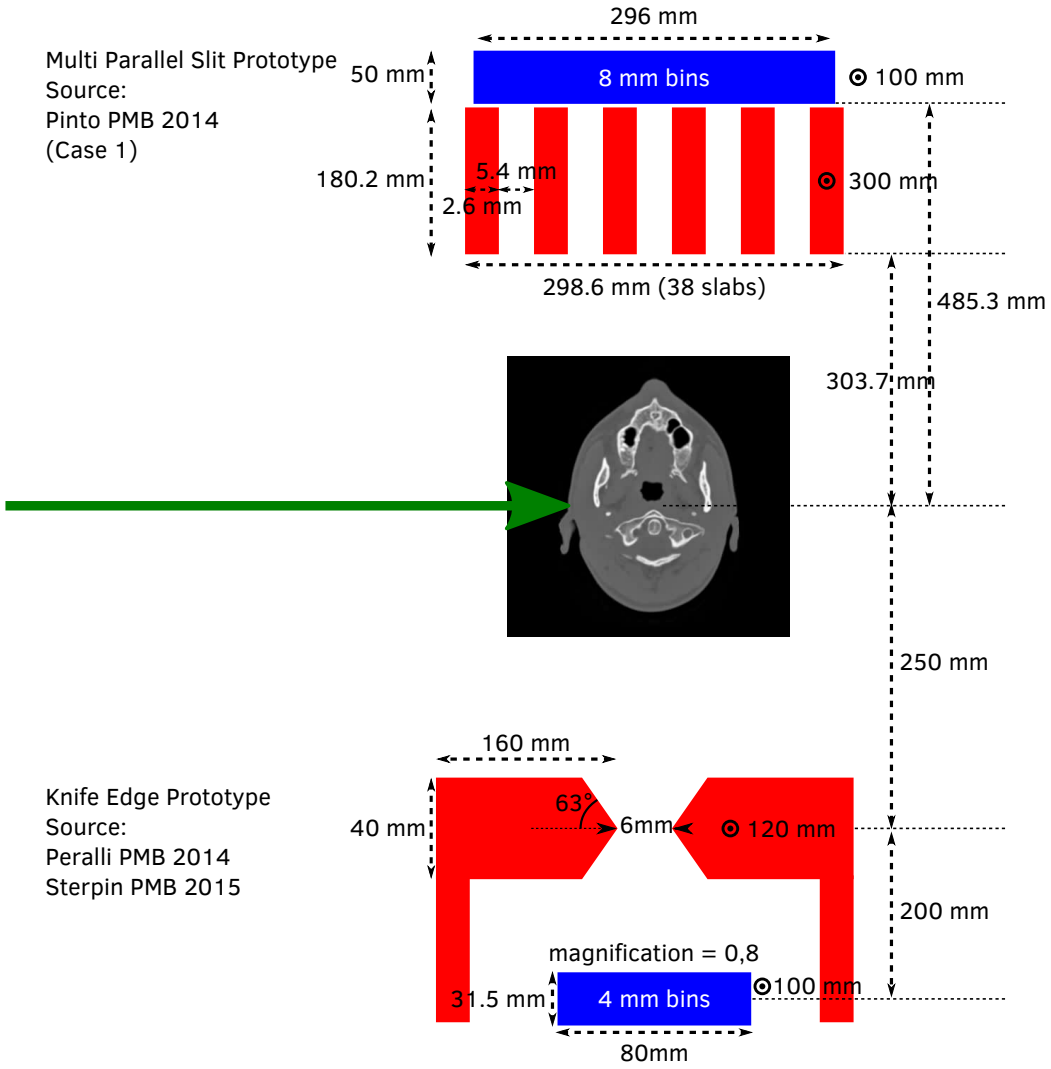


Figure 5: Schematic presentation of the two PG cameras considered in this study. The green arrow represents the proton beam. In red the collimation elements and in blue detection elements. The dimensions were taken from Pinto et al. (2014) and Perali et al. (2014); Sterpin et al. (2015). Note that the two cameras were positioned at an identical location above the head during all simulations.

2.3.1 Background estimation

Background estimation in PG simulation is a difficult and largely unsolved issue (Huisman et al., 2016; Perali et al., 2014; Pinto et al., 2014; Sterpin et al., 2015). Simulations would ideally include beam nozzle and whole room modeling, but these are habitually omitted because of the laborious work involved. ToF selection techniques can improve the signal-to-noise ratio (SNR) (Testa et al., 2008), but then depend on the proper simulation of the beam accelerator time structure, which the Gate Scanning Pencil Beam class currently does not provide. As noted in Huisman et al. (2016), no validation for background in PG simulations has been performed at this time. In this study, we will therefore assume the stable time structure of current generation cyclotrons, where the neutron background is largely constant. Estimates of background counts in the detector are taken from Perali et al. (2014); Pinto et al. (2014), which are both based on measured data:

- MPS: Pinto et al. (2014) fig. 9: $1 \cdot 10^3 \pm 1 \cdot 10^2$ per $4 \cdot 10^9$ primary protons per 8mm bin
  Converted to per primary proton: $2.5 \cdot 10^{-7} \pm 0.25 \cdot 10^{-7}$
- KES: Perali et al. (2014) fig. 11: $5 \cdot 10^{-7} \pm 0.5 \cdot 10^{-7}$ per primary proton per 4mm bin

Per unit of bin length, the background yield of the MPS with ToF is therefore 4 times as low as the background seen with the KES.

2.3.2 Auger Actor

Both PG camera prototypes have different photo-multiplication tubes and different detector electronics. In this study, these differences are not implemented. We instead use the same method as described in Gueth et al. (2013) to obtain a hit from an impinging photon. Using the lifetime of a Geant4 particle, one can set the ToF window around the time PG are expected to arrive. When the ToF window closes, if the integrated energy deposited in a crystal lies in the acceptable energy window, the event is recorded. The position of the event in the crystal is considered as the energy weighed barycenter of all interactions in the crystal, plus a random value taken from a 5mm FWHM Gaussian to simulate the electronics and the detector resolution.

2.3.3 CT data, treatment plans

The Centre Léon Bérard in Lyon, France, is a cancer-specialized hospital equipped with X-ray imaging and treatment facilities. We searched patient records for cases where a follow-up CT was made, in order to study the effect of morphological change over the time period of a (normally fractionated) treatment. A follow-up CT is made for cases where replanning is deemed necessary, based on externally observable morphological changes: i.e. for patients where a larger than expected or usual change is easily visible from the outside. It could be argued therefore that our selection of data leads us to study more severe cases of morphological change, not typical. While cone-beam images are more widely available because their lower imparted dose allows them to be recorded as part of regular clinical protocol, they are of sufficiently different quality from the planning CT that a comparison would be dominated by the quality difference, not the morphological change. Cone-beam assisted virtual CTs are a promising way of recreating CT quality images at later stages during treatment, but we did not have access to such software. The planning and replanning CT (PCT and RPCT) were co-registered on bony structures.

A CT and RPCT set was chosen for a patient with a head and neck tumor, wrapped around the trachea. Since the Centre Léon Bérard is an X-ray facility, we made our own proton treatment plans based on the contours provided by the clinicians. The hospital has a research machine with Elekta XiO proton treatment planning software, which can create treatment plans for a proton beam which is modeled in Gate (Grevillot et al., 2012). Under supervision of clinical physicists, realistic plans were produced on the PCT and exported to a format readable by Gate. A two field plan was created to prevent overdosing in AORs. In this paper, only the main field, field 2 of our plan (fig. 6), will be studied. The CT with the dose due to field 2 and the PTV structure are seen in fig. 7.

2.4 Data selection and analysis

2.4.1 Fall-off position estimation procedure

Publications sometimes mix up range estimation and FOP estimation. The first requires more than the estimate of where the primary particle stops: the primary point of entry into the patient (a "fall-in" position or FIP)

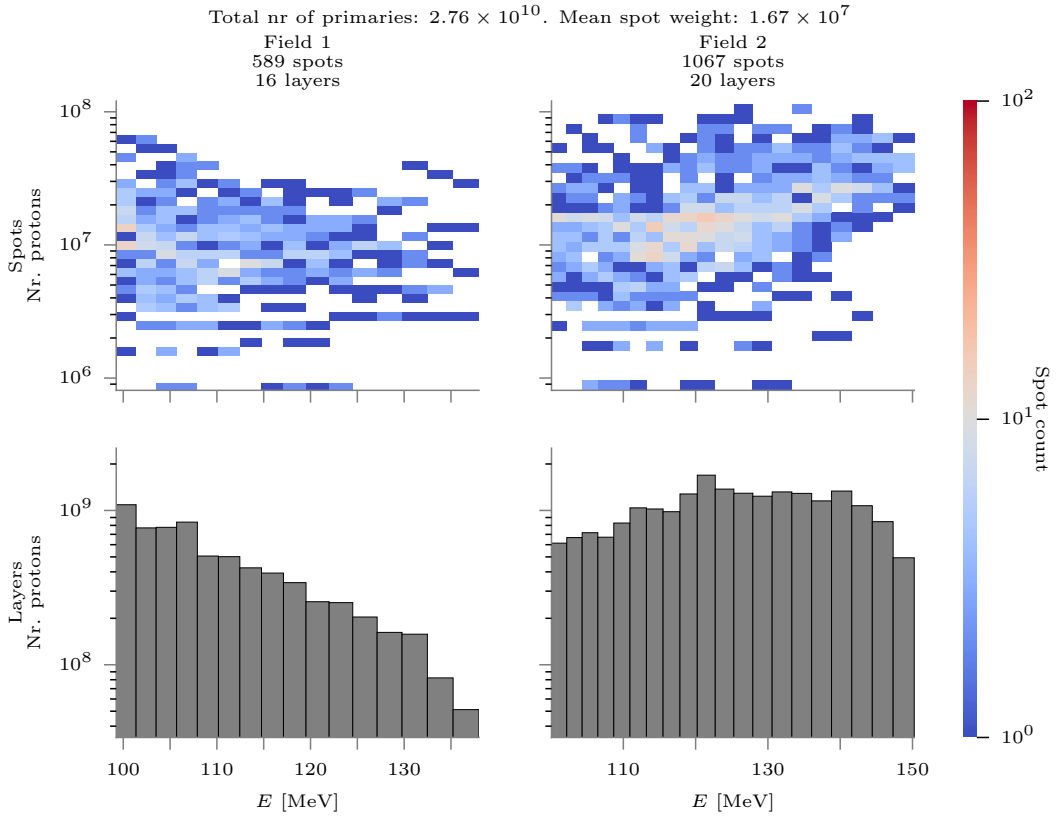


Figure 6: Structure of the 2-field treatment plan used in this study. The tumor was located in the head-and-neck region, wrapped about the trachea. In the plan shown here the first field is clearly shifted towards lower energies: this field was added to prevent field 2 from overdosing some AORs (most notably spine and optical nerves).

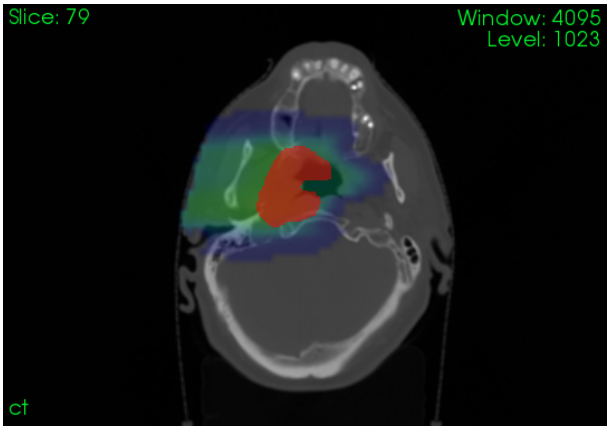


Figure 7: A slice of the CT used in this study is shown with the dose due to field 2 of our plan overlaid, with in red the PTV structure.

must also be measured, as was mentioned in Pinto et al. (2014) as the entrance position. Of course, range estimation requires that both positions are within the field of view (FoV) of the detector, which not every proposed PG camera is capable of. The KES camera has a more limited FoV, which is why we chose to only estimate the fall-off. Note that the FIP may be estimated by other means: the optical patient positioning systems present in most clinics could be used to get the FIP, although they do not provide a direct measurement. The range estimate could be more interesting than FOP from a clinical perspective. Consider the simple case where a patient was set up a centimeter further along the beam-line: the fall-off will have shifted, but the range remains unaltered. In the case studied in here, a large part of the patient's morphological change was weight loss in the subcutaneous layer. This implies a shift of the whole ion path, but not necessarily a change in range.

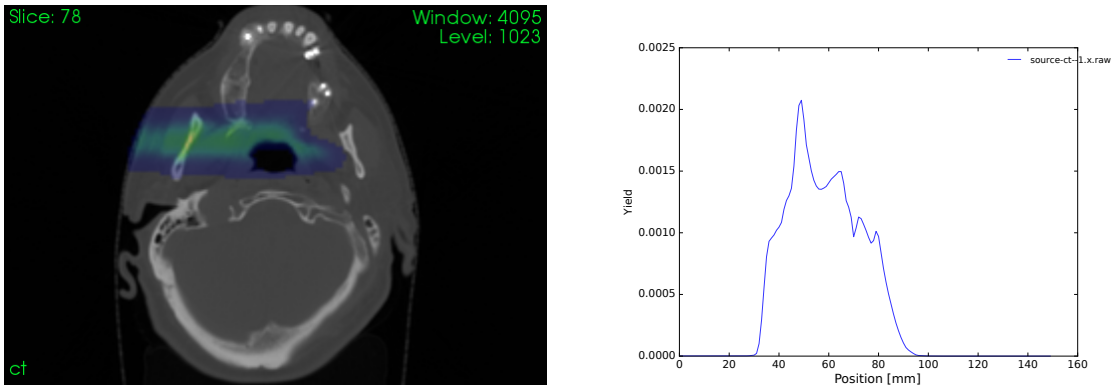


Figure 8: Left: CT with the PG production image of the most distal iso-depth layer of the second field of our TP in fig. 6. Right: the PG production profile (along the beam) of the same layer. A ragged plateau region due to inhomogeneities is sometimes called the camelback.

The guiding design principle behind collimated PG cameras is then range or FOP estimation: a 1D PG production profile is obtained, from which the FOP is extracted. Multiple approaches to extracting a FOP from the line profile have been proposed (Gueth et al., 2013; Janssen et al., 2014; Roellinghoff et al., 2014; Smeets et al., 2012; Sterpin et al., 2015). In figure 8 the peaks in the profile due to inhomogeneity are seen. Such a ragged plateau region is sometimes called camelback. No conventional fit procedure accounts for such situations. In preparatory work, we investigated a number of the proposed procedures, and all required significant tweaks to be able to deal with our data. In addition, insofar the methods had free parameters, we established significant sensitivity to such parameters on the final FOP estimates. In summary, the FOP estimate depends greatly on the procedure, the robustness relies on having yields uncommon on the spot-level in clinical TPs, and also on an absence of unavoidable inhomogeneities.

We will therefore not focus on the fitting procedure, but employ a simple method that visually works on most data we have. Our procedure is as follows:

1. The measured PG profile is smoothed and interpolated with a smoothing spline function:

$$\sum_{i=1}^{n} (Y_i - \hat{f}(x_i))^2 + \lambda \int_{x_1}^{x_n} \hat{f}''(x)^2 \, dx \tag{1}$$

   where Y_i is the measured PG profile and x_i the associated x-coordinates, $\hat{f}(x_i)$ the estimate smoothed spline function and λ a smoothing parameter that determines the penalty for deviating from measurement in exchange for smoothness (second order derivatives are close to zero on smooth functions). $\lambda = 0$ produces a perfect spline fit to the data, while $\lambda \gg 1$ produces a horizontal line. We found that $\lambda = 2$ struck the right balance between overfitting to noise and removing too many features, which tends to happen for low statistic measurements.
2. The obtained function is plotted for 1024 x_j. Any $f(x_j) < 0$ are set to 0.
3. The global maximum is found.
4. The baseline is set equal to the lowest 25% of bins.
5. From the distal end backwards, the first maximum is taken as the distal most peak position, if it is above the threshold of 30% of the difference between baseline and global maximum. If no such point is found,

the global maximum is taken as the distal most maximum.

6. The fall-off amplitude (FOA) is set to the difference between the distal maximum and baseline: $FOA = max - baseline$. The FOP is obtained by traversing the smoothed profile from the distal end towards the peak until $y_j > \frac{1}{2}FOA$.

The results of this procedure are illustrated in figure 9. Every PG profile was estimated 50 times, and so we obtained 50 estimates for the FOP. We assume that the FOPs follow a Gaussian distribution, so the mean of the 50 realizations gives us the best FOP estimate and the sigma gives the precision of the ability to estimate the best FOP. Comparing the 50 FOP estimates obtained from the CT with the 50 estimates obtained from the RPCT simulations, gives 2500 possible shift estimates. Again, the distribution of shifts should be centered at the true shift, while the sigma tells us how likely it is that this true shift is detected under the current conditions.

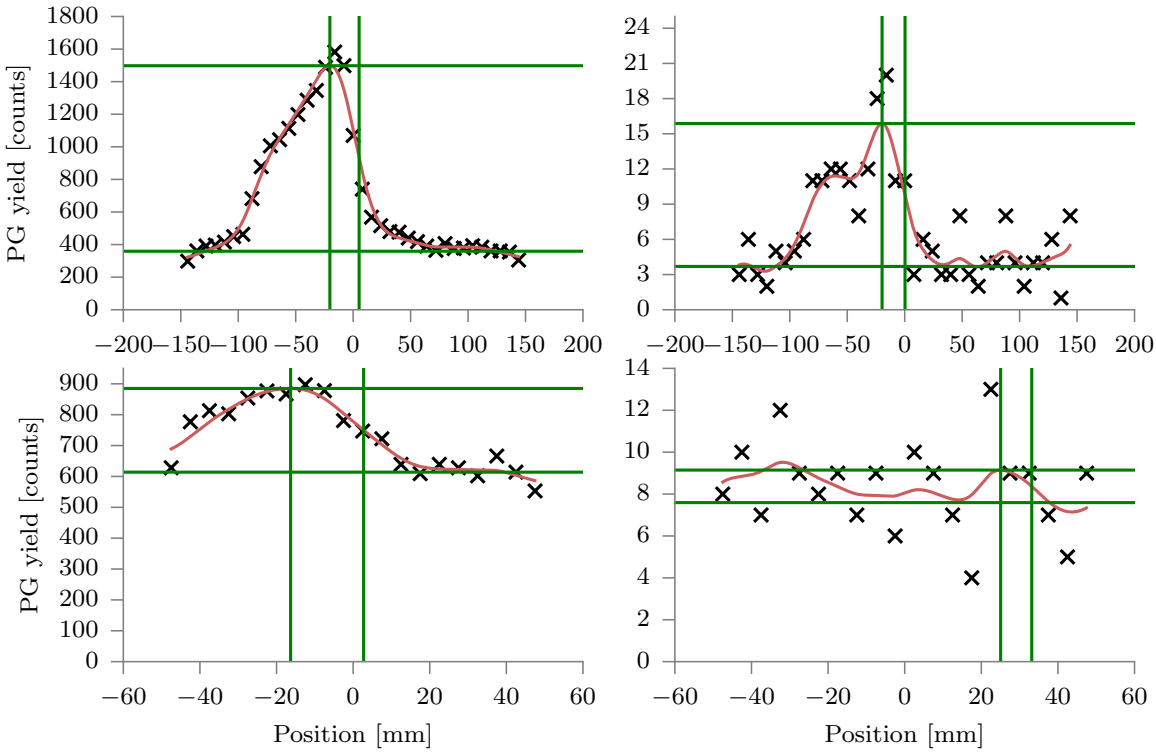


Figure 9: The top row demonstrates the fall-off determination procedure on the multi-parallel camera data; on the bottom row on knife-edge slit camera data. The left column is produced with a PG signal due to 10^9 primaries, while the right column was produced with 10^7 primary protons. In black crosses the measured PG counts are plotted. The smoothed data is shown in red. The green horizontal lines are drawn at the obtained distal maxima and baselines, while the vertical green lines shown the position of the distal maximum and the position of the fall-off. For the bottom-right plot, a history is visible where the procedure fails: the background induces an erroneous peak detection.

2.4.2 Spot selection

To understand the feasibility of spot FOP verification in clinical setting, it must be established how the quality of the shift estimate depends on the number of primaries per spot. We used two things in this search: the FOP estimation procedure detailed in the previous section, *executed on the dose image*, and a quality-of-fit estimation. The quality of fit estimation consists of testing if the FOPs at 50% is within a certain distance dx of the FOP evaluated at 20% of FOA and 80% of FOA. dx was set to 8 mm. An observation is that about 10% of spots do not pass the uniform FOP criterion: figure 10 shows two examples of the more typical case where inhomogeneities show their impact. For the spot-grouping procedure, 123 out of 1067 spots not meeting this criterion were discarded.

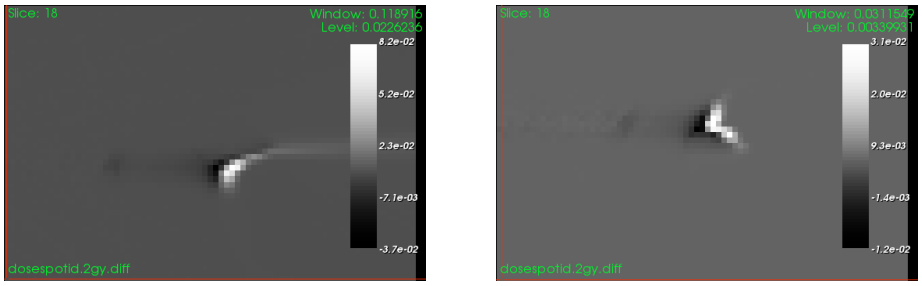
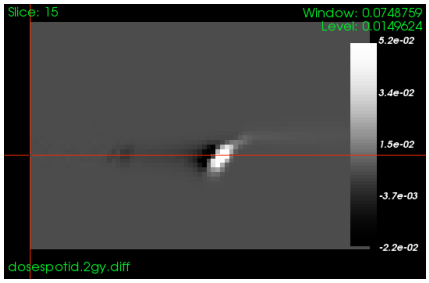
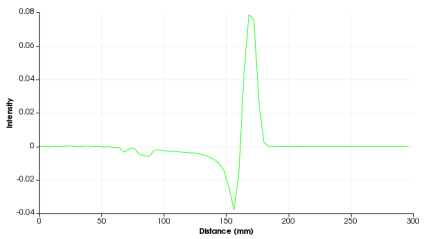


Figure 10: Axial slices of dose difference images for spots 60 and 361 demonstrate the effect of patient inhomogeneity on the range. The color scale was set to make the under- and overshoot regions clearly visible. We used a threshold to select "single range" FOPs by requiring that 20% of FOA and 80% of FOA are spaced out no more than 8 mm of 50% FOA. The proton beam is entering from the left.



Axial slice of the absolute dose difference for spot 61 of our TP, color scale chosen for best visibility. A relatively uniform forward shift may be observed.



Dose difference profile at the central line (shown in red on the left) of spot 61.

Figure 11: On the left the dose difference between CT and RPCT as seen from the top generated by irradiating spot 61. On the right the line profile corresponding to the red line (left).

We ended up choosing three spots (table 1). Spot 61, with a large shift of over a centimeter is shown in figure 11. Slightly visible on the right is a whiff of overshoot: the beam partially exits the patient causing the distal elevation seen in figure 12. An expected shift of over a centimeter should be reliably detectable for any PG camera. Spots 29 and 40 are more challenging; with expected shifts between 2 and 4 mm they represent a minimum shift that we hope PG cameras can detect.

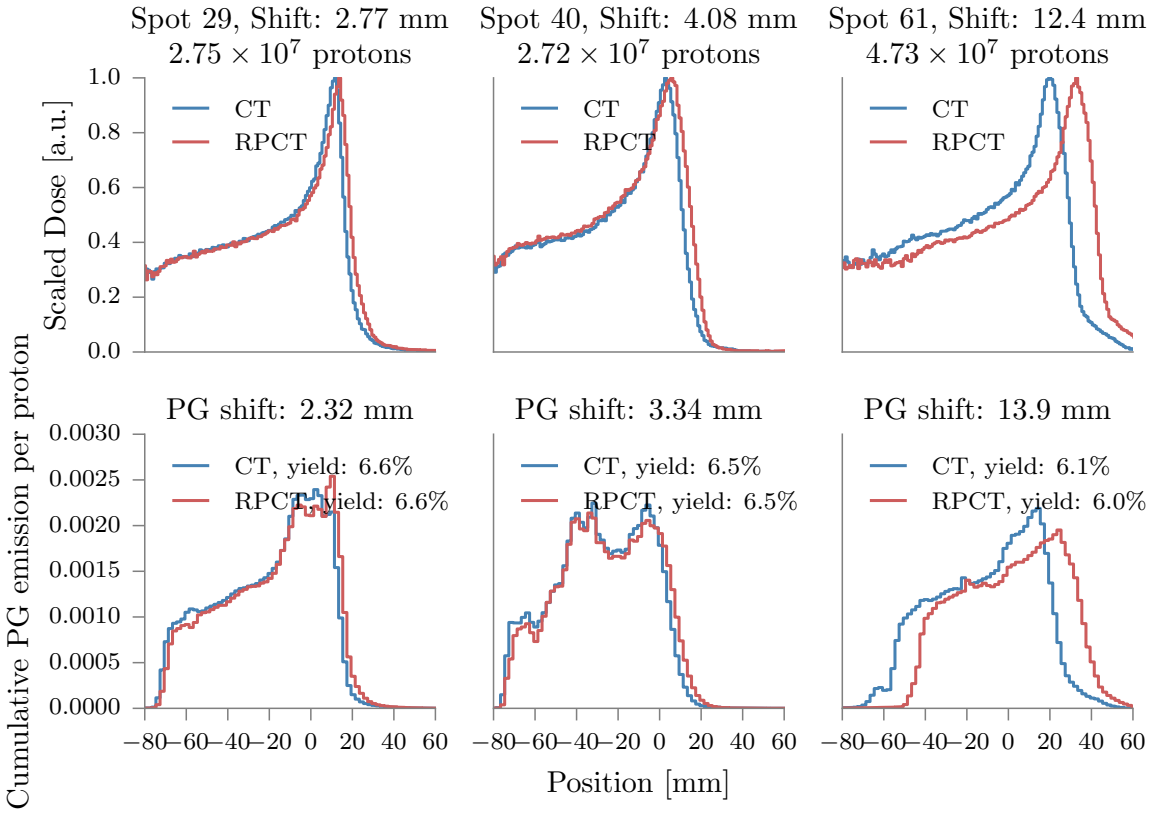


Figure 12: The three chosen spots, based on their shifts and quality of shift. Dose is normalized by mass, which explains the lack of structure. Top row: dose profiles, bottom row: PG profiles as used in the simulation. The yield is the integral yield of the whole image, so 6% means 6 out of 100 protons would generate a PG.

2.4.3 Spot grouping

As established in section 2.1.1, we expect a difference between the number of protons per spot as planned and sufficient for reconstruction. Depending on the beam line, the minimum unit of PG detection may be a spot (beam scanning) or an iso-energy layer (passive scattering). Since most new centers are of the beam scanning type, we propose an alternative spot grouping method. Our assumption is that due to inhomogeneities, protons within an iso-energy layer may have very different FOPs in the patient. We will compare grouping by iso-energy layer to grouping based on expected spot FOP. On the planning CT the 3D dose distribution is computed per spot, and then a dose FOP is determined with the same algorithm as for PG FOP, see section 2.4.1. We made the choice to use the dose FOP and not PG FOP because that is what is available as an output in TPS: dosemaps. In future TPSes, it may be imagined the (per spot) FOP could be output directly for applications such as PG monitoring.

The FOPs for each spot are binned along the beam line, and these bins we call iso-depth layers (red line in fig. 13). In terms of proton numbers (green line), this yields similar impressions as the layer views in figure 4: the distal layer(s) have higher spot statistics but lower cumulative (layer) statistics.

	Spot 29	Spot 40	Spot 61
Dose shift [mm]	2.77	4.08	12.4
PG emission shift [mm]	2.32	3.34	13.9
PG + PSF shift [mm]	2.61	2.91	11.9
Iso-energy layer [MeV]	145.86	143.02	143.02
x-pos [mm]	-24.0	-24.0	-24.0
y-pos [mm]	16.0	40.0	-40.0
Nr. protons	$2.75 \cdot 10^7$	$2.72 \cdot 10^7$	$4.73 \cdot 10^7$

Table 1: Summary of the properties of the selected spots.

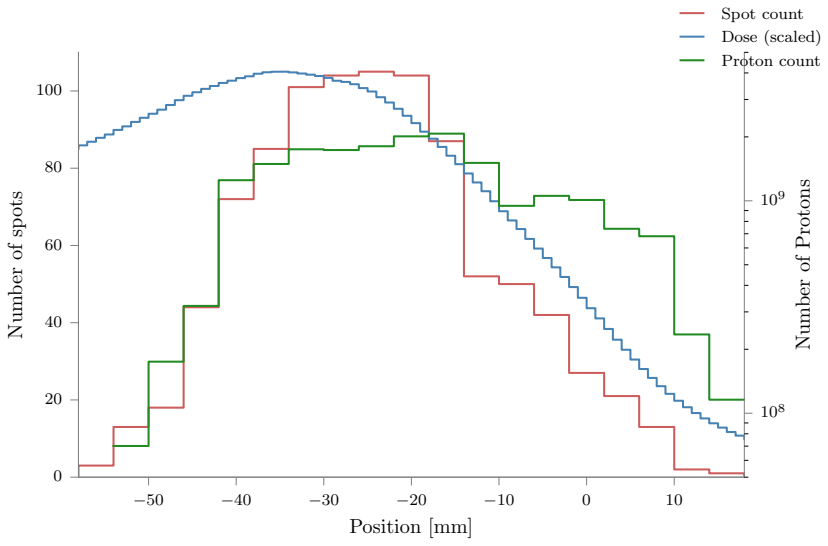


Figure 13: For the main field in the TP, per spot, a FOP is computed and binned (red, left y-axis). The depth dose distribution is overlaid in blue in linear arbitrary scale, as a reference for the shape. In green each spot is weighted with its number of protons (right y-axis) and then summed. With respect to the red line, the green line flattens out the top in the forward direction, supporting the thesis that in the geometric domain as well as the energy domain, the distal spots are individually weightier. However the cumulation of all spots leads to a larger dose in the medial layers of the target volume.

2.4.4 Figures of Merit

When evaluating the detector, principally we care about accuracy and precision: does the PG camera estimate the FOP consistently correct and does it so reliably over the course of a number of experiments? A more clinical point of view could be: does the detector correctly signal significant changes? A precise FOP estimate or FOP shift estimate is perhaps not required, but a correct indication of significant change *is*.

Employing the batch method we realize between 20 and 50 simulations for each experiment, each resulting in a FOP estimate, giving us a mean μ_{FOP} and a standard deviation σ_{FOP} for a certain experiment. Since we are studying the effect of replacing the CT with a RPCT to simulate patient change, we run each experiment with both CT and RPCT. Each FOP estimate of the N CT realizations is compared to each FOP estimate of the N RPCT realizations, resulting in N×N possible *FOP shift* measurements. These distributions have a μ_Δ and a σ_Δ. The shift initially obtained with the dose is denoted Δ_{dose}

Grading the performance of the detectors will be done according to a few figures of merit:

- Accuracy: $|\mu_\Delta - \Delta_{dose}|$.
- Precision: σ_Δ. For this estimate of the standard deviation of the Gaussian distribution a standard deviation can be computed once again based on the number of realization n used to obtain it: $\sigma(\sigma_\Delta) = \frac{\sigma_\Delta}{\sqrt{2\times(n-1)}}$, as per Leo (1994, formula 4.54).
- Confidence: the percentage of RPCT FOP realizations that fall outside $\mu_{FOP,CT} \pm 2\sigma_{FOP,CT}$ indicates the likelihood any difference from the expected FOP is measured. In other words, given that in this analysis we know that a shift should be detected, what is the probability we will? It will be denoted as P_Δ.

2.5 Verification of Knife-Edge results

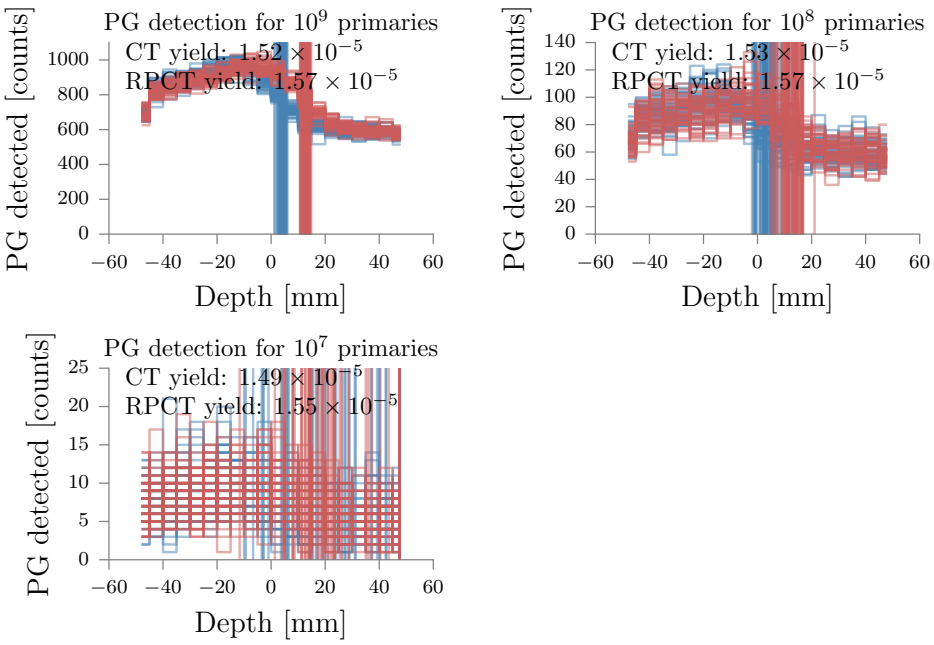


Figure 14: Knife edge FOP estimates for a simple simulation of a beam in a waterbox. In red, the beam was set at 139 MeV and 144 MeV is shown in blue. 50 realizations for each setting are overlaid, with their FOP plotted as a vertical line.

In Priegnitz et al. (2015) PG shifts due to beam energy shifts are studied. They include a study in the *detectability* of the fall-off as function of the number of primaries. To test our implementation of the knife-edge camera, we recreated the simulation: we shoot a mono-energetic beam into a waterbox at two energies with a 5 MeV shift. We created 50 realizations with a 139 MeV beam energy, and 50 realizations with 144 MeV. In figure 14 the profiles are plotted.

The results in figure 15 are very similar to those in Priegnitz et al. (2015, figure 9): at 10^9 primaries, the distributions are well separated with a shift of 8.3 mm (different from Priegnitz et al. (2015) because of the different material). In figure 13 in Perali et al. (2014) with 10^9 primaries a standard deviation of 1.5 mm is obtained, while we have 1.21 and 1.14 mm. We consider this sufficient agreement to be confident of our setup and further results. Roellinghoff et al. (2014) recalls that σ_{FOP} is proportional to the contrast-to-noise ratio (CNR) $\frac{\sqrt{\text{background}}}{\text{signal}}$, which is proportional to $\frac{1}{\sqrt{N_{\text{prim}}}}$. That means the increase in σ in figure 15 should be roughly a factor 3 as the order of magnitude in number of primaries decreases. Between 10^8 and 10^9 this holds, but between 10^7 and 10^8 it does not, likely because of the noisiness of the profiles generated with 10^7 primaries, on which the FOP estimation procedure is prone to misidentify the peak position.

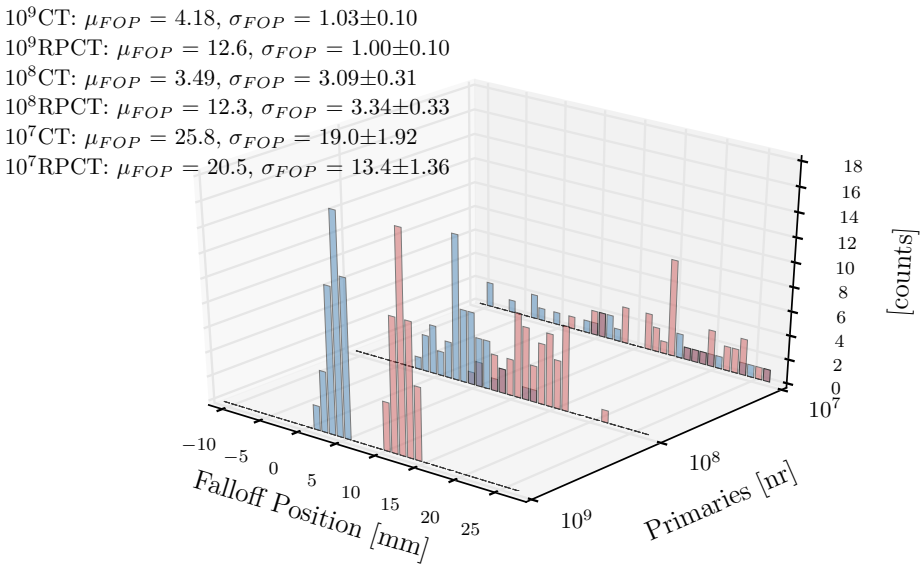


Figure 15: Knife edge FOP estimates for a simple simulation of a beam in a waterbox. In blue, the beam was set at 139 MeV and 144 MeV is shown in red. The CNR test holds between 10^8 and 10^9 holds, but between 10^7 and 10^8 it does not.

Note that the Knife-Edge prototype is very sensitive to accurate positioning with respect to the expected FOP. This was elaborated upon in Sterpin et al. (2015, Section IV.A.3): the detector response is, due to the KES collimator, not linear as with a parallel slit collimator. The response further away from the position directly in front of the KES slit is slightly lower due to the increased photon path length through the patient and the attenuation incurred by that longer path. Additionally, the solid angle per detector pixel depends on the position with respect to the KES slit. Initially we aligned the camera at the isocenter, which may be tens of centimeters away from the KES slit. We confirm the observation that this results in incorrect FOP estimates, for our spot 61 of up to half a centimeter. For clinical cases, alignment with the FOP from the TPS dose profile may be a workable alternative. In this study, to make the comparison as fair as possible and avoid any bias, alignment on the FOP specific for each spot was done as follows: the intermediate PG source image of vpgTLE (equivalent to the PG emission) was projected on the beam axis, and then convolved with a Gaussian of σ = 8.5 mm, which corresponds to the point spread function (PSF) with a FWHM of 20 mm used in Priegnitz et al. (2015) to approximate the detected profiles from the emitted profile.

2.6 Verification of Multi-Parallel Slit results

As a verification our implementation of the MPS camera, we checked that we obtain a similar precision on the FOP by repeating the procedure performed for the KES (fig. 16). In the caption of figure 9 in Pinto et al. (2014) we read that with 10^8 primaries a standard deviation of 1.3 mm for the detector design used here is obtained, which is about 20% different from our 1.63 and 1.54 mm in figure 17.

A last number perhaps of interest is the detection yield: we saw an average of $2.5 \cdot 10^{-5}$ detections per emitted PG for the MPS camera, and around $1.5 \cdot 10^{-5}$ for the KES. In combination with the number of detector

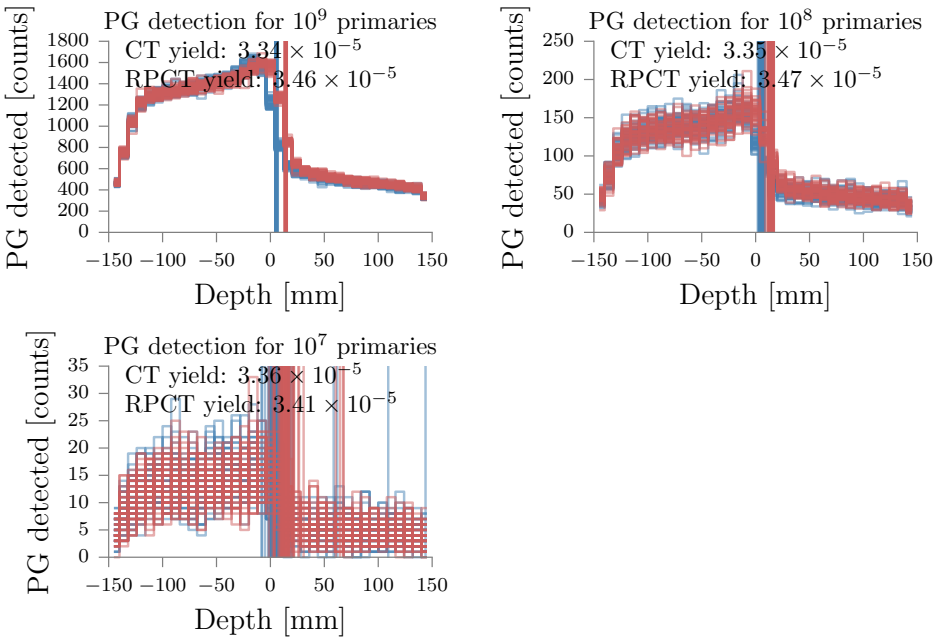


Figure 16: Multi parallel slit FOP estimates for a simple simulation of a beam in a waterbox. In red, the beam was set at 139 MeV and 144 MeV is shown in blue. 50 realizations for each setting are overlaid, with their FOP plotted as vertical lines.

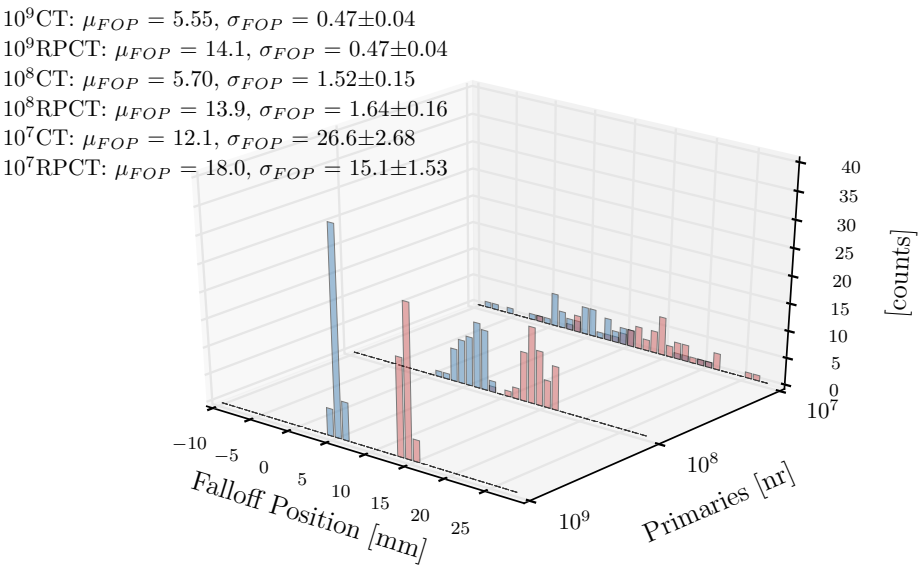


Figure 17: Multi parallel slit FOP estimates for the same setup used to verify the Knife Edge results: a beam in a waterbox. In blue, the beam was set at 139 MeV and 144 MeV is shown in red. The CNR test holds when comparing σs for 10^8 and 10^9 primaries, but again the profiles for 10^7 were too noisy to produce reliable FOP estimates.

bins (MPS: 37, KES: 20), these numbers help understand why 10^6 primaries can be dismissed right off the bat: a reasonable profile can never be estimated with a handful of counts. Therefore we omitted the 10^6 data from the analysis. The CNR test holds when comparing σs for 10^8 and 10^9 primaries, but again the profiles for 10^7 were too noisy to produce reliable FOP estimates.

3 Results

The first study is on the performance of the two collimated PG camera designs on a patient case with a CT and follow-up CT. Secondly, we investigate how two PG cameras perform as function of the number of primaries for a given spot. To conclude, we propose a new spot-summing method, we discuss why it's an improvement over iso-energy layer summing, and demonstrate its effect.

3.1 Weight-modulated single spot

Figure 18 shows all profiles, for all number of primaries, for both cameras, for spots 29, 40 and 61. The yields in the graphs correspond to the number of recorded hits anywhere in the spectrum, per launched primary proton, per realization (averaged). Based on the yields in these three spots, the first two (29 and 40) are very similar while spot 61 is a bit different. For both cameras the PG signal is lower (in particular the FO amplitude) in spot 61 than for the other two spots. Looking at the profiles, for best visibility the 10^9 realizations, we see the background-level is slightly lower and the signal to background ratio is worse for spot 61 with respect to the other spots. Note that the background is not the same as the background-level: with the MPS about 30 to 40% of the background-level is contributed by scattered PGs, while with the KES this is about 10%. With the MPS the SNR is about 13% lower when comparing spot 61 to 40, while with the KES the SNR is reduced by about 7%.

Figure 19 shows the FOPs obtained from the profiles in figure 18, binned per CT (blue) and RPCT (red). We see the distributions get narrower and exhibit less overlap as statistics improve. We see that taking the PG emission and convolute with a Gaussian is a good assumption: the μ_{FOP} is always within a millimeter of the FOP on the "PG emission + PSF". If we stick to the results obtained with 10^9 primaries, we see with spots 29 and 40 a stable $\sigma_{FOP} \approx 0.5$ - 0.6 mm is obtained for the MPS camera and 1.3 - 1.6 mm for the KES. Switching to spot 61, the lower yields obtained for spot 61 translate to a lower precision. Using that the contrast to noise ratio is proportional to σ_{FOP}, the expected difference in the loss of precision is consistent with about 30% and 17% for the MPS and KES respectively, between spots 40 and 61. Taking the mean of the $\sigma_{FOP,CT}$ and $\sigma_{FOP,RPCT}$ of the 10^9 results in, the MPS σ_{FOP} indeed show a round 30% increase and for the KES 23%. We can conclude that not all spots are created equal, and that the σ_{FOP} seems to depend on location in the target (attenuation) and heterogeneity of the tissue in which it is delivered.

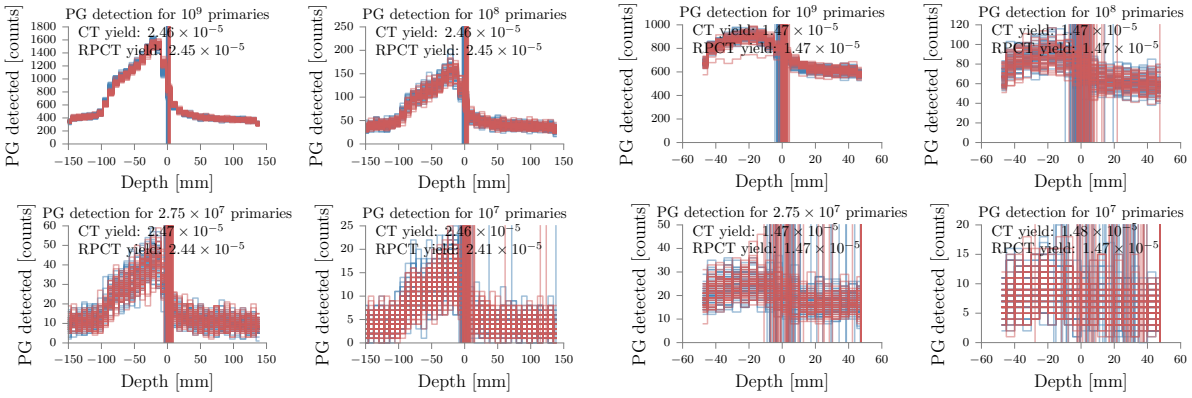
The FOPs obtained using the 50 realizations with the CT and the 50 with the RPCT can be compared with each other, thereby obtaining a distribution of 2500 possible shift estimates. The mean of the distribution of FOP differences should be centered at the true mean proton FOP difference.Figure **??** shows these FOP shift distributions per spot, per camera, per number of primaries.

Table 2 summarizes the results relevant to the ultimate estimated shifts. With 10^9 primaries, both cameras are within 2 mm of the "PG + PSF" emission shift, except the KES camera in spot 61. Here, when its σ_Δ is considered, it is within one σ of the expected value. At the prescribed spot weights, the MPS still converges on the correct μ_Δ, but the KES no longer. Considering that the σ_Δ for the KES is on the order of half the field of view of the camera, it seems fair to draw the conclusion no correct FOPs are estimated at these spot weights.

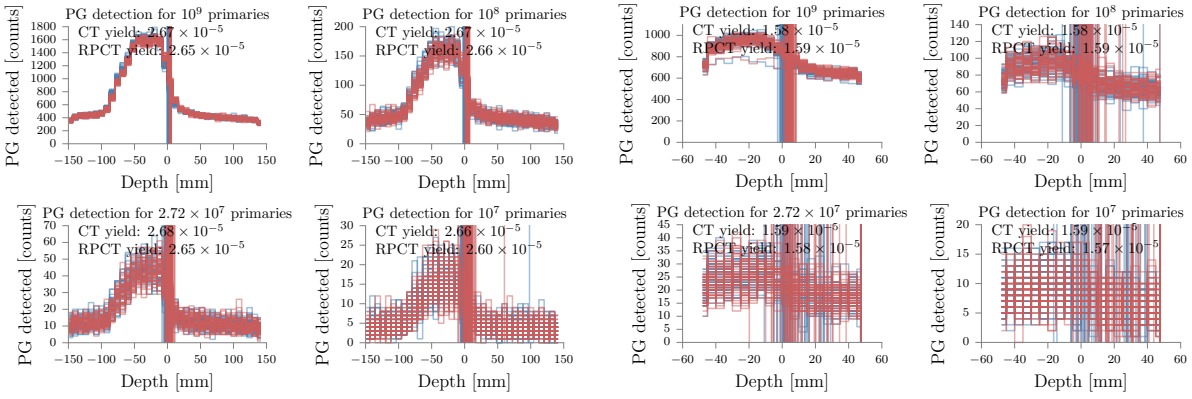
Turning to P_Δ at 10^9 primaries, we observe that the MPS performs very well: for any spot the probability that a single RPCT realization indicates a change from the CT is at least 99%. The KES reaches the same probability only for the large shift of spot 61. At their planned spot weights, neither camera provides great results except for the MPS for spot 61: even though the precision is 4.73 mm, the size of the true shift is enough to keep the CT and RPCT FOP distributions separated enough to be 96% certain a single measurement would have provided a clue that there was a significant shift.

3.2 Spot grouping

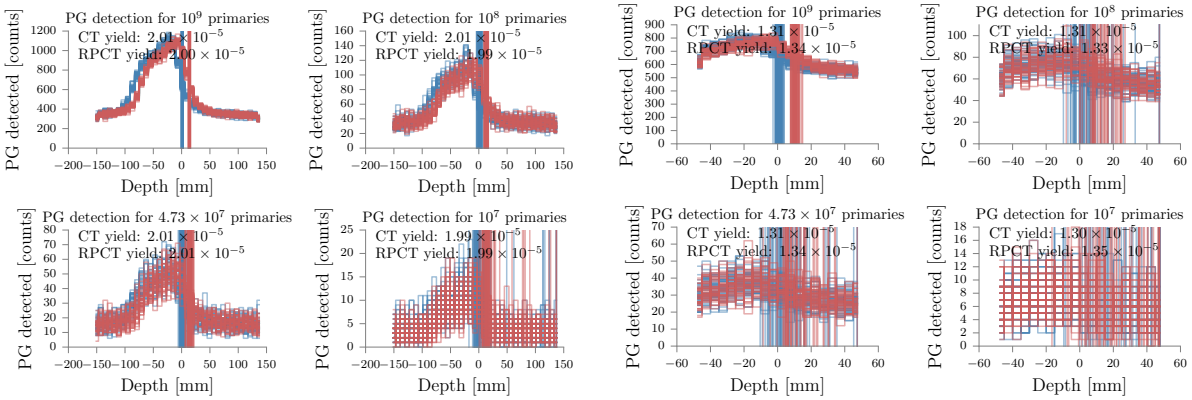
Since spot-by-spot FOP shift measurements appear to be off the table, we can consider collecting data during the delivery of multiple spots. Based on the spot modulation analysis, the minimum proton count required



MPS, spot 29 KES, spot 29

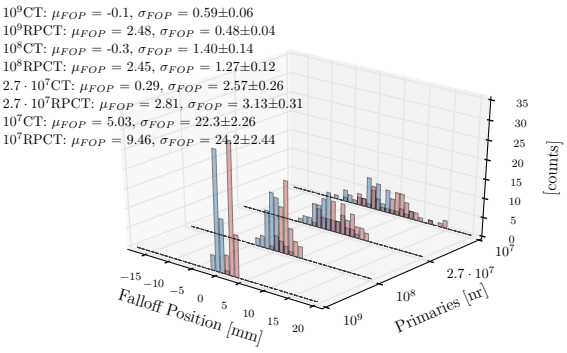


MPS, spot 40 KES, spot 40

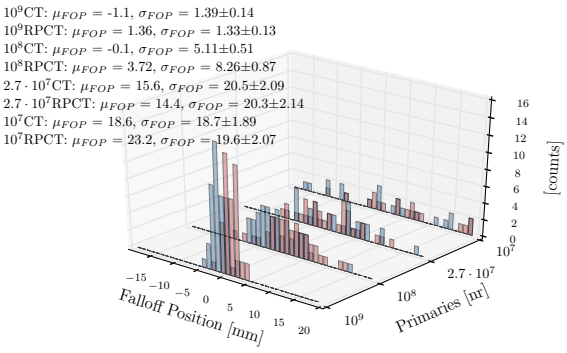


MPS, spot 61 KES, spot 61

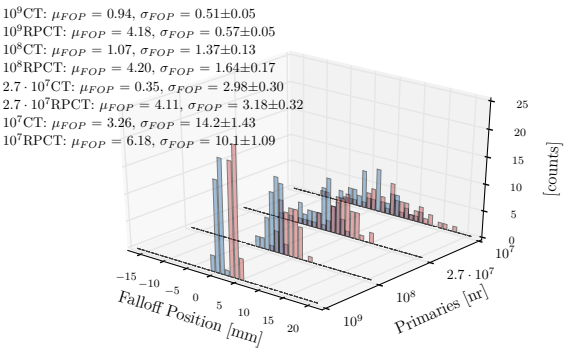
Figure 18: The profiles plotted for spot 29, 40, and 61 (row-wise) and for the MPS and KES cameras (column-wise). Per subplot, four graphs are shown, based on 10^9, 10^8, 10^7, and the original spot weight (clockwise, starting top-right). In blue the results with the CT image, red the RPCT image. The yields in the graphs correspond to the number of recorded hits anywhere in the spectrum, per launched primary proton, per realization (averaged).



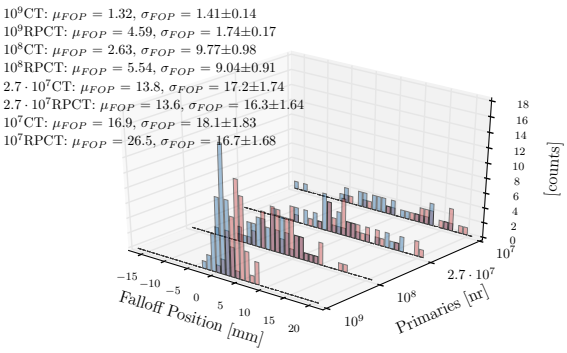
MPS, spot 29



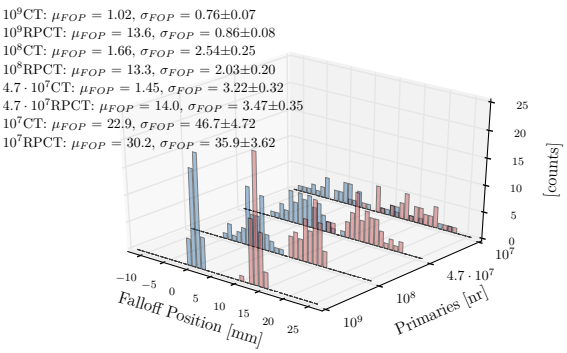
KES, spot 29



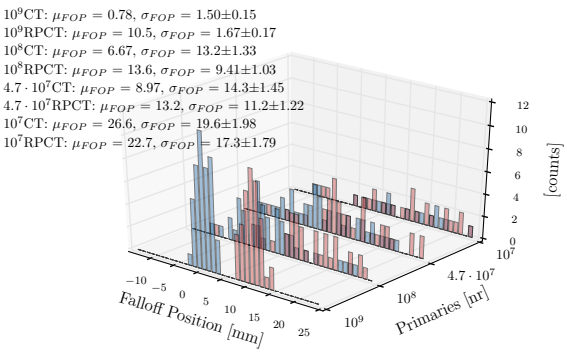
MPS, spot 40



KES, spot 40



MPS, spot 61



KES, spot 61

Figure 19: The FOP distributions are plotted for the spots (row-wise) and for the MPS and KES cameras (column-wise). Per subplot, four times two FOP distributions are shown. In blue the CT FOPs, in red the RPCT FOPs. In front the FOPs based on 10^9, 10^8, the original spot weight, and 10^7 as we go towards to back of the 3D plot.

	Spot 29 shift [mm]	Spot 40 shift [mm]	Spot 61 shift [mm]
Dose	2.77	4.08	12.4
PG emission	2.32	3.34	13.9
PG + PSF	2.61	2.91	11.9
μ_Δ MPS @ 10^9	2.68±0.77	3.23±0.77	12.6±1.15
μ_Δ KES @ 10^9	2.56±1.93	3.27±2.24	9.79±2.25
μ_Δ MPS @ spot weight	2.51±4.05	3.76±4.36	12.5±4.73
μ_Δ KES @ spot weight	-1.2±28.9	-0.2±28.9	4.30±26.2
	Spot 29 P_Δ	Spot 40 P_Δ	Spot 61 P_Δ
P_Δ MPS @ 10^9	99%	99%	100%
P_Δ KES @ 10^9	43%	59%	99%
P_Δ MPS @ spot weight	20%	24%	96%
P_Δ KES @ spot weight	1.9%	1.6%	1.4%
	Spot 29 spot weight	Spot 40 spot weight	Spot 61 spot weight
Nr. protons	$2.75 \cdot 10^7$	$2.72 \cdot 10^7$	$4.73 \cdot 10^7$

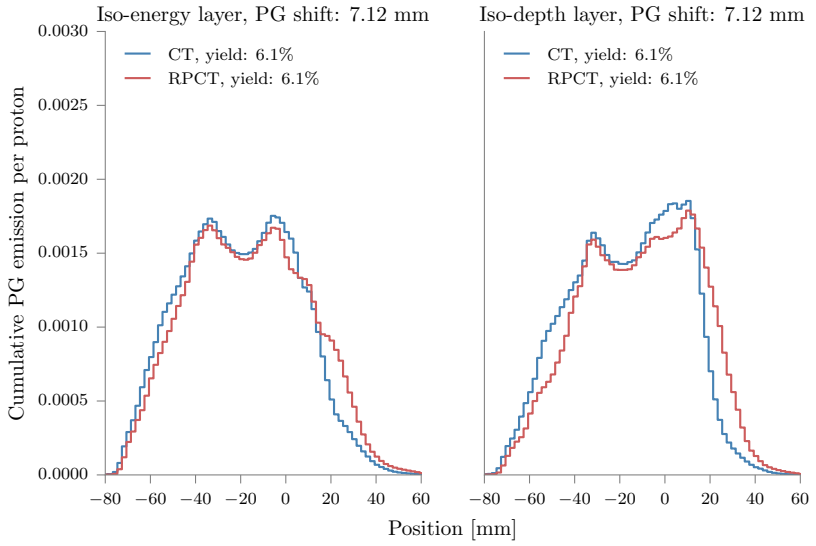
Table 2: Results from figure **??** summarized and compared to the ground truths from figure 12: the accuracy can be gleaned by comparing the shifts, the probability of measuring the change is estimated with P_Δ. With "PG + PSF" the FOP on the PG emission profile convolved with a Gaussian with a 20 mm FWHM is indicated. The geometric effect of the detector not measuring the PGs at emissions but at detection, approximated by the convolution, leads to up or downstream shifted FOPs.

should be no less than 10^9 primaries. We take spot 61 as a start, which is conveniently part of the first iso-energy layer (counted down from highest energy) having at least 10^9 protons. The geometric group is constructed by grouping spots with FOPs closest to spot 61's FOP until $10^9 \pm 5\%$ protons is reached. The final count is $1.07 \cdot 10^9$ and $0.98 \cdot 10^9$ protons in the iso-energy and iso-depth layers respectively.

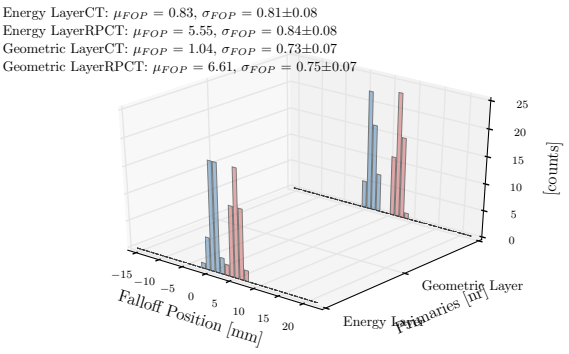
Observing figure 20, we see the fall off region is narrower, steeper and more regular with iso-depth grouping. Because of the sparse number of spots around spot 61, the grouped spots have dose FOPs of up to 9 mm different from spot 61. Altogether the iso-energy and iso-depth grouping share 30% the same spots. The camelback in the iso-depth profiles may be explained by this spread in FOPs. The iso-energy RPCT shows a clear partial forward shift with a shoulder right around the middle of the fall-off, which may translate to spread in the obtained FOPs. The fall-off slopes in the iso-depth profiles are different, which is reflected in second row of the figure: a larger σ_Δ for the RPCT is obtained. If we consider that spot 61 and spot 40 are both part of the energy layer, with very different FOPs, but the σ_{FOP} is preserved. Averaged over spots 40 and 61, CT and RPCT, σ_{FOP} is 0.56 and 1.58 mm for the MPS and KES respectively, while averaged over iso-depth and iso-energy grouping and CT and RPCT, σ_{FOP} is 0.78 and 2.4 mm resp. For the KES, σ_{FOP} is consistently better with iso-depth grouping however (averaged over CT and RPCT: 2.02 and 2.69 mm for iso-depth and iso-energy resp.).

If we make the iso-depth and iso-energy groups based on spots 29 and 40, the plots in figure 21 are obtained. Now, the number of primaries per grouping are nearly identical between the groupings ($0.84 \cdot 10^9$ for the spot 20 groupings and $1.07 \cdot 10^9$ for the spot 40 groupings). The lower number of primaries is reflected in the P_Δ for both cameras: both are no longer at a steady 99% as with the spot 61 groupings, but less. The small shift of spot 29 translates to small shifts in the groupings, especially the iso-depth grouping. Here, the FOP shift is well within the precision of the KES camera, and so it clearly loses its power to discriminate the CT signal for the RPCT (P_Δ = 14%). The MPS shows, apart from the iso-energy group about spot 29, a slight improved on the μ_Δ when compared to the spot set to 10^9 protons. For each camera and spot, the precision on the FOP difference (μ_Δ) improves with iso-depth over iso-energy grouping.

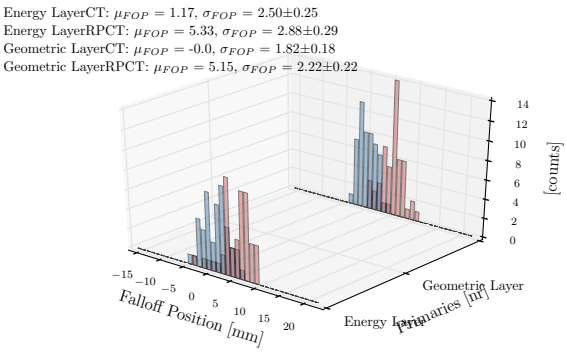
Observing table 3, the better FOP precision of the KES for iso-depth grouping is translated to a higher probability to measure a significant change: μ_Δ improved from 38% to 75%. Looking at the groupings for spot 40, we see that the PG + PSF convolution produces a very similar FOP difference, while the FOP at PG emission



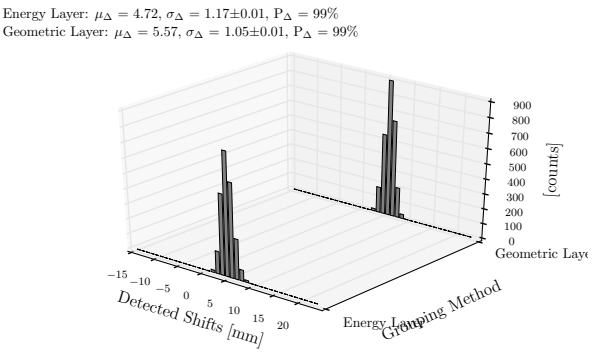
PG emission



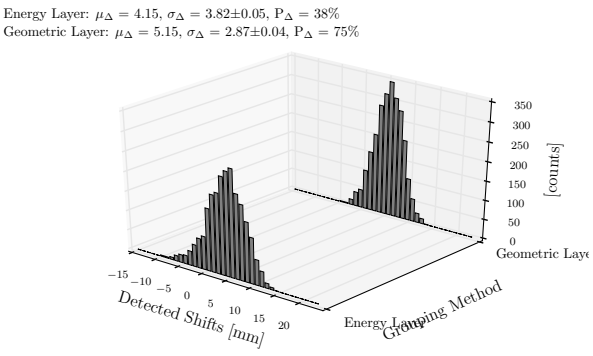
MPS, FOP distribution



KES, FOP distribution

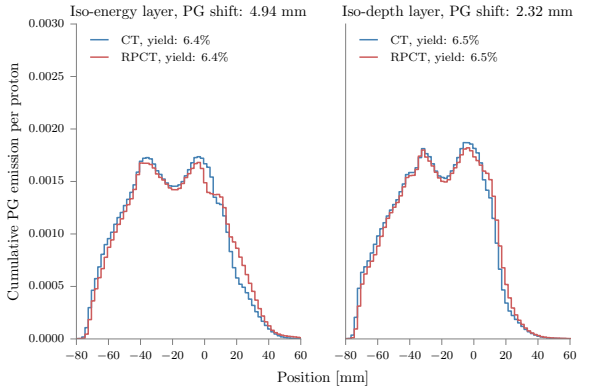


MPS, FOP shift distribution

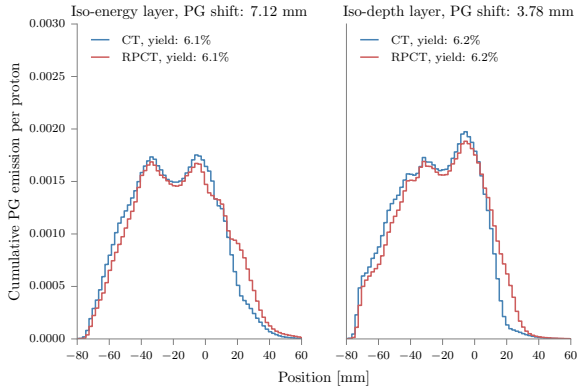


KES, FOP shift distribution

Figure 20: On top the PG emission profile (along the beam) is shown: left the iso-energy and right the iso-depth spot groupings, again in blue the CT and red the RPCT. On the second row, we see the FOP distributions with on the left the MPS and the right the KES results. The bottom row shows the FOP shift distributions.

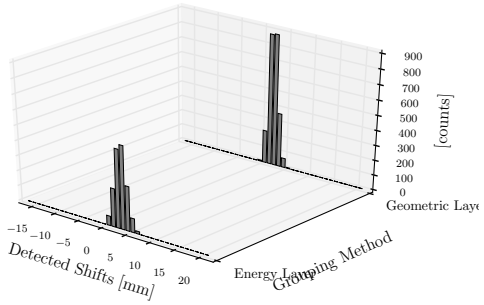


PG emission, spot 29



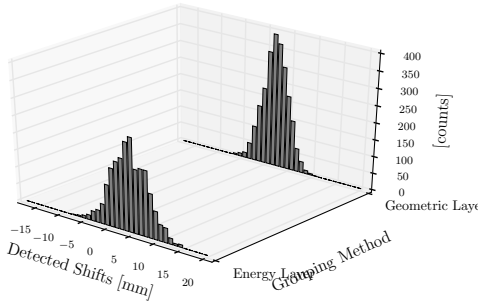
PG emission, spot 40

Energy Layer: $\mu_\Delta = 3.13$, $\sigma_\Delta = 1.16 \pm 0.01$, $P_\Delta = 96\%$
Geometric Layer: $\mu_\Delta = 2.18$, $\sigma_\Delta = 0.89 \pm 0.01$, $P_\Delta = 92\%$



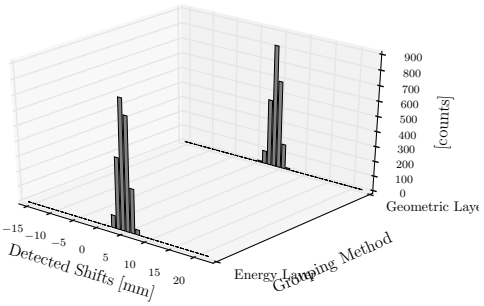
MPS, FOP shift distribution, spot 29

Energy Layer: $\mu_\Delta = 4.20$, $\sigma_\Delta = 3.61 \pm 0.05$, $P_\Delta = 34\%$
Geometric Layer: $\mu_\Delta = 1.80$, $\sigma_\Delta = 2.55 \pm 0.03$, $P_\Delta = 14\%$



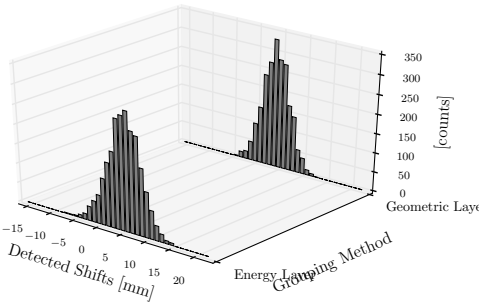
KES, FOP shift distribution, spot 29

Energy Layer: $\mu_\Delta = 4.82$, $\sigma_\Delta = 1.00 \pm 0.01$, $P_\Delta = 99\%$
Geometric Layer: $\mu_\Delta = 3.67$, $\sigma_\Delta = 0.96 \pm 0.01$, $P_\Delta = 99\%$



MPS, FOP shift distribution, spot 40

Energy Layer: $\mu_\Delta = 4.90$, $\sigma_\Delta = 3.38 \pm 0.04$, $P_\Delta = 54\%$
Geometric Layer: $\mu_\Delta = 3.75$, $\sigma_\Delta = 2.71 \pm 0.04$, $P_\Delta = 57\%$



KES, FOP shift distribution, spot 40

Figure 21: In top left the PG emission profiles are plotted for groups about spot 29, with iso-energy and iso-depth groups far left and center left respectively. On the right, the same plots for the groupings about spot 40. Blue lines show the CT PG emission, red the RPCT PG emission. The middle row shows the FOP shift distributions for the two grouping methods for the iso-energy and iso-depth (both with 0.84×10^9 primaries) groups about spot 29. The bottom row shows the results for the spot grouping about spot 40, with 1.07×10^9 primaries for both groupings. For either row the left results were obtained with the MPS camera and those on the right with the KES camera.

	Spot 29		Spot 40		Spot 61	
	iso-energy	iso-depth	iso-energy	iso-depth	iso-energy	iso-depth
Dose	2.47	1.60	6.40	4.07	6.40	6.40
PG emission	4.94	2.32	7.12	3.78	7.12	7.12
PG + PSF	3.20	2.18	4.94	4.51	4.94	6.25
μ_Δ MPS	3.13±1.16	2.18±0.89	4.82±1.00	3.67±0.96	4.72±1.17	5.77±1.05
μ_Δ KES	4.20±3.61	1.80±2.55	4.90±3.38	3.75±2.71	4.15±3.82	5.15±2.87
P_Δ MPS	96%	92%	99%	99%	99%	99%
P_Δ KES	34%	14%	54%	57%	99%	99%
Nr. protons	$0.84 \cdot 10^9$	$0.84 \cdot 10^9$	$1.07 \cdot 10^9$	$1.07 \cdot 10^9$	$1.07 \cdot 10^9$	$0.98 \cdot 10^9$

Table 3: Overview of the iso-energy versus iso-depth comparison: the accuracy can be gleaned by comparing the shifts, the probability of measuring the change is estimated with P_Δ. All shifts are given in units of millimeters. The fact that the PG emission and dose shifts for spot 61 are exactly the same with iso-energy and iso-depth grouping is purely fortuitous.

and dose are quite different.

4 Discussion

Even though cameras with different collimators were used in this study, the goal was not to compare designs per se, but to see the differences in performance in two particular cameras both optimized for FOP precision.

4.1 Spot 61

The lower detection yield for spot 61 may seem puzzling at first, but may be explained with gamma attenuation in the patient. Spot 40 and 61 share a coordinate in the plane transverse to the beam, but differ because they are almost at opposite ends of the other dimension in the transverse plane, which puts spot 61 about 5 cm deeper in the patient seen from the cameras point of view. Lin et al. (2016, Table 3) presented some results with respect to attenuation: a 10 cm increase of path lenth in a PMMA target leads to a 24% detection yield reduction with the MPS camera of that paper, and 39% with their KES.

4.2 Simulation limitations

This study was performed in silico. We used Geant4's QGSP_BIC_HP_EMY physicslist to produce our results. For Geant4 is it known that currently PG production is overestimated; about 30% according to Pinto et al. (2016). This means that the results presented here may be better than in reality. Secondly, we have used the vpgTLE method, which only simulates PGs due to the primary protons; no background estimation is provided, which has not been subjected to calibration with empirical results in any case. The room and nozzle were not modeled. Nevertheless, background estimates were taken from the original papers, and they were both based on experimental measurement.

4.3 Multiple fields and FOP

Range verification is useful for under- and overshoot detection. It does not provide checks for transverse error. If there is an organ at risk (OAR) right behind the (geometrically) distal layer, FOP verification would provide valuable information. The interplay between multiple fields is difficult to disentangle, because an overshoot in one field may be compensated by an undershoot in another field, but that was considered out of scope for this publication. For single field treatment plans, or perpendicular beams where under- and overshoot or not likely to compensate, FOP verification is easily understood.

We can expect different shifts due to the profile's various shapes, changing from spot to spot, and profile differences depending on the particular grouping. Adding the PSF blurs these features and – apparently – can

alter the shift estimate. Range mixing is inherent to spot grouping, and therefore to altering the profile shape, so it is something to be aware of.

4.4 From FOP to Range to Profile measurement

The FOP estimation algorithm presented here was not the primary focus of the study. Possibilities for improvement are likely. As far as these authors could establish, our method was broadly similar to procedures in other publications: a maximum and a floor are extracted from the (smoothed) profile, the difference of which is the FO. Then, the FOP is set to the position where the profile crosses the half-FOA threshold. This means such procedure is sensitive to the floor and maximum estimation. With low statistics, the peak may be misidentified at a noisy bin. For the KES, incorrect alignment to the expected FOP may lead to the background falling outside of the field of view and may lead to a misidentified floor. In our plan, the width of the KES FO was about 4 cm. The angular resolution of collimated cameras implies some blurring and therefore broadening of the FO width. With a FoV of 10 cm and a possible forward shift of perhaps more than a cm (spot 61) it is not uncommon that little if any post BP floor is in the FoV. Taking care to center the camera at the expected FOP (by taking the FOP of the PG emission convolved with a PSF) proved essential for the KES results. The MPS has a clear advantage here.

As mentioned in section 2.4.1, it is important to distinguish range from FOP estimation. True range estimation, determining a FIP as well as a FOP, might improve the quality of the results, because it would allow the clinician to distinguish morphological change from setup shifts. Both would still suffer from the sensitivity to peak and floor estimation. In Roellinghoff et al. (2014) fitting to a reference profile was proposed. Gueth et al. (2013) provided a correlation at registration method where shifted profiles could be distinguished from changed profiles (or bad data). Since in clinical conditions, each spot would require its own reference, the challenge becomes sourcing the references. Schumann et al. (2016) demonstrated that a per-spot PG profile may be obtained by convoluting the per-spot dose profiles. A TPS might then provide such profiles as a matter of course, which would profile the per-spot reference needed. Such profile-matching might make features such as camelbacks and no-single-range spots as in fig. 10 a feature instead of a problem: advanced profile matching might use the shape to infer structural information. A peak due to bone might function as a reference point, even if FIP and FOP change due to weight loss.

4.5 Improved Spot-grouping

Initially we aimed to include a transverse threshold in our grouping procedure: it stands to reason that patient changes are not uniform in the plane transverse to the beam. However, the minimum statistics discovered in the modulated spot analysis proved implacable: this treatment plan prescribes a total of $2.76 \cdot 10^{10}$ protons, which divided by the *minimum* 10^9 protons needed for a reliable (but still not sub-mm accurate) measurement with the MPS camera, leaves us with about 27 possible groups, if spots are grouped only once. At the proximal and distal ends the FOPs are too spread out to preserve FOP sharpness, and in any case this leaves very little room to include both FOP and transverse proximity in the grouping algorithm. Grouping on transverse position only, e.g. integrating the signal over all spots at one transverse position, is not expected to provide an improved FOP estimate, since the highest energy would determine the (distal) FOP, while all the other spots will do no more than elevate the plateau region.

We have no doubt that a more sophisticated grouping method may strike a better balance between proximity, both transverse and longitudinal, and proton count than what we were able to envisage. A fundamental inverse relationship between proton count and spot-pertinence remains: the spots at the fields edges, distal and transverse, are most pertinent to FOP detection, while simultaneously very spread out and therefore difficult to group due to *range mixing*. Moreover, the difficulty of low spot weights may be expected to increase with a trend towards smaller spots in non-isocentric planning (Grevillot et al., 2015).

A possible alternative avenue involves identifying regions of expected deviation. In the contouring stage, areas could be grouped on likelihood of deformation, for instance lung or windpipe. A pre-grouping might then provide a set of spots that should correlate to under- or overshoot in the particular region. Perhaps even the TPS could be made aware, and place weightier, but fewer, spots at the distal end of the sensitive area, so that PG FOP measurement is aided.

4.6 Improved collimated cameras

The detection yields are low with respect to the proton statistics. Since we can not launch more protons, we must improve the detection yield. The most straightforward method for improving the yield is increasing the solid angle. When the KES prototype considered here was first presented, a full ring like in PET detection was envisaged (Smeets et al., 2012, figure 2). While this does not remove the sensitivity to alignment, it seems fair to say that at minimum an order of magnitude better yield can be expected. With the MPS camera, due to its detection efficiency, a factor two or three seems to be the upper limit of the expected yield gains. In terms of absorbers the cameras perform rather similar, so it is unclear if there is room for improvement there.

4.6.1 Other PG camera designs

Compton camera designs are an alternative way to measure the PG profile and is seeing a lot of interest (Hueso-González et al., 2016; Krimmer, 2015; Nurdan et al., 2015), although at this time they are not yet competitive with the conceptually simpler collimated cameras. Apart from position, time of arrival (ToF) and energy provide information about the beam-patient interaction. PG timing (Golnik et al., 2014) cameras measure FOP by converting the ToF to a distance. Such a design lends itself well for a head-to-head comparison to collimated cameras. PG spectroscopy is being developed at Massachusetts General Hospital (Verburg and Seco, 2014), where the FOP is not considered but the target tissue is deduced from the spectrum. A different tissue indicates change. It is expected to be tested under clinical conditions soon (Verburg and Bortfeld, 2016).

5 Conclusion

We have presented the first start-to-end clinical simulations of a multi-parallel slit camera for Prompt Gamma detection. In addition, we have presented the first head-to-head comparison of two collimated PG cameras under realistic conditions. Also, we present a new spot grouping method based on the notion of iso-depth. A small study of clinical spot weights was performed. Finally, a figure of merit is presented that provides an estimate of the probability of a measured FOP falling within or outside of the expected mean $\pm 2\times$ standard deviation.

At the nominal spot weights, neither camera can be expected to measure a correct result. Typical spot weights are studied, and with a downward trend of spot weights for low tolerance plans, it seems unlikely the two PG cameras considered will be able to produce clinically relevant results on the spot level, without significant improvements. Spot grouping is a way to improve the statistics and enable millimetric precision of the FOP estimate. Two spot grouping methods are studied, and only for the KES camera did the precision on the FOP shift improve by using the new iso-depth grouping, which deserves further investigation. A better than two out of three likelyhood to obtain a meaningful indication of the shift (P_{Δ}) is likely for both cameras, when using iso-depth spot grouping instead of grouping by iso-energy layer.

6 Acknowledgements

This work was partly supported by SIRIC LYric Grant INCa-DGOS-4664, LABEX PRIMES (ANR-11-LABX-0063 / ANR-11-IDEX-0007) and Fondation ARC. The authors would like to thank Marie-Claude Biston, Thomas Baudier and Gloria Vilches-Freixas for their help finding the CT images and demonstrating and using the treatment planning software. We also thank Erik Almhagen and Uppsala University Hospital, Sweden for the treatment plan data presented in this paper. Lastly, David Boersma of Skandion, Sweden and Alessio Elia of MedAustron, Austria are thanked for their help in investigating clinical treatment plans.

References

Benjamin M Clasie, Jacob B Flanz, and Hanne M Kooy. Interpolation of tabulated proton Bragg peaks. *Phys. Med. Biol. Phys. Med. Biol*, 57(57):405–409, 2012. ISSN 0031-9155. doi: 10.1088/0031-9155/57/21/N405. URL http://iopscience.iop.org/0031-9155/57/21/N405.

Christian Golnik, Fernando Hueso-González, Andreas Müller, Peter Dendooven, Wolfgang Enghardt, Fine Fiedler, Thomas Kormoll, Katja Roemer, Johannes Petzoldt, Andreas Wagner, and Guntram Pausch. Range assessment in particle therapy based on prompt γ-ray timing measurements. *Physics in Medicine and Biology*, 59(18):5399–5422, sep 2014. ISSN 0031-9155. doi: 10.1088/0031-9155/59/18/5399. URL http://stacks.iop.org/0031-9155/59/i=18/a=5399?key=crossref.5437fcd3059992135ec2113679c7dad6.

L Grevillot, D Bertrand, F Dessy, N Freud, and D Sarrut. A Monte Carlo pencil beam scanning model for proton treatment plan simulation using GATE/GEANT4. *Physics in medicine and biology*, 56(16):5203–19, aug 2011. ISSN 1361-6560. doi: 10.1088/0031-9155/56/16/008. URL http://www.ncbi.nlm.nih.gov/pubmed/21791731.

L Grevillot, D Bertrand, F Dessy, N Freud, and D Sarrut. GATE as a GEANT4-based Monte Carlo platform for the evaluation of proton pencil beam scanning treatment plans. *Physics in medicine and biology*, 57(13):4223–44, jul 2012. ISSN 1361-6560. doi: 10.1088/0031-9155/57/13/4223. URL http://www.ncbi.nlm.nih.gov/pubmed/22684098.

L Grevillot, M Stock, and S Vatnitsky. Evaluation of beam delivery and ripple filter design for non-isocentric proton and carbon ion therapy. *Physics in medicine and biology*, 60(20):7985–8005, 2015. ISSN 1361-6560. doi: 10.1088/0031-9155/60/20/7985. URL http://www.ncbi.nlm.nih.gov/pubmed/26418366.

P Gueth, D Dauvergne, N Freud, J M Létang, C Ray, E Testa, and D Sarrut. Machine learning-based patient specific prompt-gamma dose monitoring in proton therapy. *Physics in medicine and biology*, 58(13):4563–77, jul 2013. ISSN 1361-6560. doi: 10.1088/0031-9155/58/13/4563. URL http://www.ncbi.nlm.nih.gov/pubmed/23771015.

Fernando Hueso-González, Fine Fiedler, Christian Golnik, Thomas Kormoll, Guntram Pausch, Johannes Petzoldt, Katja E. Römer, and Wolfgang Enghardt. Compton Camera and Prompt Gamma Ray Timing: Two Methods for In Vivo Range Assessment in Proton Therapy. *Frontiers in Oncology*, 6(April):1–13, 2016. ISSN 2234-943X. doi: 10.3389/fonc.2016.00080. URL http://journal.frontiersin.org/article/10.3389/fonc.2016.00080.

Brent F B Huisman, J M Létang, É Testa, and D Sarrut. Accelerated Prompt Gamma estimation for clinical Proton Therapy simulations. *Physics in Medicine and Biology*, 61:7725–7743, 2016. ISSN 0031-9155. doi: 10.1088/0031-9155/61/21/7725.

S Jan, G Santin, D Strul, S Staelens, K Assié, D Autret, S Avner, R Barbier, M Bardiès, P M Bloomfield, D Brasse, V Breton, P Bruyndonckx, I Buvat, A F Chatziioannou, Y Choi, Y H Chung, C Comtat, D Donnarieix, L Ferrer, S J Glick, C J Groiselle, D Guez, P-F Honore, S Kerhoas-Cavata, A S Kirov, V Kohli, M Koole, M Krieguer, D J van der Laan, F Lamare, G Largeron, C Lartizien, D Lazaro, M C Maas, L Maigne, F Mayet, F Melot, C Merheb, E Pennacchio, J Perez, U Pietrzyk, F R Rannou, M Rey, D R Schaart, C R Schmidtlein, L Simon, T Y Song, J-M Vieira, D Visvikis, R Van de Walle, E Wieërs, and C Morel. GATE: a simulation toolkit for PET and SPECT. *Physics in Medicine and Biology*, 49(19):4543, 2004. URL http://stacks.iop.org/0031-9155/49/i=19/a=007.

F M F C Janssen, G Landry, P Cambraia Lopes, G Dedes, J Smeets, D R Schaart, K Parodi, and F Verhaegen. Factors influencing the accuracy of beam range estimation in proton therapy using prompt gamma emission. *Physics in medicine and biology*, 59(15):4427–41, aug 2014. ISSN 1361-6560. doi: 10.1088/0031-9155/59/15/4427. URL http://www.ncbi.nlm.nih.gov/pubmed/25049223.

Antje-Christin Knopf and Antony Lomax. In vivo proton range verification: a review. *Physics in medicine and biology*, 58(15):R131–60, aug 2013. ISSN 1361-6560. doi: 10.1088/0031-9155/58/15/R131. URL http://www.ncbi.nlm.nih.gov/pubmed/23863203.

Aafke Christine Kraan. Range Verification Methods in Particle Therapy: Underlying Physics and Monte Carlo Modeling. *Frontiers in Oncology*, 5(July):1–27, 2015. ISSN 2234-943X. doi: 10.3389/fonc.2015.00150. URL http://journal.frontiersin.org/Article/10.3389/fonc.2015.00150/abstract.

Jochen Krimmer. Development of a Compton camera for medical applications based on silicon strip and scintillation detectors. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, pages 98–101, 2015.

Shunsuke Kurosawa, Hidetoshi Kubo, Kazuki Ueno, Shigeto Kabuki, Satoru Iwaki, Michiaki Takahashi, Kojiro Taniue, Naoki Higashi, Kentaro Miuchi, Toru Tanimori, Dogyun Kim, and Jongwon Kim. Prompt gamma detection for range verification in proton therapy. *Current Applied Physics*, 12(2):364–368, 2012. ISSN 15671739. doi: 10.1016/j.cap.2011.07.027. URL http://dx.doi.org/10.1016/j.cap.2011.07.027.

William R Leo. *Techniques for Nuclear and Particle Physics Experiments*. 2nd edition, 1994. ISBN 0-387-57280-5.

Hsin-hon Lin, Hao-ting Chang, Tsi-chian Chao, and Keh-shih Chuang. A comparison of two prompt gamma imaging techniques with collimator-based cameras for range verification in proton therapy. *Radiation Physics and Chemistry*, pages 1–7, 2016. ISSN 0969-806X. doi: 10.1016/j.radphyschem.2016.04.020. URL http://dx.doi.org/10.1016/j.radphyschem.2016.04.020.

Gabriela Llosá, Marco Trovato, John Barrio, Ane Etxebeste, Enrique Muñoz, Carlos Lacasta, Josep F. Oliver, Magdalena Rafecas, Carles Solaz, and Paola Solevi. First Images of a Three-Layer Compton Telescope Prototype for Treatment Monitoring in Hadron Therapy. *Frontiers in Oncology*, 6(February):14, 2016. ISSN 2234-943X. doi: 10.3389/fonc.2016.00014. URL http://journal.frontiersin.org/Article/10.3389/fonc.2016.00014/abstract.

Chul Hee Min, Chan Hyeong Kim, Min Young Youn, and Jong Won Kim. Prompt gamma measurements for locating the dose falloff region in the proton therapy. *Applied Physics Letters*, 89(18):38–41, 2006. ISSN 00036951. doi: 10.1063/1.2378561.

M Moteabbed, S España, and H Paganetti. Monte Carlo patient study on the comparison of prompt gamma and PET imaging for range verification in proton therapy. *Physics in medicine and biology*, 56(4):1063–82, feb 2011. ISSN 1361-6560. doi: 10.1088/0031-9155/56/4/012. URL http://www.ncbi.nlm.nih.gov/pubmed/21263174.

T. Conka Nurdan, K. Nurdan, a.B. Brill, and a.H. Walenta. Design criteria for a high energy Compton Camera and possible application to targeted cancer therapy. *Journal of Instrumentation*, 10(07):C07018–C07018, 2015. ISSN 1748-0221. doi: 10.1088/1748-0221/10/07/C07018. URL http://stacks.iop.org/1748-0221/10/i=07/a=C07018?key=crossref.4049c62b8b6899c51379537eeb8d482d.

Harald Paganetti. Range uncertainties in proton therapy and the role of Monte Carlo simulations. *Phys Med Biol*, 57(11), 2013. doi: 10.1088/0031-9155/57/11/R99.Range.

I Perali, a Celani, L Bombelli, C Fiorini, F Camera, E Clementel, S Henrotin, G Janssens, D Prieels, F Roellinghoff, J Smeets, F Stichelbaut, and F Vander Stappen. Prompt gamma imaging of proton pencil beams at clinical dose rate. *Physics in Medicine and Biology*, 59(19):5849–5871, oct 2014. ISSN 0031-9155. doi: 10.1088/0031-9155/59/19/5849. URL http://stacks.iop.org/0031-9155/59/i=19/a=5849?key=crossref.d1c598721e0b970b36f1c8a6ad1dd1a1.

M Pinto, D Dauvergne, N Freud, J Krimmer, J M Letang, C Ray, F Roellinghoff, and E Testa. Design optimisation of a TOF-based collimated camera prototype for online hadrontherapy monitoring. *Physics in medicine and biology*, 59(24):7653–7674, 2014. ISSN 1361-6560. doi: 10.1088/0031-9155/59/24/7653. URL http://www.ncbi.nlm.nih.gov/pubmed/25415207.

Marco Pinto, Denis Dauvergne, Nicolas Freud, Jochen Krimmer, Jean M. Létang, and Etienne Testa. Assessment of Geant4 Prompt-Gamma Emission Yields in the Context of Proton Therapy Monitoring. *Frontiers in Oncology*, 6(January):1–7, 2016. ISSN 2234-943X. doi: 10.3389/fonc.2016.00010. URL http://journal.frontiersin.org/Article/10.3389/fonc.2016.00010/abstract.

Jerimy C Polf, Stephen Avery, Dennis S Mackin, and Sam Beddar. Imaging of prompt gamma rays emitted during delivery of clinical proton beams with a Compton camera: feasibility studies for range verification. *Phys Med Biol*, 60(18):7085–7099, 2015. ISSN 0031-9155. doi: 10.1088/0031-9155/60/18/7085. URL http://dx.doi.org/10.1088/0031-9155/60/18/7085.

M Priegnitz, S Helmbrecht, G Janssens, and I Perali. Measurement of prompt gamma profiles in inhomogeneous targets with a slit camera. *Physics in Medicine and Biology*, 4849:4849, 2015. ISSN 0031-9155. doi: 10.1088/0031-9155/60/12/4849. URL http://dx.doi.org/10.1088/0031-9155/60/12/4849.

M Priegnitz, S Barczyk, L Nenoff, C Golnik, I Keitz, T Werner, S Mein, J Smeets, F Vander Stappen, G Janssens, L Hotoiu, F Fiedler, D Prieels, W Enghardt, G Pausch, and C Richter. Towards clinical application: prompt gamma imaging of passively scattered proton fields with a knife-edge slit camera. *Physics in Medicine and Biology*, 61(22):7881–7905, 2016. ISSN 0031-9155. doi: 10.1088/0031-9155/61/22/7881. URL http://stacks.iop.org/0031-9155/61/i=22/a=7881?key=crossref.142532baa3a9efe94e4b27e9ecade18a.

Christian Richter, Guntram Pausch, Steffen Barczyk, Marlen Priegnitz, Isabell Keitz, Julia Thiele, Julien Smeets, Francois Vander Stappen, Luca Bombelli, Carlo Fiorini, Lucian Hotoiu, Irene Perali, Damien Prieels, Wolfgang Enghardt, and Michael Baumann. First clinical application of a prompt gamma based in vivo proton range verification system. *Radiotherapy and Oncology*, 118(2):232–237, 2016. ISSN 18790887. doi: 10.1016/j.radonc.2016.01.004. URL http://linkinghub.elsevier.com/retrieve/pii/S0167814016000074.

C Robert, G Dedes, G Battistoni, T T Böhlen, I Buvat, F Cerutti, M P W Chin, a Ferrari, P Gueth, C Kurz, L Lestand, a Mairani, G Montarou, R Nicolini, P G Ortega, K Parodi, Y Prezado, P R Sala, D Sarrut, and E Testa. Distributions of secondary particles in proton and carbon-ion therapy: a comparison between GATE/Geant4 and FLUKA Monte Carlo codes. *Physics in medicine and biology*, 58(9):2879–99, may 2013. ISSN 1361-6560. doi: 10.1088/0031-9155/58/9/2879. URL http://www.ncbi.nlm.nih.gov/pubmed/23571094.

F. Roellinghoff, M. H. Richard, M. Chevallier, J. Constanzo, D. Dauvergne, N. Freud, P. Henriquet, F. Le Foulher, J. M. Létang, G. Montarou, C. Ray, E. Testa, M. Testa, and a. H. Walenta. Design of a Compton camera for 3D prompt-γ imaging during ion beam therapy. *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 648(SUPPL. 1):S20–S23, aug 2011. ISSN 01689002. doi: 10.1016/j.nima.2011.01.069. URL http://linkinghub.elsevier.com/retrieve/pii/S0168900211001471.

F Roellinghoff, a Benilov, D Dauvergne, G Dedes, N Freud, G Janssens, J Krimmer, J M Létang, M Pinto, D Prieels, C Ray, J Smeets, F Stichelbaut, and E Testa. Real-time proton beam range monitoring by means of prompt-gamma detection with a collimated camera. *Physics in medicine and biology*, 59(5):1327–38, 2014. ISSN 1361-6560. doi: 10.1088/0031-9155/59/5/1327. URL http://www.ncbi.nlm.nih.gov/pubmed/24556873.

David Sarrut, Manuel Bardiès, Nicolas Boussion, Nicolas Freud, Sébastien Jan, Jean-Michel Létang, George Loudos, Lydia Maigne, Sara Marcatili, Thibault Mauxion, Panagiotis Papadimitroulas, Yann Perrot, Uwe Pietrzyk, Charlotte Robert, Dennis R Schaart, Dimitris Visvikis, and Irène Buvat. A review of the use and potential of the GATE Monte Carlo simulation code for radiation therapy and dosimetry applications. *Medical Physics*, 41(6), 2014. doi: http://dx.doi.org/10.1118/1.4871617. URL http://scitation.aip.org/content/aapm/journal/medphys/41/6/10.1118/1.4871617.

A Schumann, M Priegnitz, S Schoene, W Enghardt, H Rohling, and F Fiedler. From prompt gamma distribution to dose: a novel approach combining an evolutionary algorithm and filtering based on Gaussian-powerlaw convolutions. *Physics in medicine and biology*, page 6919, 2016. ISSN 0031-9155. doi: 10.1088/0031-9155/61/19/6919. URL http://dx.doi.org/10.1088/0031-9155/61/19/6919.

J Smeets, F Roellinghoff, D Prieels, F Stichelbaut, A Benilov, P Busca, C Fiorini, R Peloso, M Basilavecchia, T Frizzi, J C Dehaes, and A Dubus. Prompt gamma imaging with a slit camera for real-time range control in proton therapy. *Physics in medicine and biology*, 57(11):3371–405, 2012. ISSN 1361-6560. doi: 10.1088/0031-9155/57/11/3371. URL http://www.ncbi.nlm.nih.gov/pubmed/22572603.

Paola Solevi, Enrique Muñoz, Carles Solaz, Marco Trovato, Peter Dendooven, John E Gillam, Carlos Lacasta, Josep F Oliver, Magdalena Rafecas, and Irene Torres-Espallardo. Performance of MACACO Compton telescope for ion-beam therapy monitoring: first test with proton beams. *Physics in Medicine and Biology*, 61 (14):5149–5165, 2016. ISSN 0031-9155. doi: 10.1088/0031-9155/61/14/5149. URL http://stacks.iop.org/0031-9155/61/i=14/a=5149?key=crossref.6bf836448c34a4bc8348edeba28d35c4.

G Sportelli, N Belcari, N Camarlinghi, G a P Cirrone, G Cuttone, S Ferretti, a Kraan, J E Ortuño, F Romano, a Santos, K Straub, a Tramontana, a Del Guerra, and V Rosso. First full-beam PET acquisitions in proton therapy with a modular dual-head dedicated system. *Physics in medicine and biology*, 59(1):43–60, dec 2013. ISSN 1361-6560. doi: 10.1088/0031-9155/59/1/43. URL http://www.ncbi.nlm.nih.gov/pubmed/24321855.

E Sterpin, G Janssens, J Smeets, François Vander Stappen, D Prieels, Marlen Priegnitz, Irene Perali, and S Vynckier. Analytical computation of prompt gamma ray emission and detection for proton range verification. *Physics in Medicine and Biology*, 60(12):4915–4946, 2015. ISSN 0031-9155. doi: 10.1088/0031-9155/60/12/4915. URL http://stacks.iop.org/0031-9155/60/i=12/a=4915?key=crossref.aabd8815e135401a22d165e343a7bac4.

E. Testa, M. Bajard, M. Chevallier, D. Dauvergne, F. Le Foulher, N. Freud, J. M. Letang, J. C. Poizat, C. Ray, and M. Testa. Monitoring the Bragg peak location of 73 MeVu carbon ions by means of prompt ?? -ray measurements. *Applied Physics Letters*, 93(9):1–10, 2008. ISSN 00036951. doi: 10.1063/1.2975841. URL http://arxiv.org/abs/0809.0185http://dx.doi.org/10.1063/1.2975841.

P.G. Thirolf, S. Aldawood, M. Böhmer, J. Bortfeldt, I. Castelhano, G. Dedes, F. Fiedler, R. Gernhäuser, C. Golnik, S. Helmbrecht, F. Hueso-González, H. v.d. Kolff, T. Kormoll, C. Lang, S. Liprandi, R. Lutter, T. Marinšek, L. Maier, G. Pausch, J. Petzoldt, K. Römer, D. Schaart, and K. Parodi. A Compton camera prototype for prompt gamma medical imaging. *EPJ Web of Conferences*, 117:05005, 2016. ISSN 2100-014X. doi: 10.1051/epjconf/201611705005. URL http://www.epj-conferences.org/10.1051/epjconf/201611705005.

J Verburg and T Bortfeld. TU-FG-BRB-06: A Prompt Gamma-Ray Spectroscopy System for Clinical Studies of in Vivo Proton Range Verification. *Medical Physics*, 43(6):3757–3757, jun 2016. ISSN 00942405. doi: 10.1118/1.4957546. URL http://doi.wiley.com/10.1118/1.4957546.

Joost M Verburg and Joao Seco. Proton range verification through prompt gamma-ray spectroscopy. *Physics in Medicine and Biology*, 59(23):7089–7106, 2014. ISSN 0031-9155. doi: 10.1088/0031-9155/59/23/7089. URL http://stacks.iop.org/0031-9155/59/i=23/a=7089?key=crossref.2c380dcfa8360cb6e6defe62dd49f24a.

5.8 Notes on the article

5.8.1 Comparison to other comparative studies

An earlier comparison between two collimated PG camera designs was presented in Lin et al. (2016). That paper was a geometrical study into the difference of the two methods of collimation without optimisation, which may be different for the two types. As such, not actual (optimized) prototypes were compared; only the collimators were swapped. In addition, the target was a "perfect line beam" in a PMMA block. All plots have 1 mm binning for the PG profiles, which indicates the mechanical and electronic constraints of a detection stage may not have been implemented. Although an interesting study, these points are significant reasons to be careful drawing conclusions about the comparative efficacy of either a knife-edge or multi-parallel-slit design.

5.8.2 Treatment planning

In particle therapy and particle therapy imaging literature we often see it mentioned that a spread out Bragg Peak (SOBP) is achieved by giving the most distal energy layer the highest intensity and each successive energy layer is of lower energy and of lower intensity. This view makes sense for PTVs with a rectangular shape and rectangular targets, but of course patients do not fit that model. The surface where the beam impinges the patient is, seen perpendicular to the beam path, not flat, but curved, as is the entrance and exit surfaces of the PTV. This leads to the effect that the geometrically distal layer requires various energies (energy layers). In other words, energy-layers alone are not a suitable proxy for geometrical position, and therefore not for range of FOP verification.

5.8.3 Fall-off retrieval procedure

Range estimation using PGs requires observing the PG production rate along the beam path. This is the guiding design principle behind the collimated PG cameras: a 1D line profile is obtained. PG range estimation then requires that a single range is extracted from such a 1D PG profile. Multiple approaches have been proposed (Gueth et al., 2013; Janssen et al., 2014; Roellinghoff et al., 2014; Smeets et al., 2012; Sterpin et al., 2015), all ad hoc, and their performance is usually presented on a spot composed of 10^8 protons, which, as seen in the previous section, is an exceptional condition in modern treatment plans. Moreover, the algorithms are tested on 10^8 particle spots in artificial and thus relatively simple phantoms, sidestepping clinically common inhomogeneous circumstances. The concept of a single range per spot does not hold up in laterally inhomogeneous targets, and no fit procedure accounts for such situations. For the IBA Knife-Edge camera a fit algorithm is proposed Smeets et al. (2015) for range retrieval from full energy layers, due to the passive scattering nozzle at OncoRay, Dresden. By integrating the PG yields over an energy layer, the issue of low yields is solved, and inhomogeneity side-stepped: the signals recorded are convolutions of all lateral spots in the layer, which diminishes localized range differences but improves the ability of reconstruction. Another point that all fit methods invented for the IBA knife-edge design share, is the requirement that the camera is centered on the expected fall-off position.

As a small illustration, figure 5.1 shows how the choice of parameters affects the obtained fall-off position for certain spot groups in an early version of our FOP estimator. We hoped to see convergence towards a single value, but it proved tricky to ensure this worked well for an arbitrary spot (group).

5.8.4 Absolute Fall-off determination

As mentioned in the Camera alignment section under the Knife-Edge verification, correct alignment of the camera is important to avoid the edge-effects mentioned in Sterpin et al. (2015, Section IV.A.3).

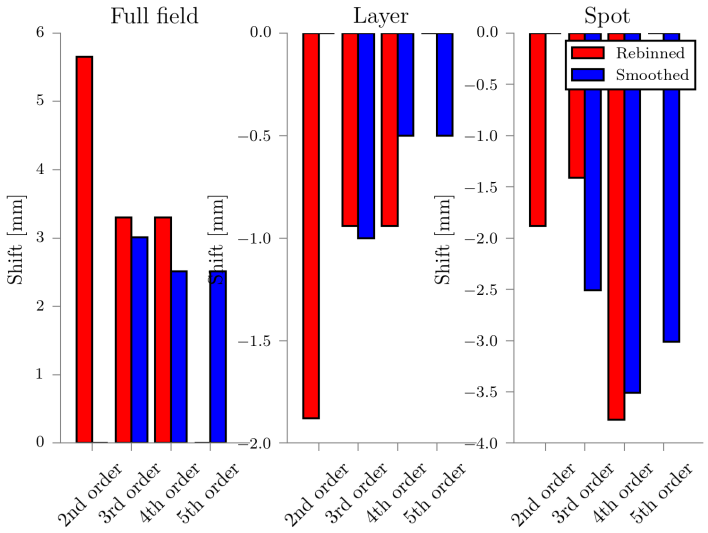


Figure 5.1: Performance of spline fit of various orders (labels on x-axis) on the PG yield of a full field (left), a selected energy layer (center) and a selected spot (right). The collimated camera used in the simulation detected the position of the PG with 1 mm bins. In red the retrieved range is obtained by first binning the data to 8 mm (as in the IPNL collimated camera) and in blue data is smoothed before the fit is applied. Plotted is the PG fall-off shift between a treatment plan irradiated on the planning CT and a RPCT recorded a few weeks into the treatment. This patient had seen serious weight loss, causing the plans as a whole to shift forward (decrease in subcutaneous fat), but selected layers and spots to decrease (the wind pipe was more closed due to tumor growth in the RPCT). While the retrieved range difference converges for the full field and energy layer, for the spot the noise prevents convergence to any particular range.

Ensuring correct positioning, was trickier than anticipated. After ensuring our simulations and Gate correctly placed the geometry at the intended locations, figure 5.2 could be produced.

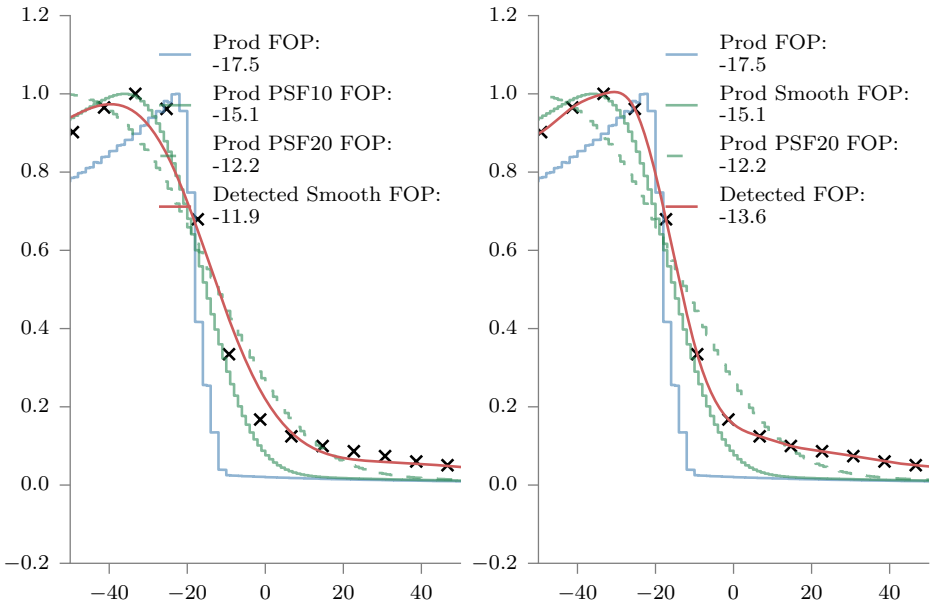


Figure 5.2: On the left and right the PG production (Prod) is given in blue, that signal convoluted with a Point Spread Function (PSF) with smoothing parameter 10 mm and 20 mm (dashed) in green, and the detected response in black crosses. Red are the fitted curves, on the left with $\lambda = 2$ and on the right $\lambda = 0$. The cost of smoothing, required especially for lower statistics, comes at the price of a FOP shift.

Slit cameras have collimators and collimators have a finite depth. The depth and the distance to the photosensitive detector determine the penumbra of the detected signal. For a perfect collimator a tophat-function, with the width of the slit, is expected, while for real collimators the tophat 'sags' into a roughly trapezoid function. Scatter in the collimator smooths the trapezoid out to some degree, approaching roughly a Gaussian. We simulate that effect by convolving the PG production curve with a Point Spread Function (PSF) with smoothing parameter 10 mm and 20 mm. Depending on the smoothing factor λ in eq. 1 in the paper, a curve is obtained on which we compute the FOP according to the procedure described earlier. As figure 5.2 shows, both adding the smoothing kernel to the PG production profile and the smoothing kernel fitted through the measured data each introduce their own forward change in the FOP. Here is it demonstrated for a pencil beam in a waterbox, the same used to validate the knife edge results in the paper. In real circumstances, the noise of the detected signal makes the bias even larger, and with the KE camera again larger than for the MPS camera, perhaps also because we have not reached sufficient statistics even with 10^9 primaries and our results are noise dominated. The slight tail for the datapoints past $x = 0$ may be explained by high energy gammas penetrating the collimator leaves.

In conclusion, an improved reconstruction might include a smoothing factor λ dependent on total PG counts, because that implies a more converged signal and less need to smooth out noise. Note that this bias is consistent, so when searching for FOP *differences*, the effect does not matter. Only for cameras that require precise FOP alignment is this an issue.

5.8.5 Improvements to vpgTLE

At the time of writing this thesis, colleague Thomas Cajgfinger has implemented Fixed Forced detection and an Angular Responce Function to speed up generating PET detector output, within Gate. The premise of this technique is to save computation time by propagating particles only if they will be detected, and then to skip propagation through the collimator and absorber by having a function that relates incident particle positions, angles and energies to the detector response. Roughly speaking, the detector response is estimated and stored for later use, after which a plausible signal is generated with the correct magnitude and noise as would have an analog Monte Carlo simulation. A master student, Gunjan Kayal, was attracted to validate the procedure for Stage 2 of vpgTLE. If successful, it should be made available in Gate in the summer or autumn of 2017.

Chapter 6

Conclusion

6.1 Summary of accomplishments

The vpgTLE variance reduction method was implemented and validated as part of this thesis program. The method was found to be compatible with inhomogeneities, flexible in terms of output and offers a speed-up of a factor 10^3. We also noted that no work has been done to validate PG backgrounds in PG simulations, which is essential if simulations are used to predict real world results. The vpgTLE method is publicly available as part of Gate version 7.1 and up, enabling researchers from over the globe to benefit without any further requirement. Some researchers have already shown interest and I eagerly await any results that may be obtained with this method.

Secondly, the first fully clinical proton therapy simulation monitored by PG collimated detectors was presented, in which a real clinical deviation was investigated. vpgTLE made it possible to complete this study in a reasonable time frame, and enabled spot-by-spot analysis. A new spot-grouping method was presented and demonstrated, and we think the study will move the discussion in the PG monitoring forwards in terms of how these detectors can be used to obtain clinically relevant results. Based on a study of treatment plans, spot-by-spot PG verification is unlikely with the current state of PG collimated cameras, but we showed that grouping geometrically close spots does not change the σ of the FOP appreciably and thus spot grouping enables PG verification.

The title of the PhD research proposal is "Towards real-time treatment control in protontherapy using prompt-radiation imaging: simulation and system optimization". To what extent did we advance towards that goal? I think we did, by producing a faster tool for PG simulations, and showing how it can be used to study various aspects of PG detection, and it helped us identify areas in which improvements are necessary (PG profile processing and spot grouping).

6.2 Further thoughts

If I would have 6 months more, I would take on the research of Chapter 5 in a more systematic way. Right now we depended on a spot selection method resulting in three spots, which we studied in detail. Perhaps it was necessary to start small, because we learned a lot even from only three spots, but in the end a more statistical approach where we obtain FOPs, their precisions and FOP-shift precisions for all spots in the plan would have made us more confident on whether or not we happened to choose outliers or not. Currently, the main limitations are disk space required for the intermediate image (one PG source is ≈1.5 GB, so if a PG source for each individual spot is desired, 1.5 TB of disk space is required) and the simulation time of vpgTLE Stage 2 (the propagation of photons into the detector). The first stage needs to be only run once, and takes about 3 hours for 10^6 primary protons (equivalent to 10^9 protons with analog MC) on one core of a 4 year old laptop (Intel Core i7-3740QM CPU @ 2.70 GHz), while stage two takes ≈ 14 hours, on a single core, to reach the statistics for 10^9

protons. Reducing the size of the intermediate image may be possible by exploiting the sparseness of the 4D image, as well as reducing the voxel sizes in dimensions that are not measured, and reducing the sampling resolution of the PG production cross sections in Stage 0. A master student was attracted to help validate a recently developed Fixed Forced Detection and Angular Response Function components in Gate as a replacement for Stage 2. This will hopefully speed-up the slowest part of a vpgTLE simulation, enabling even faster iteration of research simulations. It would have been interesting to simulate the detection profiles per spot, which could then be summed in post-processing as desired, depending on the precise grouping method. We iterated over the grouping method, and that required generation of new sources and simulation of the 50 CT and 50 RPCT Stage 2 simulations, slowing us down a bit with respect to finding an optimal grouping strategy. In addition, this was one patient with one field of one plan; we did gather more CT+RPCT images but did not have time to perform analysis across plans.

6.3 Clinical use of Prompt Gamma detection

We have seen various PG detector designs, each revealing different properties of the PG. Heavier ions produce more PGs: a carbon primary has a PG production probability of about 40% whereas for protons this number is 5-10%, which must be balanced by the fact that fewer carbons are required to reach a certain dose (a factor 5 fewer carbon ions is used in a comparative study with protons in Pinto et al. (2015)). Perhaps because the first step in designing a detector is whether or not it can be done at all, not much thought has gone into how a clinician could use these outputs. Clinicians working in radiotherapy and radiology are used to work with (CB)CT and occasionally MRI images. Adding a modality is a major challenge, because thousands of clinicians would have to be trained in understanding it. Moreover, at this time providing PG detection outputs in the form of images seems a distant prospect. It also may turn out spot-by-spot statistics are too low and collecting data over larger spot groups is a necessity. What kind of spot groups make most clinical sense? It can be imagined that by grouping spots in geometric proximity and in clinically related structures are a way to provide more targeted feedback. Perhaps a treatment plan could output a list of range-sensitive spots stopping near AORs, or crossing structures likely to deform. A geometric criterion seems a good way to ensure coherence between the deviations of the spots within a group, instead of canceling out when spots in different structures suffering from different deformations are grouped.

If we keep in mind the outputs of collimated and timing cameras, the point of emission of the PG and from that an inferred fall-off position and perhaps range, the information could be presented in terms of thresholds or tolerances: how many spot or spot groups overshoot and undershoot? Perhaps we can measure the full PG profile, and together with a planned profile (Schumann et al., 2016) we could perform profile matching (Gueth et al., 2013) to produce per spot(group) a likelihood of shift and likelihood of no shift estimates. Going even further, the profiles might be converted back to dose, as demonstrated in Schumann et al. (2016). Then, the dose can be summed over all spot(group)s, from different fields, and overlaid on a CT. That would provide the type of image data clinicians are familiar with, and can work with right away. A dose-passing rate of a gamma-index are established methods to compare dosemaps. If a CBCT is present in the treatment room, together with improved deformable registration techniques (Park et al., 2017; Yan, 2010), a CT of the day can be produced for more accurate matching of the PG-derived dosemap and the patients' structures. Supplement this with adaptive particle therapy, also based on the CT+CBCT deformable registration (Veiga et al., 2015), real gains may be within reach in terms of lowering acceptable tolerance margins.

Radiomics is an upcoming medical method constituted on the idea that machines can now be trained to deduce *much more* from clinical data such as CTs (Yip and Aerts, 2016). Based on the vast volumes of medical data that every hospital collects during the patients treatments, a big data approach may result in software that can learn to extract more from the available data than (human)

clinicians currently can. Machine learning as proposed in Gueth et al. (2013) seeks to use such methods for PG to fall-off reconstruction. Unfortunately, the amount of available training data for PG detection is extremely limited, but it is a promising start to have PG detector monitor treatments as a matter of course (Richter et al., 2016; Verburg and Bortfeld, 2016). A similar approach might be considered for treatment planning: software may be much better at real-time dose painting based on live data from a CT+CBCT deformable registration before and PG detection during treatment.

Appendices

Appendix A

Résumé Français

A.1 L'Hadrontherapie

Les cancers sont des cellules cancéreuses anormales malignes qui limitent le fonctionnement humain normal et peuvent entraîner un handicap permanent ou la mort. La plupart des cancers sont environnementaux (pollution, choix de mode de vie, infections) tandis que certains sont dus à une prédisposition héréditaire (International Agency for Research on Cancer, 2014). Avec l'âge, divers facteurs de risque s'accumulent, ce qui explique l'incidence plus élevée de cancer chez les personnes âgées. Avec les progrès continus dans les soins de santé, le cancer est de plus en plus la principale cause de décès. Des mesures préventives peuvent arrêter ou ralentir le développement de certains types de cancers. Si un cancer parvient à se développer, divers types de contre-mesures existent pour le combattre. Les traitements les plus courants comprennent la chirurgie, les produits chimiques anticancéreux ou les médicaments, éventuellement aidés par des biomarqueurs, et la radiothérapie (RT). Souvent, diverses modalités sont combinées pour améliorer les chances de récupération.

Dans la RT, les dépôts de rayonnements ionisants se déposent dans des tissus, où la dose quantifie le dommage au tissu. Le rayonnement peut être appliqué en introduisant des sources radioactives dans le corps, par ingestion, injection ou placement par chirurgie. La radiothérapie par faisceau externe (EBRT), la forme de radiothérapie la plus courante, est appliquée en dirigeant des faisceaux de rayonnements ionisants à travers le corps. Les cancers localisés sont communément appelés tumeurs (malignes), et l'EBRT exige que le cancer soit localisé, car il cible les régions spatiales. Environ la moitié de toutes les tumeurs impliquent un traitement par EBRT. Différents types de rayonnements ionisants sont ou ont été utilisés, mais de loin le plus commun est le rayon X: les photons avec des énergies dans la gamme de 1 à 20 méga de volt électronique (MeV). Les rayons X ont une autre caractéristique utile: ils peuvent être utilisés pour imaginer le corps car différents tissus atténuent le faisceau différemment.

La thérapie par particules est un type d'EBRT plus récent, et il se réfère à l'utilisation médicale d'ions accélérés. *Hadron* ou *ion therapie* le différencierait plus correctement de la thérapie par rayons X du point de vue de la physique, mais comme la thérapie implique de multiples disciplines (physique, biologie, médecine), le mot *particule* Semble être le terme le plus largement compris pour le différencier du traitement par rayons X. En dehors des installations expérimentales, deux espèces d'ion sont utilisées dans la thérapie par particules médicales: les protons (environ 90% des patients traités) et les carbones (10%). Dans le reste de cette thèse, le terme *thérapie de particules* sera utilisé.

La thérapie par particules montre un nombre croissant de patients traités (figure A.1). Les particules sont accélérées dans les cycles synchro ou cyclotrons et ensuite «jetées» dans la tumeur. Cette modalité peut être considérée comme un spin-off de la physique des énergies élevées, qui utilise des faisceaux de particules pour sondage de la matière et des interactions physiques. Les accélérateurs médicaux ne sont conceptuellement pas différents des accélérateurs physiques expérimentaux. Il

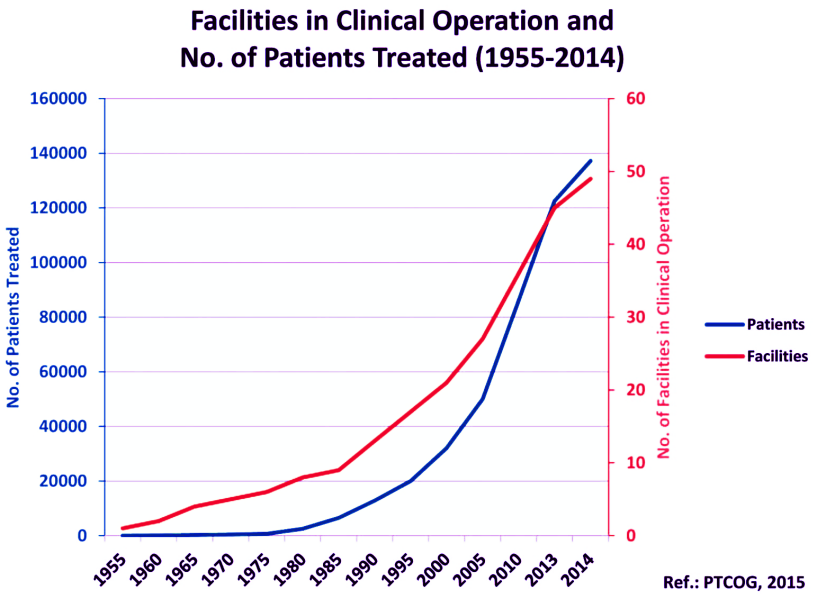


Figure A.1: Les cliniques de particules et les patients traités par thérapie de particules en fonction du temps (Jermann, 2015).

existe une raison physique très intéressante pour choisir des particules sur des photons dans l'EBRT: ces ions présentent une gamme définie dans la matière, par opposition à une diminution exponentielle de la dose par absorption, comme on le voit avec les rayons X (figure **??**) . De plus, juste avant la fin de leur gamme, les pointes de dose déposées: le pic de Bragg (BP). En plaçant cette pointe sur le site de la tumeur, pour un niveau de dose donné dans la tumeur, le tissu environnant reçoit une dose réduite par rapport au traitement par rayons X.

EBRT cible une région spatiale, et pour les photons, cela signifie tous les tissus dans le chemin des poutres. Pour déconcentrer la dose dans le tissu sain entourant la tumeur, les traitements par rayons X comportent habituellement des faisceaux à partir de directions multiples, de champs, en faisant tourner un portique ou en déplaçant la table du patient (radiothérapie à modulation d'intensité, IMRT). En thérapie par particules, le nombre de champs peut être réduit à un ou deux, alors que IMRT peut avoir plus de 10.

En résumé, la thérapie par particules semble être une amélioration par rapport à la thérapie par rayons X (fig. A.3). Un inconvénient important à ce moment est le coût du traitement: la thérapie par particules nécessite des accélérateurs de particules coûteux. Une installation de traitement de protons coûtera, au moment de l'écriture, entre une et deux ordres de grandeur plus qu'une installation de traitement par rayons X et des installations capables d'accélérer encore plus les ions plus lourds comme le carbone. De plus, une importante expansion de l'hôpital lui-même est souvent requise pour héberger l'accélérateur, et il faut également recruter plus de personnel pour faire fonctionner l'accélérateur. Des travaux sont réalisés pour rendre possibles des accélérateurs de particules médicales plus compacts et abordables. Les «solutions de chambre simple», le Varian ProBeam et l'IBA Proteus One sont des exemples de ces conceptions.

A.2 Incertitudes

Un défi dans l'EBRT en général et la thérapie de particules en particulier est de développer notre compréhension de la façon dont le rayonnement affecte les tissus, comment les rayonnements peuvent être appliqués dans les bons endroits, et pas sur les faux, et quels effets secondaires sont inévitables

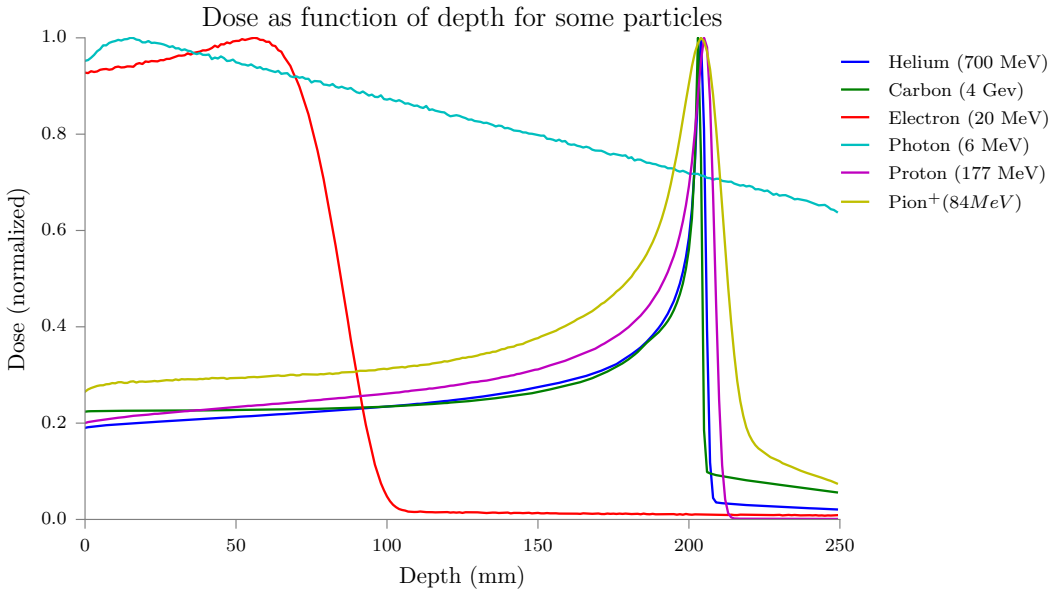


Figure A.2: Profils de profondeur pour particules communes dans la matière. Notez les queues après le pic de Bragg pour les ions plus lourds. Obtenu par simulation avec Gate / Geant4.

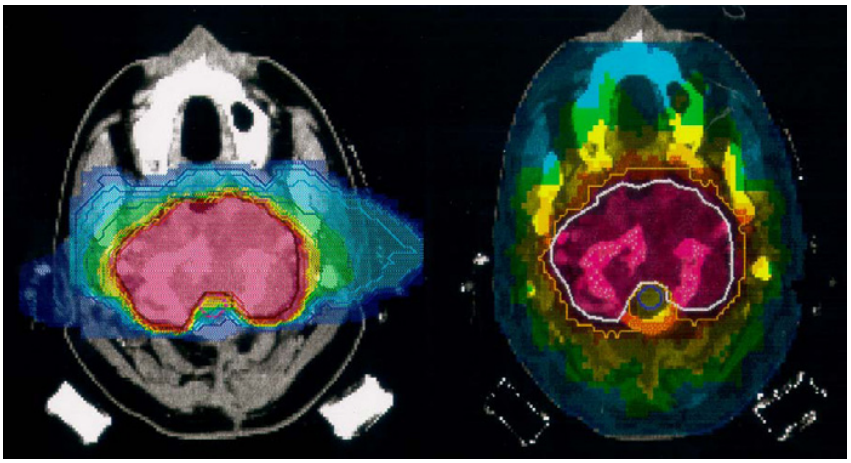


Figure A.3: À gauche, une irradiation au carbone utilisant 2 champs. À droite, un plan de traitement pour une irradiation par rayons X utilisant 9 champs. En utilisant moins de champs, le plan de carbone gère une administration de dose beaucoup plus confinée. Notez qu'en plus, le plan de carbone dépasse pratiquement aucune dose dans les nerfs optiques, ce qui est un type de tissu bien connu extrêmement sensible aux rayonnements. Source: Amaldi and Kraft (2005).

ou comment ils peuvent être Minimisé. Le traitement robuste et réussi et la survie des patients sont les objectifs, mais doivent être divisés en une multitude de facteurs qui l'influencent si nous voulons améliorer la façon dont nous traitons le cancer. La figure A.4 est un résumé approximatif des incertitudes rencontrées dans la thérapie *proton* et calcule leur gravité par ordre décroissant. L'ordre des causes est différent pour différentes particules; Pour la thérapie au carbone, l'incertitude due à RBE serait plus élevée par exemple.

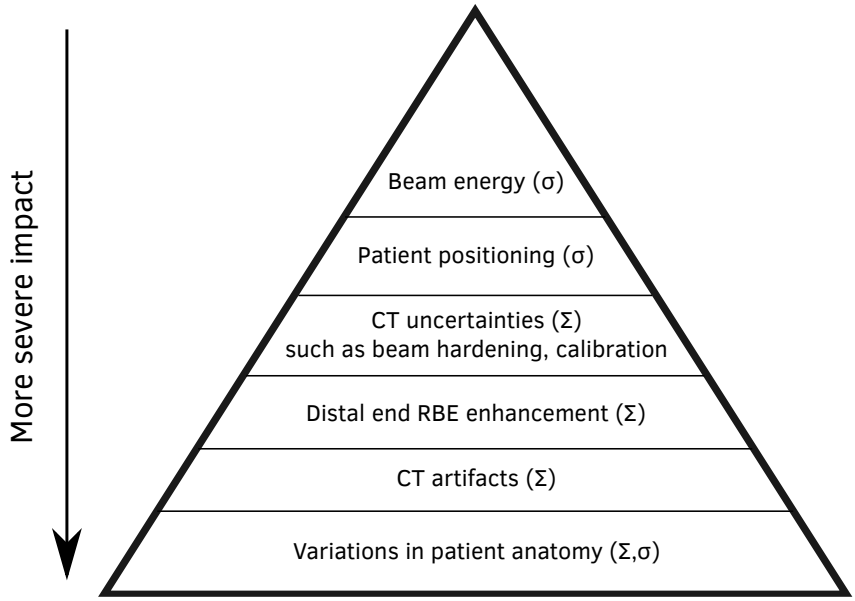


Figure A.4: Le Triangle des incertitudes des Bermudes, tel que présenté par Tony Lomax à l'école de Hadronothérapie à CNAO, 2014. Σ se réfère à des incertitudes systématiques et σ à des effets statistiques.

Actuellement, les incertitudes empêchent tout potentiel de traitement des particules. L'exploitation de la grande précision du traitement des particules est un équilibre entre la robustesse et l'épargne tissulaire. Le principal défi est que la dose n'est pas observable et certainement pas en temps réel pendant les traitements cliniques. Cependant, un traitement de particules peut être considéré comme une expérience de physique fondamentale, où par une bonne modélisation des interactions, la cible, le faisceau, un détecteur et les secondaires peuvent être utilisés pour déduire des quantités pertinentes lorsque le traitement a été ou est mené. Cette façon de recueillir les commentaires de traitement activés par la simulation est cruciale pour les comparaisons entre signaux attendus et mesurés. Proton CT peut aider à éliminer les erreurs de conversion du flux de travail de planification de traitement. La radio-thérapie adaptative vise à combiner les mesures avec une planification de traitement améliorée, et de nouvelles méthodes telles que les theranostics cherchent à élargir la précision accrue en s'assurant que la dose est localisée et suivie pendant le traitement. Étant donné que les traitements par particules sont moins robustes en ce qui concerne le traitement par rayons X, la surveillance et éventuellement le contrôle sont pertinents pour réduire les marges tout en améliorant la robustesse en même temps.

Une notion évoluée d'ART est l'IGRT, la radiothérapie guidée par image. L'IGRT prend un peu plus loin: l'imagerie, en quelque sorte, pourrait être une procédure normale pour tout traitement. Cela nécessite des dispositifs d'imagerie dans la salle de traitement. Portal Imaging (van Elmpt et al., 2008) ou EPIM dosimétrie (Rozendaal et al., 2015) sont des implémentations dans lesquelles l'absorption des rayons X est mesurée pendant le traitement et, par conséquent, une certaine information sur la déposition de la dose est reconstruite. Pour les traitements de particules, PET (Dendooven et al., 2015) est un moyen d'image de l'activation du tissu pendant ou immédiatement après

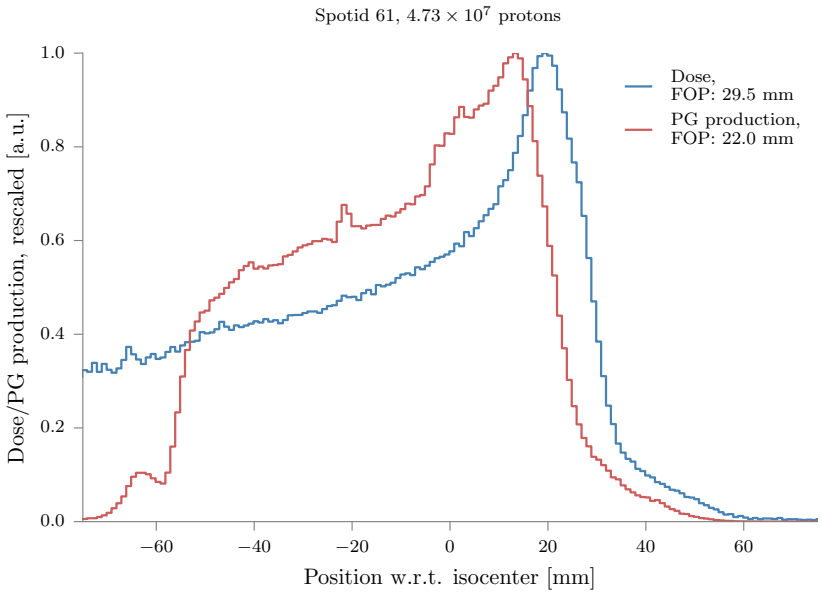


Figure A.5: Une simulation clinique de thérapie de proton dans Gate/Geant4 où 10^6 protons sont lancés et la dose résultante et les distributions PG sont tracées, rééchelonnées à un maximum de 1.

le traitement. Bien que la fragmentation soit généralement considérée comme un inconvénient, Kraft (2000) la décrit dans le cas des ions carbone comme avantage potentiel: généralement β^+ émettant ^{10}C et ^{11}C Sont produits qui pourraient être utilisés pour la vérification directe avec un scanner PET. L'orientation de l'IRM en ligne montre beaucoup d'intérêt dans la communauté de thérapie par rayons X (Lagendijk et al., 2016), et certaines études in silico démontrent qu'elle est également possible dans la thérapie des particules (Hartman et al., 2015). Des travaux en phase précoce sont réalisés sur l'imagerie neutronique, un sous-produit naturel du traitement des particules et donc aucune dose supplémentaire donnée sur le patient (Valle et al., 2016). Le suivi des protons secondaires (Henriquet et al., 2012) en thérapie aux ions lourds est à l'étude. L'intégration précitée du PET dans le portique de livraison a conduit les gens à spéculer que l'imagerie PET pourrait être réalisable (Parodi et al., 2005), et il y a quelques années, la première acquisition de PET en faisceau lors d'un faisceau complet a été présentée (Sportelli et al., 2013).

A.3 Gammas Prompts

Ce section détaille l'état des applications cliniques de Gamme rapide (PG). Comme mentionné au section 1, les PG sont des photons libérés dans les interactions nucléaires entre les particules et les noyaux entrants et secondaires dans la cible. Figure A.5 montre comment la dose (physique) et la production de PG sont corrélées, pour un seul point dans un traitement simulé de proton clinique. La corrélation est évidente. Peut-être que la différence la plus significative est la fin du signal: si nous définissons la position d'arrêt (FOP) comme $\frac{\text{max - ligne de base}}{2}$ (voir fig. A.6), puis le PG FOP est environ un demi-centimètre avant la dose FOP pour cette tache distale. Sous l'influence de l'énergie primaire et de la composition du matériau cible, la profondeur de pénétration et donc la FOP vont changer. Les PG sont produits dans les interactions nucléaires (fig. A.7), mais seulement si la particule incidente a suffisamment d'énergie. À l'extrémité distale, le primaire n'a que quelques MeV à gauche et cela ne suffit pas à la fois pour surmonter le potentiel électrique et produire des fragments tels qu'une PG, d'où la différence de PG et de dose de FOP. Étant donné que la PG est produite dans une interaction nucléaire, le délai entre l'interaction des ions cible primaire et la production PG est

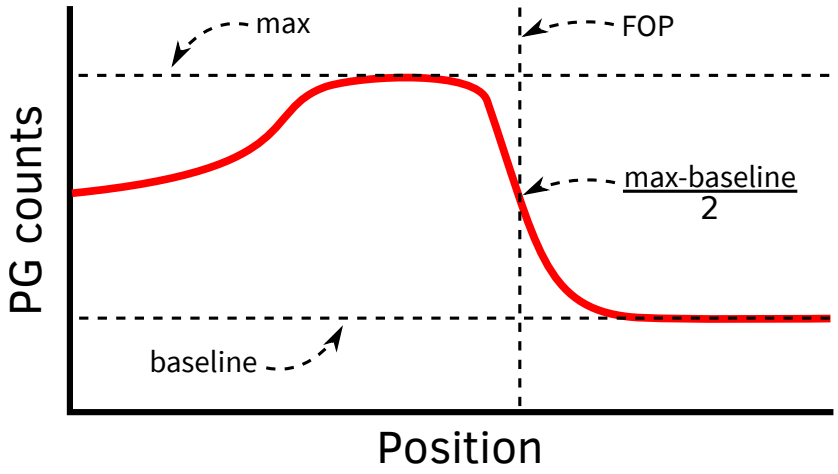


Figure A.6: Une mesure de profil PG hypothétique le long de la ligne du faisceau (rouge). Les lignes pointillées indiquent le maximum, la ligne de base et la position de chute, ou FOP, qui est obtenue en dessinant une ligne verticale à la position où $\frac{\text{max - ligne de base}}{2}$ coupe le profil. Notez que $\frac{\text{max - ligne de base}}{2}$ est un choix arbitraire.

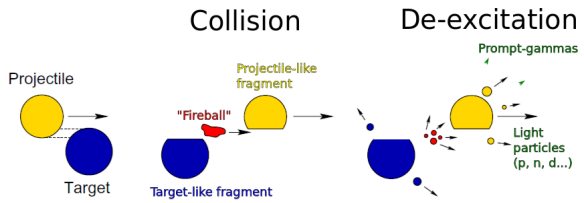


Figure A.7: Une illustration schématique d'une interaction nucléaire entre un ion incident et un ion dans la cible. Ici, une collision frontale partielle est esquissée, où une boule de feu est coupée à la fois des ions entrants et cibles. Cette boule de feu hérite parfois d'une grande partie de l'énergie cinétique de l'ion incident, ce qui lui permet d'avoir une portée plus longue que la BP des ions primaires, ce qui entraîne la dose-queue. La cible, le projectile et la boule de feu peuvent se détendre grâce aux émissions de PG.

bien inférieur à 1 ns et provient directement du chemin des particules. Les PG sont donc un très bon indicateur des pistes primaires et c'est pourquoi elles sont un sujet d'investigation.

Une règle générale est que, pour une étendue de 10 cm dans l'eau, un proton a une chance de 10% pour produire une PG et un ion-carbone d'environ 40%. Sur les fantômes cliniques avec des plans de protons cliniques, j'observe des facteurs de production entre 3-10%. Figure A.8 montre la production de PG pour les protons primaires de 100k dans Geant4, décomposé par la multiplicité du numéro de production PG. Environ la moitié des interactions nucléaires ("fHadronInelastic" dans le langage Geant4) ne produisent pas de PG, un peu moins produit 1 PG, et il y a une queue d'interactions nucléaires produisant plus d'une PG. Les PG sont produits avec des énergies relativement élevées, jusqu'à environ 8 MeV, beaucoup plus énergétiques que les photons produits par la plupart des radionucléides utilisés pour l'imagerie PET et SPECT. Tout en enregistrant des photons à haute énergie est un défi, ils offrent un avantage significatif: les énergies supérieures ont des probabilités inférieures à absorber ou dispersées chez le patient. Un signal moins bruyant par rapport aux images PET ou SPECT peut donc être attendu. Moteabbed et al. (2011) met le rapport de transmission entre PET et PG à un facteur 5 en faveur de PG. En outre, le rapport de production est plus favorable que le PET, et avec l'acquisition par lavage et retardée, on peut s'attendre à une différence entre 60-80.

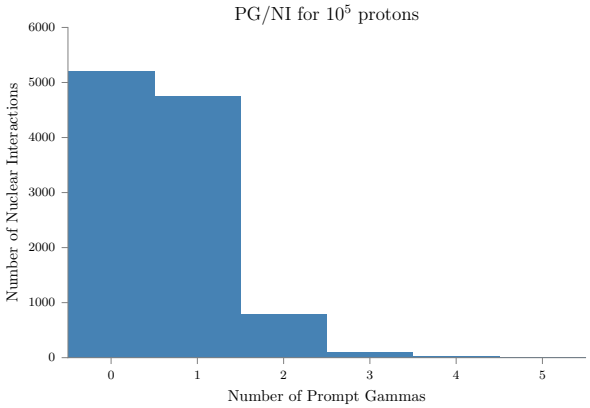


Figure A.8: Un histogramme du nombre de PG émises par réaction nucléaire pour une simulation clinique de la protonothérapie dans Gate/Geant4 où des protons de 10^5 sont lancés. Le nombre total d'interactions nucléaires produisant une PG est la somme des données et représente environ 6% pour cette image de CT clinique plus un plan.

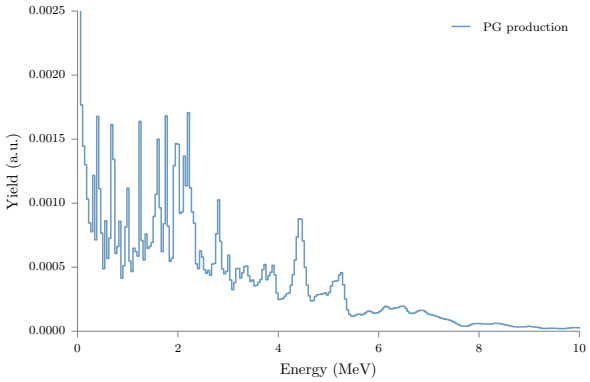


Figure A.9: Spectre d'énergie de production de PG d'une simulation de traitement clinique de protons

Une considération importante est que nous supposons actuellement que les PG sont émis spatialement par isotropie. Geant4 adhère également à cette hypothèse. Cependant, il existe des preuves du contraire (Sheldon and Van Patter, 1966; Verburg et al., 2012). En outre, les espèces de noyau cible ont une influence sur l'énergie PG. Certains matériaux ont tendance à produire certaines énergies PG. Par exemple, le pic de 4.44 MeV dans le spectre d'énergie de production PG montré dans A.9 est attribuable à ^{12}C. Par conséquent, l'enregistrement de l'énergie PG révèle quelque chose sur la matière dans laquelle est produite: 4.44 MeV PG indiquerait un tissu riche en carbone.

A.3.1 Collimation

L'incident *direction* d'une PG peut être établi en limitant le champ de vision (FoV) de la zone photosensible, voir la figure A.10. Ceci est possible avec des matériaux à haute teneur en Z, comme les scintillateurs. La différence est que maintenant une absorption complète est nécessaire, pas un matériau qui réémette l'énergie comme lumière. Les métaux comme le plomb et le tungstène sont un choix commun. Selon l'épaisseur du collimateur et la largeur du trou ou de la fente, l'angle d'incidence est limité à une plage limitée. Le fait que l'acceptation angulaire est une gamme finie et qu'une seule valeur se traduit par une fonction d'étalement du point dans le détecteur: les particules émises sous

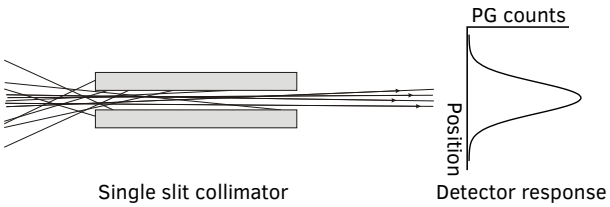


Figure A.10: En plaçant les matériaux absorbant les photons de manière appropriée, l'angle d'incidence est limité à une plage (très limitée). Le fait que l'acceptation angulaire est une gamme finie se traduit par une fonction d'étalement du point (PSF) dans le détecteur. Dans la figure, les photons viennent de la gauche et à droite le PSF est esquissé.

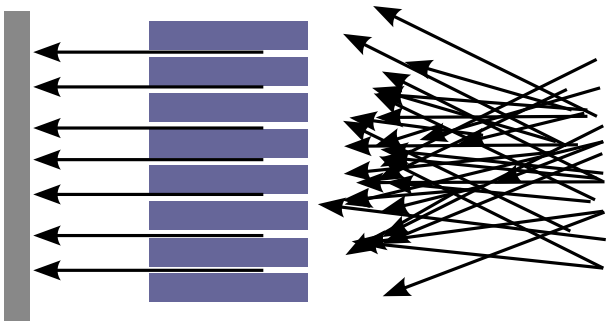


Figure A.11: Les fentes parallèles dans un matériau absorbant fournissent plusieurs bacs ou pixels avec le même angle d'incident contraint, ce qui permet d'créer une image 1D de la source. Ce type de collimètre s'appelle le colimateur Söller ou le collimateur à fente multi (parallèle). Image CC-BY-SA autorisée par Wikipedia Auteurs.

un angle «voir» une fente plus étroite et donc le collimateur transmet moins de photons comme angle d'incidence du photon Augmente la pénombre. L'effet est décrit par une fonction de propagation de points (ou PSF, voir le côté droit de A.10). Une fente plus étroite signifie une acceptation angulaire plus petite, ce qui traduit une meilleure précision sur le trajet incident (contribution plus faible du pénombre), mais un nombre plus élevé de photons rejetés (absorbés) et donc un signal inférieur. Puisque les sources de bruit sont généralement constantes, un signal inférieur signifie un rapport signal / bruit pire. Le nombre de photons transmis par photon émis peut être appelé efficacité de collimation.

Une seule fente ou trou ne donnerait que des informations sur la production d'un point ou d'une ligne. Dans la vérification PG, une des caractéristiques cliniques est la gamme des protons. Dans ce cas, selon la figure A.5, une image de la dimension le long du faisceau doit être enregistrée. Ce but peut être atteint en plaçant plusieurs fentes le long du faisceau, en se rassemblant dans un bac, ou en utilisant un collimateur relativement mince avec une seule fente et projettent la dispersion angulaire sur une surface de détection segmentée.

La géométrie de collimateur la plus directe est le collimateur à plusieurs vitesses. Chaque fente est collectée dans un bac unique, et en ayant ces fentes et bacs le long de la ligne de faisceau, on obtient un profil d'émission PG 1D. Les dimensions perpendiculaires (la longueur des fentes) peuvent être utilisées pour augmenter le signal collecté, au prix de la perte de l'information angulaire dans cette direction.

Dans la configuration de fente parallèle, chaque fente est censée fournir une projection du segment de ligne directement devant. Chaque segment de ligne de fentes est placé séquentiellement pour obtenir un profil complet. La pénombre de chaque fente est minimisée par les fentes relativement profondes, tout en augmentant la taille et le poids de ces conceptions. L'avantage est que

Case	Precision fall-off	Precision entrance	Precision profile length	FOV	Slit	Slab	Thickness	Distance axis-collimator	Distance axis-detector
1	0.59(0.06)	0.66(0.09)	0.88(0.11)	23.6	5.4	2.6	180.2	303.7	485.3
2	Same as Case 1								
3	0.70(0.08)	0.87(0.12)	1.12(0.14)	13.1	3.0	2.1	190.9	322.3	516.5

Table A.1: Tableau 6 prise de Pinto et al. (2014), contenant la géométrie du collimateur et la performance de l'optimisation réalisée avec $5{\cdot}10^8$ protons. Les unités sont en mm sauf indication contraire.

la réponse de détection est constante sur la portée du profil, et le champ de vision de la caméra est déterminé par le nombre de fentes. Une autre approche est de prendre un collimateur à fente unique et de façonner les bords des fentes. La largeur de fente effective en fonction de l'angle d'incidence est maintenue constante, jusqu'à l'angle des bords des fentes. Une image d'un segment de ligne beaucoup plus large qu'une seule fente dans une conception MPS peut alors être vue. Cela permet un collimateur plus fin, au prix d'une réponse de détection dépendant de la position: les segments de ligne plus éloignés de la fente (dans la direction transversale) sont plus éloignés du collimateur et de la cible. Parce que la fente résultante ressemble à un bord des couteaux, la conception du collimateur s'appelle la fente Knife-Edge (KES).

CLaRyS

Une conception collimée MPS est actuellement en construction à l'Institut de Physique Nucléaire de Lyon, dans le cadre du projet CLaRyS (Contrôle en Ligne de l'HAdronthérapie par RaYonnements Secondaires), une coopération entre différents laboratoires français: IPNL à Lyon, LPC à Clermont -Ferrand, LPSC à Grenoble, CREATIS à Lyon, LIRIS à Lyon, CPPM à Marseille et le Centre Antoine Lacassagne à Nice, qui héberge un faisceau de protons médicaux. Le détecteur se compose d'unités Scintillator-PMT réutilisées et rénovées provenant d'un scanner PET Phillips. Le scintillateur est un bloc strié d'oxyde de germanite de bismuth (BGO) fournissant 8 par 8 pseudo pixels. Il existe quelques choix préférés pour les matériaux de scintillation (notamment les matériaux dopés au césium tels que LaBr3:Ce et LYSO:Ce), mais BGO a été déterminé comme étant comparable aux deux, et mieux en raison de l'absence de radioactivité intrinsèque Richard et al. (2012). Chaque bloc scintillateur est lu par quatre PMT qui permettent une reconstruction de la position d'impact via le calcul du centroïde. Bien que cela ne soit pas fait à ce moment-là, les mesures de la position de la sous-fente peuvent améliorer les résultats ultérieurement. Bien que (Krimmer, 2015, Section 2.2) décrit les blocs BGO dans le contexte d'une caméra Compton, les blocs sont en fait les mêmes que ceux utilisés pour la caméra MPS. L'électronique de lecture pour les blocs sera également la même. De plus, la mesure du temps de vol (ToF) est activée par l'inclusion d'un hodoscope, lequel affiche les protons incidents. Sur la base de l'observation que le fond est en grande partie un bruit constant, alors que le signal PG arrive à l'intérieur de nanose-secondes après une impulsion de proton, environ 60% de l'arrière-plan est supprimé en définissant une fenêtre d'acceptation ToF sur la sortie du détecteur. La fenêtre d'énergie PG est réglée entre 1 et 8 MeV.

La conception du collimateur de tungstène a été déterminée en grande partie par l'étude exposée dans Pinto et al. (2014). Ici, pour la première fois, une étude d'optimisation in silico (Geant4) a été effectuée en variant de manière aléatoire les paramètres du collimateur (largeur de fente, largeur de la dalle, épaisseur de la dalle, distance à la source) et testant ses performances sur quelques indicateurs clés. Cela a entraîné trois configurations performantes pour trois indicateurs: la meilleure précision

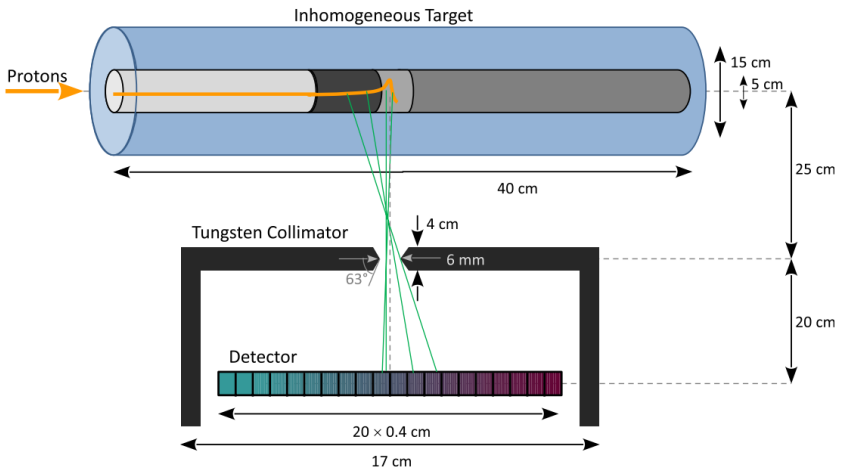


Figure A.12: La caméra IBA-couteau comme montré dans cite Priegnitz2015. Le matériau collimateur est le tungstène. Le fantôme est une coque de PMMA externe dans laquelle différents inserts («masque», air, poumon, adipeux, muscle, cerveau et os) peuvent être placés dans un ordre varié, pour étudier leur effet sur la portée et le profil PG.

sur la position de chute (cas 1), la meilleure précision sur la longueur du profil (distance à FoP, Case 2) et la meilleure précision sur la position de chute avec un FoV inférieur à 15 mm (FoV ici se réfère à la somme de la largeur de fente + deux fois le penumbra, Case 3). Les résultats de l'optimisation sont présentés dans le tableau A.1. Ces résultats ont été obtenus avec une cible homogène et un faisceau de crayons comptant $5 \cdot 10^8$ protons. Le décalage est récupéré en prenant un profil statistique élevé en tant que référence, et s'adapte aux profils par une fonction B-spline rationnelle non uniforme (NURBS) dans la région de la chute du profil.

En résumé, la géométrie du collimateur optimisée pour une meilleure précision de chute atteint une précision de 0.59 ± 0.06 mm dans une cible homogène, en utilisant des protons $5 \cdot 10^8$ avec la sélection de ToF précitée. Le faisceau de crayon 160MeV monoénergique a été centré dans un fantôme cylindrique de PMMA avec un rayon de 75 mm et une longueur de 200 mm.

IBA

Entièrement caractérisé dans Perali et al. (2014), et la performance étudiée en détail par Priegnitz et al. (2015), le prototype IBA est un design KES et la première et unique caméra au moment de l'écriture avec les résultats cliniques publiés (Richter et al., 2016). Le collimateur est un bloc de tungstène de 4 cm d'épaisseur, avec une fente de 6 mm avec des arêtes de couteau de 63 degrés (fig. A.12). Le scintillateur est LYSO (fourni par Saint Gobain) et a été sélectionné pour sa densité, son rendement léger et son faible temps de désintégration de 41 ns. LYSO souffre de la présence de l'isotope de dégradation bêta ^{176}Lu, qui a une limite supérieure de l'énergie de désintégration de 1,193 MeV. En mettant une fenêtre d'acceptation d'énergie sur la sortie du détecteur de 3 à 6 MeV, ces événements devraient être éliminés. La hauteur a été choisie à 30 mm, ce qui facilite un couplage 1: 1 à la lecture, qui est une matrice SiPM de 31,5 mm. Le LYSO et la lecture sont segmentés tous les 4 mm le long de la direction du profil du faisceau. Le rapport entre les limites de la source au collimateur et du collimateur aux détecteurs implique un facteur de mise à l'échelle de 0,8, ce qui signifie que la longueur totale du détecteur, 8 cm, correspond à une capacité à mesurer 10 cm des émissions de PG à la source. Il convient de noter que si la caméra n'est pas déplacée pendant un traitement clinique, chaque point aurait un facteur de grossissement légèrement différent, en raison des différentes positions transversales et de la position de la source à la collimation différente. Au-

cune solution pour cette implication n'est fournie. Étant donné que le déplacement de la caméra vers ou en l'absence du patient synchrone sur chaque point est une solution impraticable, la traduction de la distance transversale à la caméra au facteur de grossissement correct est une solution qui peut être considérée.

Pour récupérer la chute, Priegnitz et al. (2015) décrit une procédure qui implique de prendre les dérivées de premier et second ordre du profil obtenu:

1. Normalisation et soustraction d'arrière-plan: tous les profils sont mis à l'échelle de sorte que la moyenne des trois points de données les plus proximaux soit une et la moyenne des trois points les plus distal soit zéro.
2. Définition de bord distal: elle est définie comme la partie la plus distale du profil, où les points de données sont supérieurs à un certain seuil (ici, nous avons choisi 0,1 pour supprimer les fluctuations de fond statistique), le profil diminue (première dérivée <0) , Et seul un point d'inflexion se produit (la dérivée secondaire est zéro une seule fois). Ce qui n'est pas décrit, c'est ce qui se passe avec les données qui ne passent pas ce critère.
3. Normalisation pour l'estimation de portée: les bords distaux définis sont normalisés à 1 et 0.
4. Estimation de l'écart de plage avec l'algorithme de correspondance de pente distale, où la correspondance est effectuée par des profils obtenus avec une mesure de simulation statistique élevée.

Ensuite, on étudie les effets sur la gamme et le profil PG de remplacement de divers inserts dans le déplacement fantôme et d'introduction en raison d'une modification de l'énergie du faisceau. Pour chaque «configuration de changement» (qu'il s'agisse d'un matériau remplacé ou d'une énergie de faisceau modifiée), une mesure physique avec des protons de $2.6 \cdot 10^{11}$ est effectuée au Centre de thérapie de protons, à Prague, en République tchèque, et à des estimations in silico au 10^7, 10^8 et 10^9 sont effectués. Les estimations in silico sont fournies par la méthode de réduction de la variance analytique pour l'estimation PG présentée dans Sterpin et al. (2015). Comme décrit là-bas, au moment où l'algorithme ne traitait pas d'inhomogénéités latérales, la raison probable pour l'étude d'inclure un fantôme 1D. Chaque estimation in silico PG est exécutée à des milliers de fois et la chute obtenue avec la procédure décrite ci-dessus histogramée. Les écarts moyens et standard de ces histogrammes fournissent la précision de l'estimation de la chute. Ensuite, en fonction de la vérité au sol mesurée dans l'expérience physique, des courbes de caractéristique d'exploitation du récepteur (ROC) sont créées et fournissent une estimation de l'opportunité ou non du décalage en fonction du nombre de primaires lancées, et pour chaque changement dans Matériau ou changement d'énergie.

La meilleure précision de chute (écart type sur les mesures FOP) atteint environ 2 mm pour les cibles homogènes dans le champ de vision de la caméra, avec des primaires de 10^9. On peut s'attendre à ce que les deux conditions soient rares dans les conditions cliniques. Peut-être plus intéressant est l'implication de (Priegnitz et al., 2015, fig 8): les changements, les énergies du faisceau et les inhomogénéités interfèrent de manière intéressante, ce qui pourrait être exploité pour apprendre plus que seulement sur les positions de chute.

Certaines observations du même groupe publié dans Smeets et al. (2012) sont intéressantes:

- Les profils PG après une sélection d'énergie de 4 à 5 MeV ont le meilleur rapport signal sur fond,
- Les profils PG ne se rapportent pas à la dose déposée avec un seuil d'énergie supérieur de 1 MeV, ce qui complète les observations concernant l'activité intrinsèque de LYSO,
- Les énergies de détection inférieures montrent un pic de Bragg moins pointu, complétant les instructions concernant les PG dispersées dans Min et al. (2012).

A.4 Accelerated Prompt Gamma estimation pour les simulations cliniques de Proton Therapy

Dans ce section, l'effort sur la simulation accélérée des estimations PG est présenté. Les difficultés expérimentales avec des rendements faibles discutés dans le section précédent posent également un problème pour les simulations: la convergence d'une estimation acceptable d'un processus rare nécessite beaucoup de calcul et donc de longs temps de mur, même à l'aide de l'informatique en grappe. Les détails sur ces difficultés sont également discutés dans le document, qui a été publié en 2016 dans *Physics in Medicine and Biology* (Huisman et al., 2016).

A.4.1 Introduction

La gamme bien définie de particules dans la matière est la principale raison pour laquelle elles sont utilisées dans le traitement du cancer aujourd'hui. Malheureusement, nous ne sommes pas en mesure de profiter pleinement de cette propriété, en raison des incertitudes dans le positionnement du patient, des incertitudes sur la gamme de protons en raison de déplacements inconnus ou de déformations d'organes, de remplissage malin pulmonaire, intestinal ou veineux et l'incertitude inhérente à la Houndsfield unité à la conversion de puissance d'arrêt de particules. Souvent, la pratique médicale consiste à planifier de façon conservatrice, à savoir ajouter des marges autour de la tumeur, réduisant considérablement les avantages potentiels du traitement des particules (Knopf and Lomax, 2013). Une certaine forme de surveillance en ligne in vivo est généralement considérée comme une issue à cette situation. La surveillance en ligne permettrait de mesurer les incertitudes telles que mentionnées ci-dessus et permettrait ainsi une planification plus précise qui pourrait tirer le meilleur parti possible de la dégradation escarpée du pic de Bragg (BP) et réduire les dommages aux tissus entourant la tumeur.

Les paradigmes d'imagerie tels que la détection PG sont validés contre des expériences, et souvent aussi avec des simulations Monte Carlo (MC) (Golnik et al., 2014; Gueth et al., 2013; Janssen et al., 2014; Moteabbed et al., 2011; Robert et al., 2013). Les méthodes MC conventionnelles propagent les particules selon un ensemble de processus physiques à travers les matériaux. Le suivi d'une particule est divisé en étapes, où à chaque point, une liste pondérée de toutes les prochaines étapes possibles est construite et une option est sélectionnée par un nombre aléatoire. Pour les processus qui se produisent rarement, la convergence vers le modèle de la vérité à une erreur statistique acceptable peut être lente. Les émissions de PG dans la thérapie par particules, lorsqu'elles sont vues sur une base voxel-by-voxel, sont également un processus rare et lentement convergent (Gueth et al., 2013; Pinto et al., 2015; Sterpin et al., 2015). Cela a des implications importantes: d'abord pour les concepteurs de détecteurs, d'autre part pour ceux qui simulent et, troisièmement, pour ceux qui s'intéressent à la comparaison des deux conditions cliniques, par exemple cliniques. Dans le premier cas, les détecteurs sont optimisés pour minimiser la perte de signal (voir fig. A.13) et une reconstruction avancée peut être utilisée pour maximiser l'utilisation de l'information dans le signal. Gueth et al. (2013) a démontré une méthode qui fonctionne autour du faible rendement de PG en utilisant l'apprentissage par machine pour corréler les traductions prédéfinies des patients (erreurs de configuration) aux signaux de sortie PG, ce qui réduit le temps de production d'une traduction estimée basée sur les PG détectées. Puisque la convergence vers le modèle de la vérité nécessite des temps d'exécution de simulation longs, nous pouvons compenser les méthodes de réduction de variance ou les seuils dans les domaines temporel, spatial ou spectral (par exemple, temps d'exécution fixe, taille voxel plus grande, bacs spectraux plus grands). Une telle méthode de réduction de variance est le fractionnement MC, où le moment où un processus rare ne se produit pas un, mais plusieurs futurs possibles pour ce processus sont calculés. Le total pondéré de ces contrats à terme est ensuite stocké, de sorte que la convergence s'est accélérée.

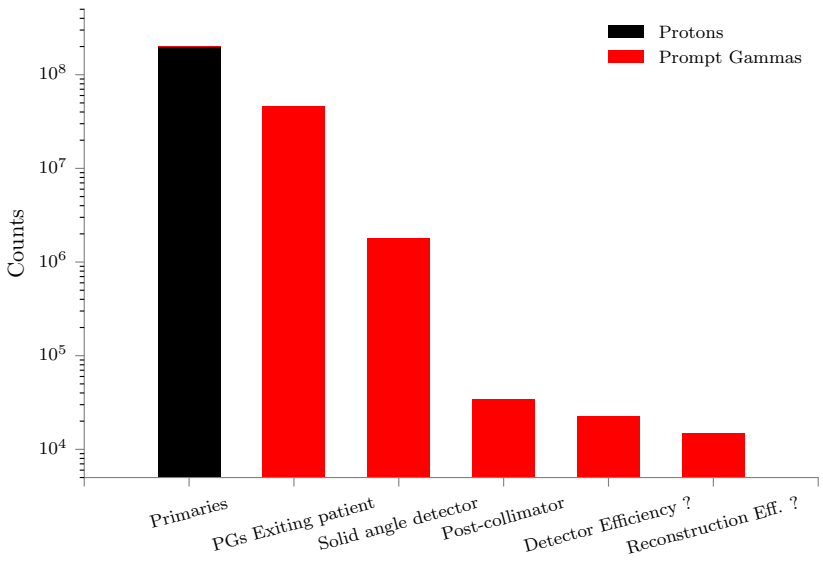


Figure A.13: Perte progressive approximative du signal PG par point. Les particules ont été enregistrées dans une simulation avec une CT clinique de tête et de cou et un plan clinique associé, avec un collimateur de 30 par 30 cm avec un facteur de remplissage de 43%, à 40 cm du patient, avec une efficacité estimée (66%) Pour la détection et la reconstruction. Notez que ce n'est pas la configuration étudiée dans le reste de ce document.

A.4.2 Materiels et Methodes

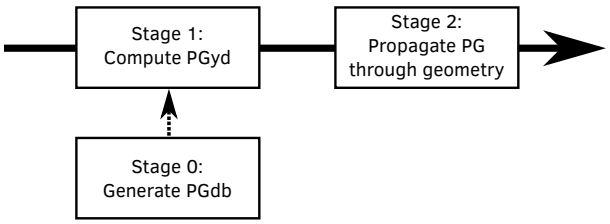


Figure A.14: Flow chart de la méthode vpgTLE. L'étape 0 est une base de données PG initiale (PGdb) et est calculée une fois. Chaque simulation subséquente est divisée en étape 1 et étape 2. L'étape 1 génère une distribution de rendement PG spécifique au CT (PGyd) en utilisant un nombre limité de primaires. Dans l'étape 2, PGyd est utilisé pour générer et propager un ensemble représentatif de PG dans toute la géométrie.

Une simulation complète *voxelized Prompt-Gamma Track Length Estimator* (vpgTLE) est divisée en deux étapes (fig.A.14). Le processus suppose l'existence d'une base de données préparée (PGdb) qui est une estimation du coefficient de production PG linéaire efficace modulo la densité, par élément ($\frac{\Gamma_Z}{\rho_Z}$), par bac d'énergie PG par bac à énergie primaire. Le PGdb est calculé une fois pour une liste d'éléments communs, puis peut être réutilisé. Dans la première étape, une image de distribution de rendement PG (PGyd) est créée, spécifique à un fantôme particulier (ou image CT) et à une source primaire (par exemple, un plan de traitement, un point unique). Cette image PGyd stocke, par voxel par bac d'énergie PG, le rendement par primaire. La deuxième étape utilise le PGyd et l'hypothèse d'une émittance PG isotrope pour générer et propager les PG dans le reste de la géométrie définie par l'utilisateur (par exemple, le détecteur PG).

Chaque simulation est exécutée à la fois avec la méthode analogique Monte Carlo et avec la méth-

ode vpgTLE. Le MC analogique sert de référence. Pour obtenir une estimation de l'incertitude statistique, nous employons la technique de lot et exécutons chaque type de simulation 10 fois. Lors de l'étude du biais et de l'incertitude relative dans les sous-régions sélectionnées des fantômes, σ est calculé sur la projection considérée. C'est-à-dire, σ représente l'écart type sur le rendement moyen sur les 10 simulations. Pour l'étude sur l'efficacité et la convergence de l'incertitude relative, la variance est calculée par voxel. En prenant la médiane de (une sous-région de), cette 4D «image de variance» (c'est-à-dire la médiane de la variance) fournit un test plus fort. Nous prenons la médiane plutôt que la moyenne parce que la distribution de la variance tend à une distribution log-normale. Pour les distributions fausses comme celle-ci, la médiane est une meilleure mesure de la tendance centrale.

Deux cas de test sont présentés. Le but de Test case 1 est de vérifier que la transition vers voxels a été effectuée correctement. Le fantôme proposé par Parodi et al. (2005) et utilisé par El Kanawati et al. (2015) est converti en une représentation voxélisée, voir figure **??**. Dans le cas d'essai 2 (fig. **??**), une image clinique de la tête et du cou avec un plan de traitement de proton approprié est examinée et est destinée à démontrer les possibilités de la méthode vpgTLE. Par rapport à El Kanawati et al. (2015), le nombre d'primaires analogiques utilisées pour la référence est augmenté de 10^7 à 10^9. Ceci est nécessaire pour obtenir une figure suffisamment silencieuse pour la projection spatiale le long de l'axe du faisceau. Pour les simulations vpgTLE, quatre simulations sont exécutées avec des primaires de 10^3, 10^4, 10^5 et 10^6 respectivement. Ensuite, nous définissons certaines fenêtres d'intérêt. Sachant que la plupart des détecteurs de PG ne mesurent pas en dehors de la gamme d'énergie de 1 à 8 MeV (Testa et al., 2013), ou même plus étroits (Smeets et al., 2012), nous limitons notre analyse à cette fenêtre d'énergie. En outre, le rendement de PG en dehors de la région spatiale qui représente 90% du rendement total dans la simulation de référence est éliminé. Pour toutes les analyses, ces deux coupures sont appliquées.

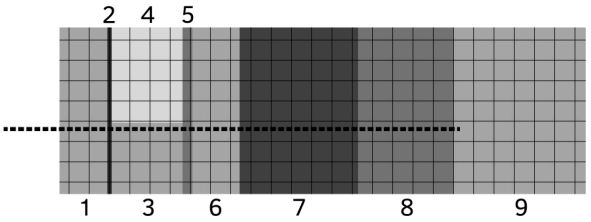


Figure A.15: Vue descendante du Parodi phantom (Parodi et al., 2005), où l'ombrage représente la densité du matériau. Les parties 1, 3, 6, 9 sont PE; 7 est PMMA; 2 est Os; 5, 8 sont Muscle; 4 est le poumon. Le faisceau est illustré par la ligne pointillée venant de gauche. Une version voxélisée de cette image est créée, avec une taille de 2^3 mm ^{3} voxel.

A.4.3 Résultats

Test case 1

D'abord, nous avons vérifié que les rendements de vpgTLE sont identiques aux résultats produits avec pgTLE, présentés dans El Kanawati et al. (2015). Ensuite, nous avons comparé notre méthode à la référence analogique. La figure A.17 représente le rendement de la première rangée, en fonction de la profondeur (colonne gauche) et de l'énergie PG (colonne droite). La différence relative du rendement PG est montrée, toutes deux intégrées sur l'ensemble du plan coronal (deuxième rangée) ou sur la ligne voxel sur le chemin du faisceau (troisième rangée). La première ligne, un tracé du rendement, montre un chevauchement parfait de vpgTLE par rapport à la référence. Nous devons examiner les

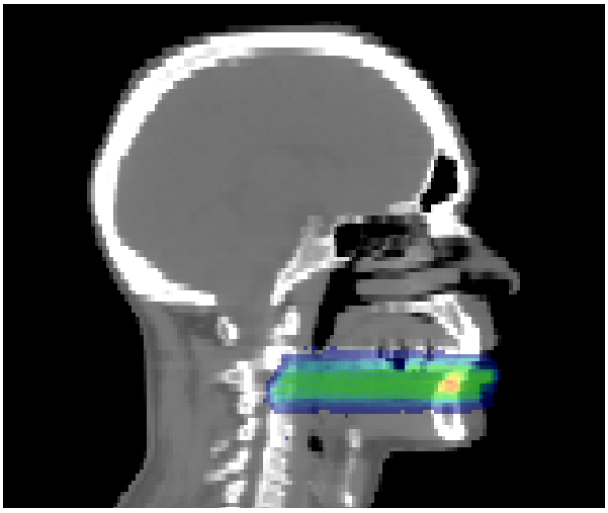


Figure A.16: Vue sagitale de l'image CT du patient, illustrée du rendement PG généré par le plan de traitement associé. Un faisceau d'un plan de traitement original a été tourné pour s'aligner avec les axes d'image, afin de rendre les projections plus faciles et d'augmenter l'hétérogénéité des matériaux, ce qui augmente le défi de vpgTLE. Le rayonnement appliqué est la couche distale de l'un des faisceaux du plan d'origine, de sorte que le décalage distal est mieux défini en contraste avec le sommet de Bragg émergé (SOBP). La taille du voxel est de nouveau 2^3 mm ^{3}.

différences relatives des diverses sorties vpgTLE par rapport à la référence dans la rangée 2 pour respecter les différences. Les zones ombrées représentent 2 σ bandes d'erreur.

Avec des particules primaires de 10^3, la moyenne est plus bruyante, comme prévu. Une surestimation au-delà de 170 mm est visible, ce qui concerne l'emplacement du pic de Bragg. La différence relative moyenne par rapport à la profondeur est de 4,0 $\times 10^{-4}$ le long du faisceau, ce qui est une bonne performance, mais en raison de la différence relative avec la référence supérieure à 1 % dans la région distale et les bandes d'erreur très larges , Nous discuterions 10^3 les primaires sont insuffisantes pour une prédiction fiable. Le décalage systématique distal a diminué en utilisant des primaires de 10 4 ou plus. Deux régions avec un biais restent: une surestimation constante d'environ 0,5 % à environ 160 mm de profondeur, puis, après le pic de Bragg, un moyen erratique avec de larges bandes d'incertitude. Ce dernier peut s'expliquer par des événements nucléaires. Une fois qu'un proton entre en collision et est absorbé, il ne peut plus produire de PG. Vers la fin, le nombre précis de protons restants devient plus incertain, et seulement 10 3 des primaires ne suffisent pas pour une bonne estimation de la variance. L'augmentation du nombre de primaires réduit l'incertitude et améliore le rendement moyen, mais l'effet demeure. La différence relative moyenne sur la profondeur est de l'ordre de 10^{-4} pour tous les ensembles primaires.

La colonne spectrale à droite démontre que vpgTLE est proche de la référence analogique sur l'ensemble du spectre, avec une petite fluctuation au haut de gamme du spectre. Le motif présent pour tous les ensembles primaires doit être dû au PGdb, et c'est le seul biais que nous observons. La différence relative moyenne varie entre -1.3×10^{-4} avec 10^3 primaires et 6.0×10^{-5} pour 10^6 primaires. Ceci soutient l'hypothèse selon laquelle nous avons convergé vers le biais introduit par l'incertitude du PGdb. Les bandes d'erreur étendues pour 10^3 sont à nouveau visibles, les bandes d'erreur pour des primaires de 10^4 ou plus restant dans 1% de la moyenne sur toute la gamme.

La ligne inférieure, la ligne de voxels centrée sur le chemin du faisceau, montre un comportement plus erratique. Une différence majeure est la surestimation proximale et la sous-estimation distale avec des primaires de 10^3. Avec 10 5 de primaires ou moins, la différence relative moyenne est de l'ordre de 10^{-3} ou plus, tandis que 10^6 primaires aboutit à 3.6×10^{-4}. L'incertitude est, naturellement,

plus grande. La vue spectrale est stable et la vue en profondeur a une variance accrue vers la fin de la gamme de protons.

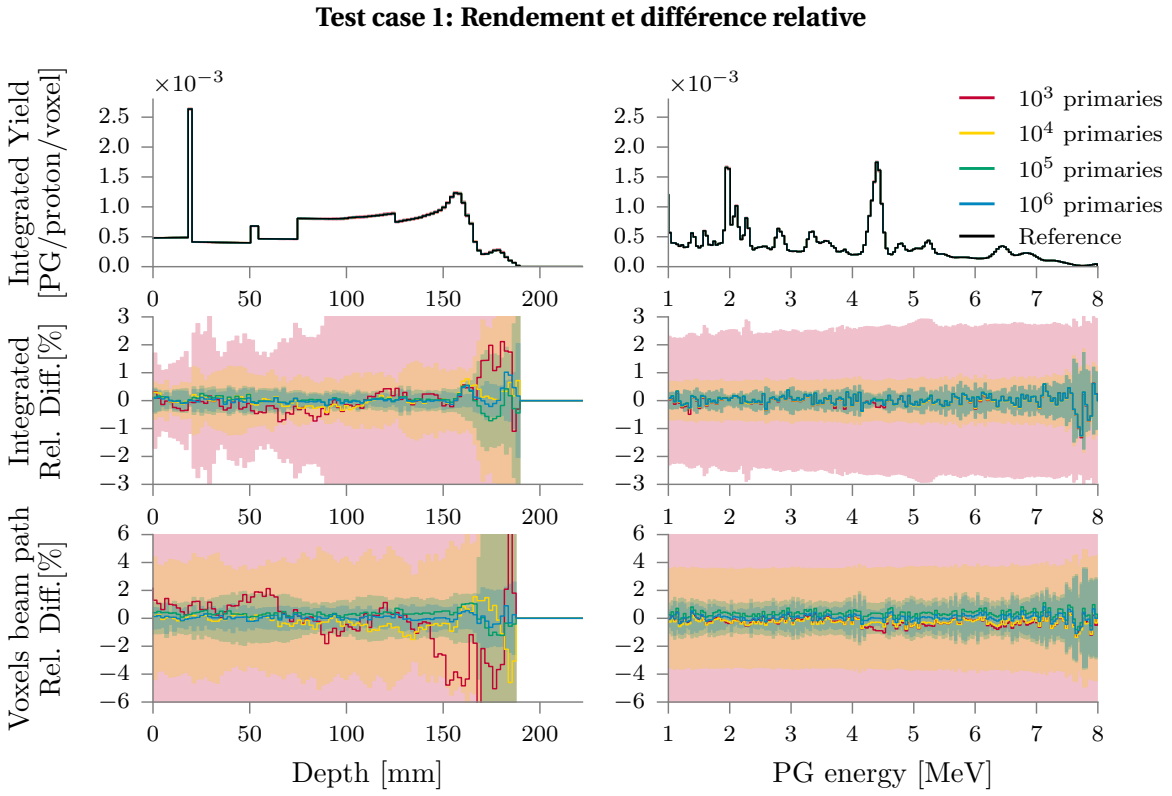


Figure A.17: Test case 1: la ligne 1 montre les rendements de PG et la rangée 2 la différence relative par rapport à la référence ($\frac{vpgTLE-Reference}{Reference}$). Le rendement correspond à la moyenne sur 10 simulations. Pour les deux rangées, notez que le rendement a été intégré sur toutes les autres dimensions. La ligne 3 montre la différence relative sur la ligne des voxels au centre du chemin du faisceau, où nous ne nous sommes pas intégrés sur toutes les autres dimensions. La colonne de gauche est une projection le long de l'axe du faisceau, tandis que la colonne de droite montre les bacs spectraux intégrés sur tous les voxels considérés. Les zones ombrées représentent 2 σ bandes d'erreur, où σ est l'écart type par rapport à la moyenne de 10 simulations. Notez que le motif de covariance dans la différence relative est dû au bruit du signal analogique et ne représente aucun problème avec l'implémentation vpgTLE.

La figure A.18 montre côte à côte la convergence de l'incertitude relative médiane et un histogramme des gains. Nous voyons qu'une convergence médiane à 2% est atteinte en environ 3 minutes et environ 68 heures avec vpgTLE et analogique MC respectivement. Au niveau de convergence 2 %, le gain est de 1,55 $\times 10^3$. Les histogrammes à gauche montrent que les gains sont stables par rapport au nombre de primaires vpgTLE. Cela signifie que vpgTLE n'a pas de problèmes systématiques. Nous pouvons clairement voir l'inclinaison des distributions (notez l'échelle logarithmique sur l'axe x). Le pire gain est un facteur de 6,19 $\times 10^1$, tandis que le meilleur voxel s'allume à 5,21 $\times 10^4$, avec une médiane de 1,40 $\times 10^3$.

Test case 2

Regardez pour calculer **??** pour le rendement et la différence relative de vpgTLE par rapport à la référence. La largeur des bandes 2σ a augmenté par rapport au cas de test 1. Sur le faisceau, colonne

Test case 1: Gain distribution et Convergence

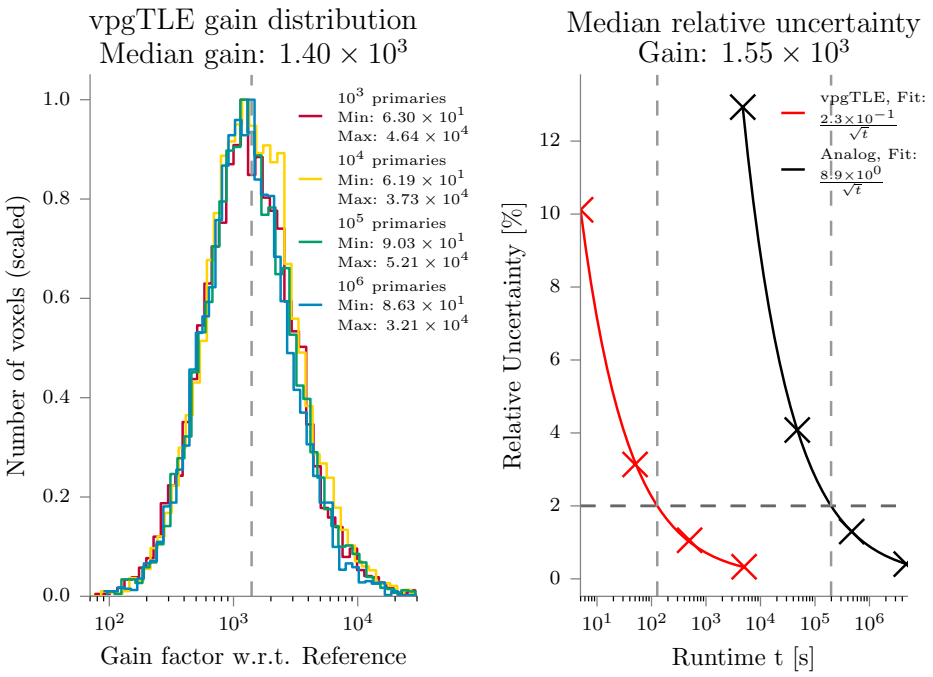


Figure A.18: Cas d'essai 1. À gauche, l'efficacité est calculée, par voxel, pour tous les ensembles primaires vpgTLE par rapport à la référence. À droite, l'incertitude relative médiane en fonction du temps d'exécution t, tant pour les méthodes analogiques que pour les méthodes VpgTLE. Chaque point successif est généré avec des primaires de 10^3 - 10^6 pour vpgTLE, et avec 10^6 - 10^9 primaires pour MC analogique. Comme t augmente, l'incertitude relative diminue comme $\frac{C}{\sqrt{t}}$, où C est un facteur d'ajustement. Pour calculer le gain, nous prenons le rapport des temps de fonctionnement au niveau 2 %, indiqué par la ligne horizontale pointillée, généralement considérée comme suffisamment convergente.

gauche de la figure, nous voyons que 10^3 primaires produisent une ligne erratique, tandis que 10^5 et up sont proches de 0 %. Passé l'extrémité distale, nous voyons une divergence significative comme dans le cas de test 1. Alors que le biais moyen est d'une grandeur similaire à celui du cas de test 1 (10^{-4}, à l'exception de 10^3 primaires), à l'extrémité distale, le biais n'a pas été très dissipé avec des primaires de 10^6. Une explication probable pour la pire performance est le faisceau différent par rapport au cas de test 1: maintenant nous regardons le rendement causé par une couche d'énergie complète composée de 7 points, ce qui entraîne l'étalement des primaires sur un volume plus important et résulte donc Des statistiques inférieures par voxel. Nous pouvons encore attribuer une partie de l'augmentation de l'erreur à l'erreur systématique induite par PGdb. La vue spectrale est aussi stable que dans le cas de test 1, qui est réparti sur les mêmes 250 bacs, et diffère uniquement par des bandes d'erreur plus larges.

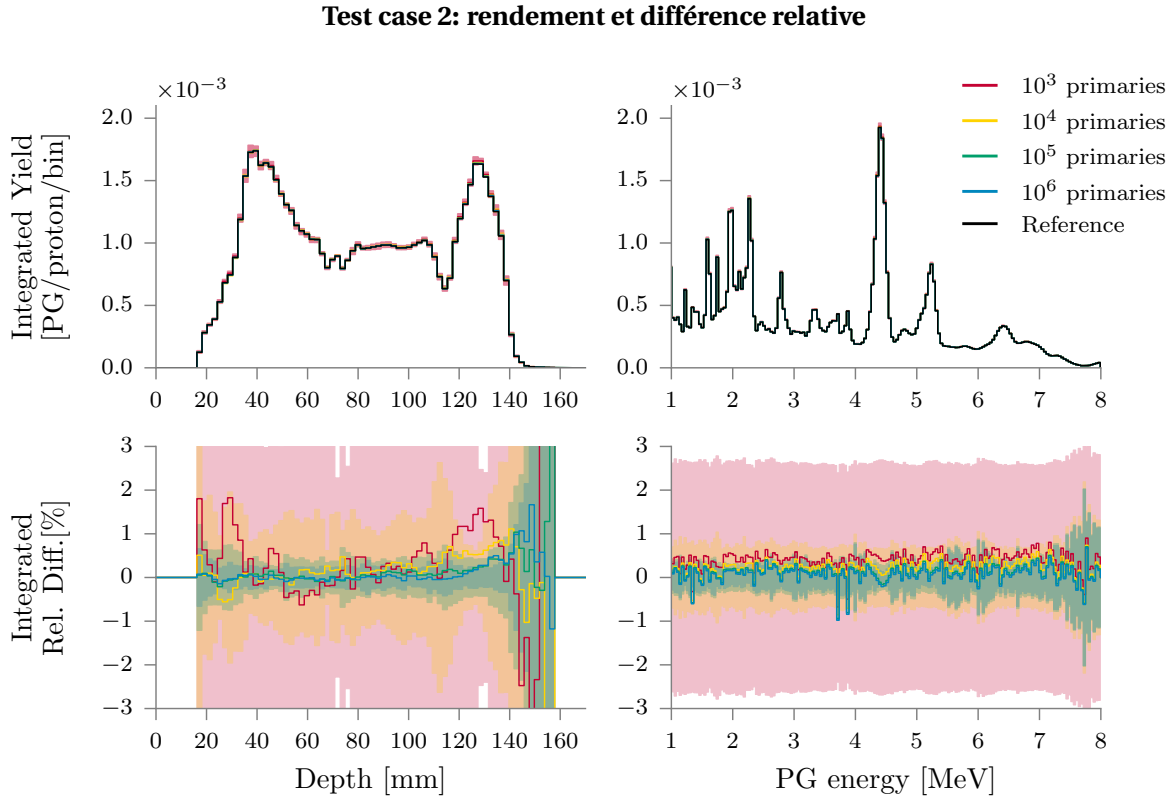


Figure A.19: Test case 2. La ligne 1 montre les rendements de PG et la rangée 2 la différence relative par rapport à la référence ($\frac{vpgTLE-Reference}{Référence}$). Pour les deux rangées, notez que le rendement a été intégré sur toutes les autres dimensions. La colonne de gauche est une projection le long de l'axe du faisceau, tandis que la colonne de droite montre les bacs spectraux intégrés sur tous les voxels considérés. Les zones ombrées représentent 2 σ bandes d'erreur, où σ est l'écart type par rapport à la moyenne de 10 simulations.

La figure A.20 montre la convergence et le gain de vpgTLE pour Test case 2. Le gain est légèrement inférieur à celui du cas de test 1. Une PGyd suffisamment convertie nécessite peu plus de 4 heures sur un seul noyau avec vpgTLE et environ 180 jours avec MC analogique. En excluant le montant de 10 3 en raison de ses valeurs aberrantes, le pire gain est de 2,70 $\times 10^1$ et le meilleur est de 8,96 $\times 10^4$, avec un gain médian de 9,98 $\times 10^2$ et un gain De 1,03 $\times 10^3$ à 2 % d'incertitude relative.

Test case 2: Gain distribution et Convergence

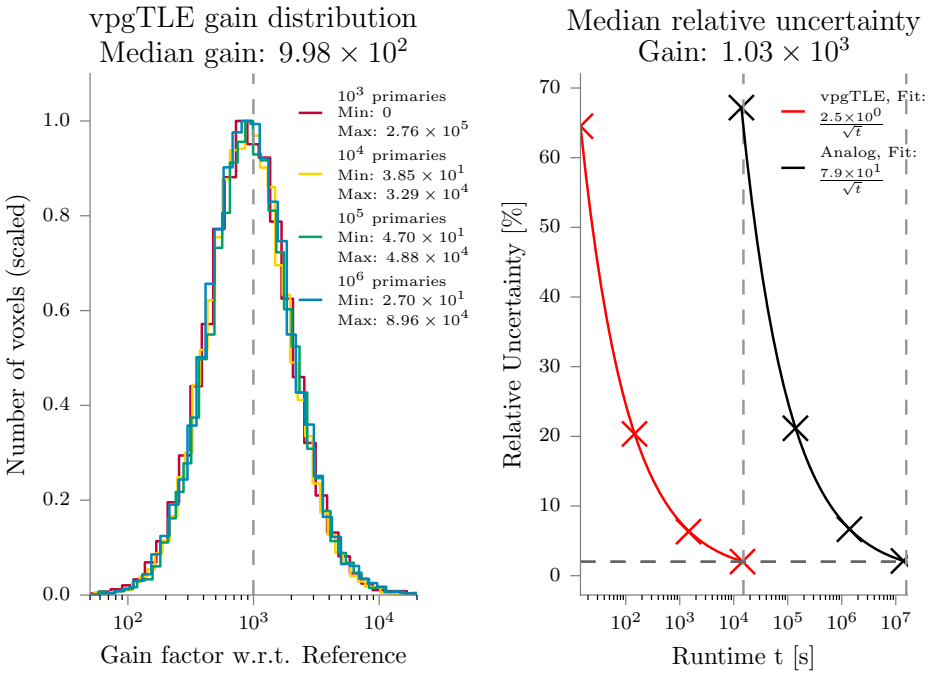


Figure A.20: Test case 2. À gauche, l'histogramme de gain est affiché, pour tous les ensembles primaires de vpgTLE par rapport à la référence. À droit, l'incertitude relative moyenne est tracée en fonction du temps d'exécution, tant pour les méthodes analogiques que pour les méthodes VpgTLE. Chaque point successif est généré avec des primaires de 10^3 - 10^6 pour vpgTLE, et avec 10^6 - 10^9 primaires pour MC analogique. Nous prenons le ratio des temps d'exécution au niveau 2 % pour obtenir le gain.

A.4.4 Conclusions

VpgTLE est une option alternative générique pour calculer la sortie PG attendue dans les géométries voxélisées. La méthode a une exigence de mémoire fixe (une image 4D) avec une taille de mémoire typique de l'ordre de quelques centaines de Mo. La méthode atteint un facteur de gain global de 10 3 pour une image CT clinique et un plan de traitement par rapport à la MC analogique. Une convergence médiane de 2 % pour la couche d'énergie la plus distale est atteinte en environ quatre heures sur un seul noyau, auquel cas la sortie s'est stabilisée à moins de 10^{-4} d'une simulation de référence analogique, lorsque la portée ou Le spectre est considéré. Les auteurs pensent que la méthode est intéressante pour ceux qui développent et simulent des dispositifs de détection PG, ainsi que des cliniciens qui étudient des cas cliniques complexes nécessitant la précision et la précision des simulations de niveau MC non offertes par des algorithmes analytiques.

La méthode vpgTLE est open source et entièrement intégrée dans Gate. Il est disponible à partir de la version 7.2.

A.5 Performance of Prompt Gamma fall-off detection in clinical simulations

Dans ce section, la première simulation clinique de bout en bout de deux caméras PG collimées optimisées sera présentée. En outre, les méthodes de regroupement des points basées sur la notion d'iso-profondeur et d'iso-énergie sont comparées. Une petite étude sur les poids cliniques des taches a été effectuée. Enfin, une figure de mérite est présentée qui fournit une estimation de la probabilité d'une FOP mesurée tombant à l'intérieur ou à l'extérieur de l'écart type prévu de $\pm 2\times$.

A.5.1 Introduction

Ici, une étude *in silico* est présentée sur la pertinence clinique de l'estimation de la position de chute (FOP), basée sur deux modèles de caméra PG collimatés: une conception multi-parallèle à fente (MPS) et à lame (KES) . Premièrement, on analyse la composition des plans de traitement clinique (TP). Le principal déterminant pour le signal PG disponible est le nombre de protons primaires par quantité d'intérêt (traitement complet, couche iso-énergétique, spot ou autre). Selon le type de cible, l'emplacement et la taille, la répartition par points varie. C'est-à-dire: le nombre de taches, leur nombre moyen de protons et leur distribution d'énergie ont des conséquences pour la détection PG. Ensuite, une brève analyse des données et de certains effets sur la planification du traitement sont présentés dans le contexte de ces distributions ponctuelles. Dans la deuxième partie de cet article, les performances des deux caméras PG sont étudiées sur un cas clinique, entièrement mises en œuvre dans une simulation Monte-Carlo réalisée avec Gate / Geant4. Pour un patient atteint d'une tumeur crânienne, une perte de poids significative a été observée et une CT de replantation (RPCT) a été obtenue. Les CT de planification et de replantation fournissent un cas «facile» où la caméra PG devrait détecter de manière fiable la différence. Le nombre minimum de protons requis pour une récupération FOP fiable est établi. Pour conclure, deux méthodes de groupage de points qui améliorent la fiabilité de la récupération en conditions cliniques sont proposées et comparées.

A.5.2 Materiels et Methodes

Simulation

Les paradigmes d'imagerie tels que la détection PG sont évalués à partir d'expériences, et souvent aussi avec des simulations Monte Carlo (MC) (Golnik et al., 2014; Gueth et al., 2013; Janssen et al., 2014; Moteabbed et al., 2011; Robert et al., 2013). Pour les processus qui se produisent rarement tels

que la simulation PG, la convergence vers le modèle de la vérité à une erreur statistique acceptable peut être lente. Cet article présente une étude *in silico* de la faisabilité de la pertinence clinique de l'estimation PG FOP à l'aide de caméras collimatées et utilise la méthode de réduction de la variance vpgTLE décrite dans Huisman et al. (2016). VpgTLE est un processus en deux étapes, où premièrement une image de distribution de rendement PG est estimée, qui dans la deuxième étape est utilisée comme source PG avec laquelle des détecteurs peuvent être étudiés. Gate 7.2 (Sarrut et al., 2014) avec Geant 4.10.02 et la liste de physique QGSP_BIC_HP_EMY, couramment utilisée pour les études PG, sont utilisées dans cette analyse. Grâce à vpgTLE, des simulations pour environ 10^9 protons (environ 6×10^8 photons) ont pris 1-2 heures sur un Intel (R) Core (TM) i7-3740QM.

PG camera modeling

Deux détecteurs PG adaptés à la vérification FOP (illustrés dans la figure A.21) ont été choisis:

- la caméra CLARyS multi-parallèle-fente (MPS), Case 1 (Pinto et al., 2014)
  Cette caméra a l'intention de mesurer le profil PG complet pour contrôler les gammes d'ions dans le patient avec un champ de vision (FoV) de 300 mm. Il utilise la sélection de ToF pour réduire le fond de neutrons. Dans l'optimisation effectuée par Pinto et al. (2014), des paramètres tels que le pitch du collimateur, l'axe au collimateur et l'axe au détecteur ont été variés et leurs impacts ont été évalués en termes de précision de récupération (FRP) et de résolution spatiale (Netteté de la région tombante). Ici, la configuration 1 (avec des contraintes détendues sur la résolution spatiale) a été choisie pour ses performances optimales en FRP. Comme cela a été fait par Pinto et al. (2014), les longueurs de la caméra (volume du collimateur et du scintillateur) sont choisies *jusqu'à* 300 mm, de sorte que la longueur est un multiple entier de la taille du pas, avec pour le collimateur un collimateur- Largeur de feuille supplémentaire, pour s'assurer que chaque pixel comporte une feuille des deux côtés. Avec le pitch de 8 mm et les colimateurs de 2,6 mm, il en résulte un volume de scintillateur de longueur 296 mm et une longueur de collimateur de 298,6 mm.
- la caméra IBA (KES) camera (Perali et al., 2014; Sterpin et al., 2015)
  Le but de cette caméra consiste à vérifier la position BP avec un FoV de 100 mm. Richter et al. (2016) fournit les premiers résultats cliniquement obtenus. À l'heure actuelle, aucune autre caméra n'a été soumise à des tests cliniques, c'est pourquoi nous considérons ce prototype comme une référence.

En ce qui concerne la sélection ToF d'arrière-plan, pour l'accélérateur IBA C230 avec une période de 10 ns, Pinto et al. (2014) a choisi une fenêtre de 4 ns autour du maximum PG, en fonction des spectres ToF expérimentaux. Cela signifie qu'environ 60 % du bruit pourrait être supprimé. Pour le prototype KES ToF n'est pas utilisé, ce qui donne lieu à un fond plus élevé, comme il est évident lorsque l'on compare les arrière-plans publiés dans les deux publications. Une deuxième différence est la fenêtre de sélection d'énergie. Le groupe IBA emploie une fenêtre de 3 à 6 MeV, alors que la collaboration CLaRyS a produit son optimisation avec une fenêtre de 1 à 8 MeV. Nous allons comparer chaque caméra avec ses propriétés publiées, c'est-à-dire: une fenêtre de 1 à 8 MeV et un ToF pour MPS et une fenêtre de 3 à 6 MeV sans ToF pour KES.

Les deux prototypes d'appareils PG ont des photodétecteurs différents et différents appareils électroniques détecteurs. Dans cette étude, ces différences ne sont pas mises en œuvre. Au lieu de cela, la méthode décrite dans Gueth et al. (2013) a été utilisée pour obtenir le point d'interaction d'un photon incident. Si l'énergie intégrée déposée dans un cristal se trouve dans l'énergie acceptable et la fenêtre ToF, l'événement est enregistré. La position de l'événement dans le cristal est considérée comme le barycentre pesé en énergie de toutes les interactions dans le cristal, plus une valeur aléatoire prise à partir d'un Gaussian FWHM de 5 mm pour simuler l'électronique et la résolution du détecteur.

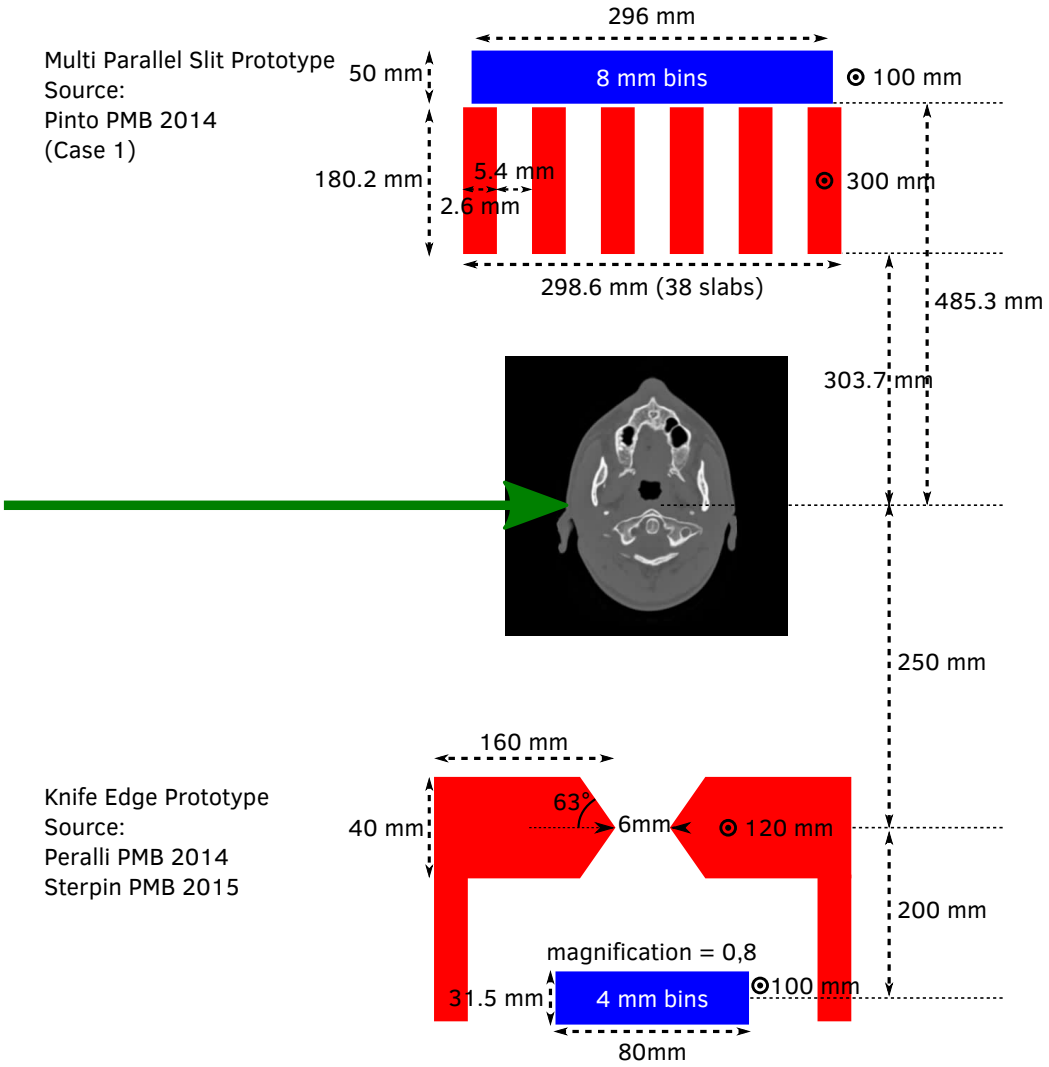


Figure A.21: Présentation schématique des deux caméras PG considérées dans cette étude. La flèche verte représente le faisceau de protons. En rouge, les éléments de collimation et les éléments de détection bleus. Les dimensions ont été prises de Pinto et al. (2014) et Perali et al. (2014); Sterpin et al. (2015). Notez que les deux caméras ont été placées à un emplacement identique au-dessus de la tête pendant toutes les simulations, et que, ici, elles ne sont pas dessiné à l'échelle.

Spot groupage

Dans la littérature d'imagerie par thérapie par particules et par thérapie par des particules, on mentionne souvent que le sommet de Bragg étalé (SOBP) est obtenu en donnant à la couche d'iso-énergie la plus distale le poids le plus élevé et chaque couche d'iso-énergie successive est d'une énergie inférieure et de poids inférieur . Ceci est correct lorsque des positions dans le plan transversal (à la direction du faisceau) sont considérées, pour des fantômes homogènes, et lorsque l'interpolation n'est pas nécessaire. L'interpolation se produit lorsque les niveaux d'énergie du système de délivrance du faisceau ne correspondent pas au contour de la tumeur distale (Clasie et al., 2012), auquel cas le contour de la dose est approximatif en répartissant la dose distale des deux plus proches Énergies possibles (voir fig. **??**).

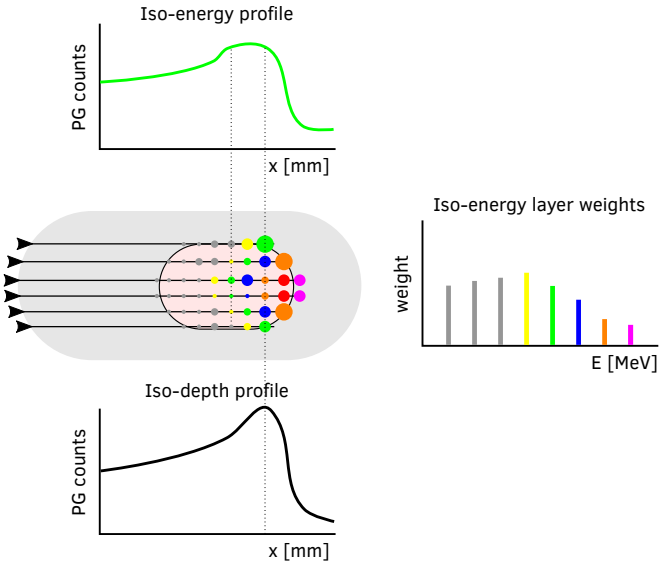


Figure A.22: Une vue schématique d'un patient (gris clair), d'un volume de traitement planifié (rose-pêche) et d'un faisceau de protons (lignes horizontales noires venant de gauche). Superposés sont des taches variées, où la taille des cercles indique des poids ponctuels. Les couleurs ponctuelles indiquent la couche d'iso-énergie dont le spot fait partie. À droite, les poids (comptage de protons) par couche iso-énergie sont tracés. En haut, le profil PG est esquissé pour une couche iso-énergie. En bas, le profil PG est esquissé pour la couche iso-depth (spots de différentes couleurs sur la ligne pointillée). En violet et en rouge, les quatre points sur l'extrémité distale des deux lignes de faisceau les plus centrales, l'effet de l'interpolation est illustré: chaque position distale n'est pas toujours prévue avec un poids élevé.

Une différence était attendue entre le nombre de protons par tache comme prévu et suffisant pour la reconstruction. En fonction de la ligne du faisceau, l'unité minimale de détection PG peut être une tache (balayage des faisceaux) ou une couche iso-énergie (diffusion passive). Étant donné que la plupart des nouveaux centres utilisent la numérisation active du faisceau, de nouvelles méthodes de regroupement sont possibles. Ici, une méthode de groupage instantané est proposée en fonction de la notion de proximité approfondie. Sur la CT de planification, la distribution de dose en 3D est calculée par point, puis une dose de FOP est déterminée avec le même algorithme que pour PG FOP. Les FOP pour chaque point sont binned le long de la ligne du faisceau, et ces bacs nous appelons iso-couches de profondeur. L'hypothèse est que, en raison des inhomogénéités, les protons dans une couche d'iso-énergie peuvent avoir des FOP très différents chez le patient. Cette méthode de regroupement iso-depth sera comparée au groupement par iso-energy.

	Spot A	Spot B	Spot C
Dose shift [mm]	2.77	4.08	12.4
PG emission shift [mm]	2.32	3.34	13.9
Iso-energy layer [MeV]	145.86	143.02	143.02
x-pos [mm]	-24.0	-24.0	-24.0
y-pos [mm]	16.0	40.0	-40.0
Nr. protons	$2.75 \cdot 10^7$	$2.72 \cdot 10^7$	$4.73 \cdot 10^7$

Table A.2: Summary of the properties of the selected spots.

Figures du mérite

Lors de l'évaluation d'un détecteur, principalement la précision et la précision sont intéressantes: la caméra PG estime-t-elle correctement le FOP (précis) et est-ce le résultat reproductible (précis)? Un point de vue plus clinique pourrait être: le détecteur signale-t-il correctement des changements significatifs? Une estimation FOP précise ou une estimation du décalage FOP n'est peut-être pas nécessaire, mais une indication correcte d'un changement significatif *est*.

En utilisant la méthode par lots, nous réalisons entre 20 et 50 simulations pour chaque expérience, chacune ayant pour résultat une estimation FOP, ce qui nous donne un μ_{FOP} et un écart type σ_{FOP} pour un certaine expérience. Puisque nous étudions l'effet de remplacer le CT par un RPCT pour simuler un changement de patient, nous effectuons chaque expérience avec CT et RPCT. Chaque estimation FOP des réalisations N CT est comparée à chaque estimation FOP des réalisations N RPCT, entraînant des mesures N × N possibles emph FOP shift. Ces distributions ont μ_Δ et σ_Δ. Le décalage initialement obtenu avec la dose est noté Δ_{dose}

Le classement des performances des détecteurs se fera selon quelques chiffres de mérite:

- Précision: $|\mu_\Delta - \Delta_{dose}|$.
- Précision: σ_Δ. Pour cette estimation de l'écart-type de la distribution gaussienne, un écart type peut être calculé à nouveau en fonction du nombre de réalisation *n* utilisé pour l'obtenir: $\sigma(\sigma_\Delta) = \frac{\sigma_\Delta}{\sqrt{2\times(n-1)}}$, comme Leo (1994, formula 4.54).
- Confiance: le pourcentage de réalisations RPCT FOP qui se situent en dehors de $\mu_{FOP,CT} \pm 2\sigma_{FOP,CT}$ indique la probabilité qu'une différence par rapport au FOP attendu soit mesurée. En d'autres termes, étant donné que dans cette analyse, nous savons qu'un changement doit être détecté, quelle est la probabilité que nous allons faire? Il sera désigné par P_Δ.

Selection spot

Trois points (table A.2) ont été sélectionnés pour être présentés dans cette étude. Spot C, avec un grand décalage de plus d'un centimètre avec une odeur de dépassement: le faisceau sort partiellement le patient en provoquant l'élévation distale vue dans la figure **??**. Un déplacement prévu de plus d'un centimètre devrait être détectable de manière fiable pour toute caméra PG. Les points A et B sont plus difficiles; Avec des décalages prévus entre 2 et 4 mm, ils représentent un décalage minimal que les caméras PG devraient être capables de détecter.

A.5.3 Résultats

La première étude porte sur la performance des deux modèles de caméra PG collimatés sur un cas de patient avec une CT et une CT de suivi. Deuxièmement, nous étudions comment deux caméras PG

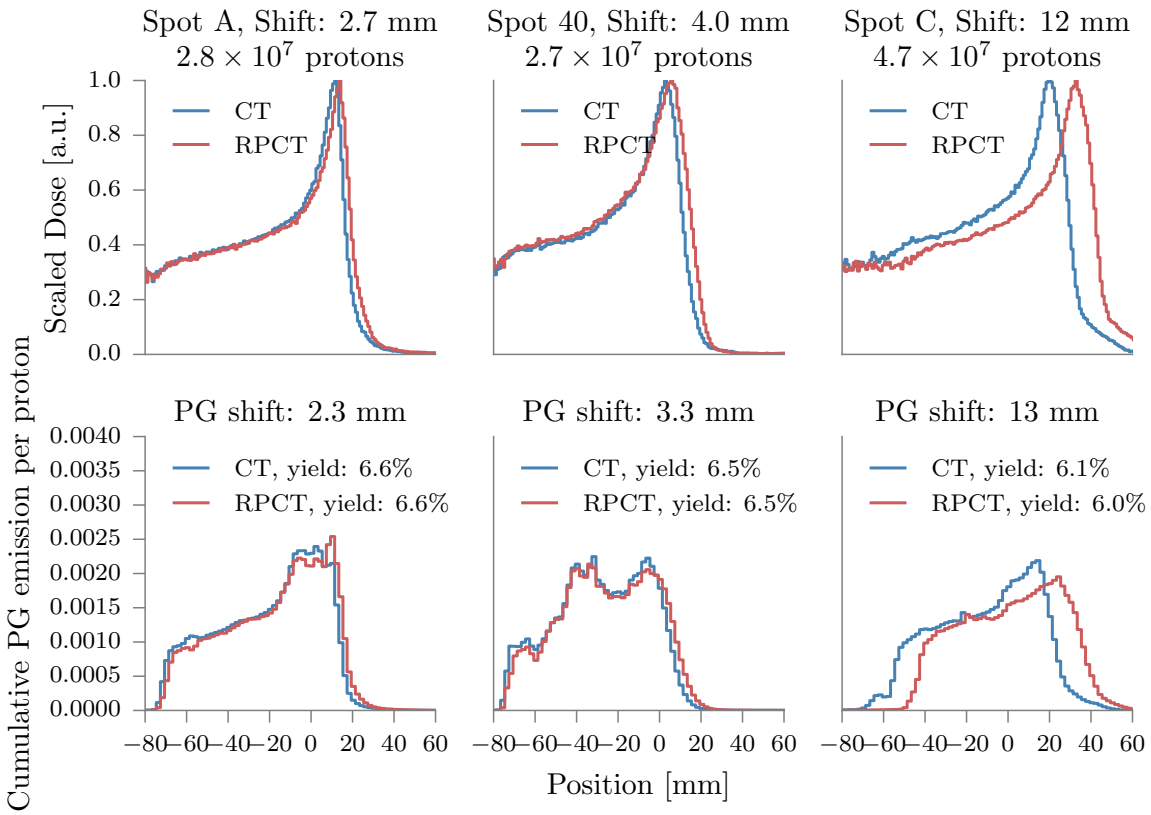


Figure A.23: Les trois points choisis, en fonction de leurs déplacements et de la qualité du changement. La dose est normalisée par masse, ce qui explique le manque de structure. Ligne supérieure: profils de dose, rangée inférieure: profils d'émission PG utilisés dans la deuxième étape de la simulation (détection PG). Le rendement est le rendement intégral de l'image entière, donc 6 % signifie que 6 protons sur 100 génèrent une PG.

	Spot A shift [mm]	Spot B shift [mm]	Spot C shift [mm]
Dose	2.77	4.08	12.4
PG emission	2.32	3.34	13.9
PG + PSF	2.61	2.91	11.9
μ_Δ MPS @ 10^9	2.68±0.77	3.23±0.77	12.6±1.15
μ_Δ KES @ 10^9	2.56±1.93	3.27±2.24	9.79±2.25
μ_Δ MPS @ spot weight	2.51±4.05	3.76±4.36	12.5±4.73
μ_Δ KES @ spot weight	-1.2±28.9	-0.2±28.9	4.30±26.2
	Spot A P_Δ	Spot B P_Δ	Spot C P_Δ
P_Δ MPS @ 10^9	99%	99%	100%
P_Δ KES @ 10^9	43%	59%	99%
P_Δ MPS @ spot weight	20%	24%	96%
P_Δ KES @ spot weight	1.9%	1.6%	1.4%
	Spot A spot weight	Spot B spot weight	Spot C spot weight
Nr. protons	$2.75 \cdot 10^7$	$2.72 \cdot 10^7$	$4.73 \cdot 10^7$

Table A.3: Résultats de la figure A.24 résumés: la précision peut être obtenue en comparant les changements, la probabilité de mesurer le changement est estimée avec P_Δ. Avec "PG + PSF", le FOP sur le profil d'émission PG converti avec un Gaussien avec une FWHM de 20 mm est indiqué. L'effet géométrique du détecteur qui ne mesure pas les PG aux émissions, mais à la détection, approché par la convolution, conduit à des FOP déplacés vers l'avant ou vers l'aval.

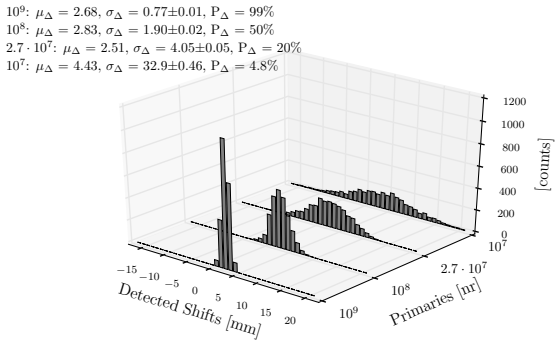
fonctionnent en fonction du nombre de primaires pour un point donné. Pour conclure, nous proposons une nouvelle méthode de sommation, nous discutons pourquoi il s'agit d'une amélioration par rapport à la sommation de la couche d'iso-énergie et démontrent son effet.

Un seul point modulé en poids

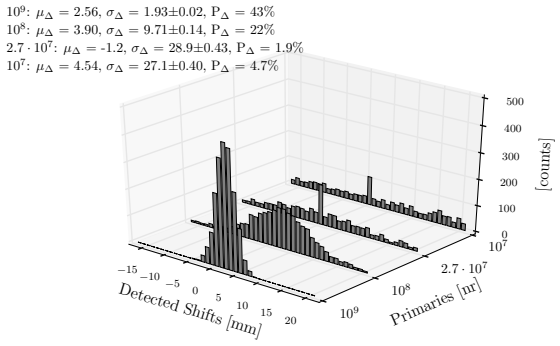
La procédure décrite pour les figures du mérite et démontrée dans l'étape de validation, est effectuée pour tout le nombre de primaires, pour les deux caméras, pour les points A, B et C. Les graphiques des résultats finaux, les distributions de décalage FOP et leurs μ_Δ, σ_Δ et P_Δ sont présentés dans la figure A.24.

La table **??** résume les résultats pertinents pour les changements estimés. Avec des primaires de 10^9, les deux caméras sont à moins de 2 mm du changement d'émission "PG + PSF", à l'exception de la caméra KES dans Spot C. Ici, lorsque son σ_Δ est considéré, il est dans un Σ de la valeur attendue. Aux poids ponctuels prescrits, le MPS converge toujours sur le bon μ_Δ, mais le KES n'est plus. Étant donné que le σ_Δ pour le KES est de l'ordre d'un quart de la FoV de la caméra, il semble juste de tirer la conclusion qu'aucun FOP correct n'est estimé à ces poids ponctuels.

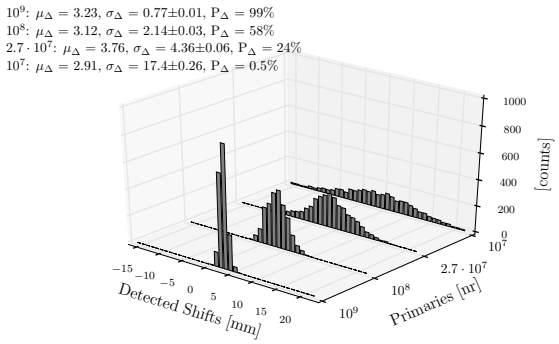
En se tournant vers P_Δ à 10^9 primaires, nous observons que le MPS fonctionne très bien: pour n'importe quel endroit, la probabilité qu'une seule réalisation RPCT indique un changement de CT est d'au moins 99 %. Le KES atteint la même probabilité que pour le grand décalage de Spot C. À leurs points de pointage prévus, aucune caméra ne donne de bons résultats à l'exception du MPS pour Spot C: même si la précision est de 4,73 mm, la taille du changement réel est suffisante Pour garder les distributions CT et ROPT FOP suffisamment séparées pour être à 96 %, une mesure unique aurait



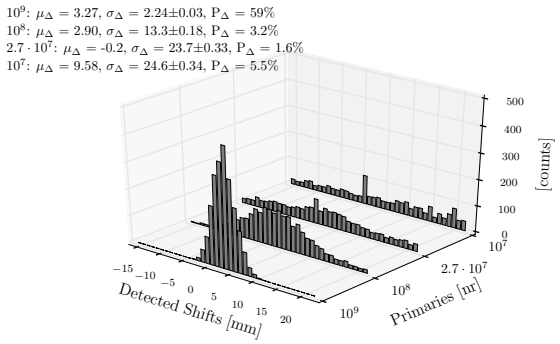
MPS, Spot A



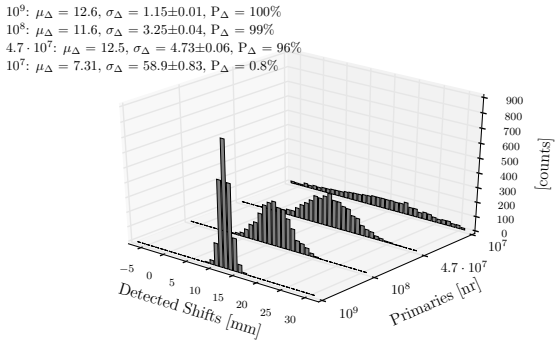
KES, Spot A



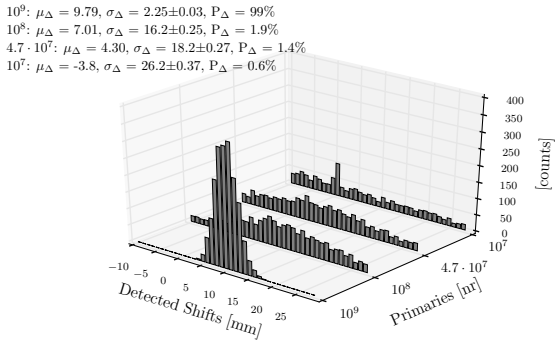
MPS, Spot B



KES, Spot B



MPS, Spot C



KES, Spot C

Figure A.24: Les distributions de différence FOP sont tracées pour les spots (en rangée) et pour les caméras MPS et KES (colonne). Par sous-trame, quatre fois deux distributions FOP sont affichées. À l'avant, le FOP se déplace en fonction de 10 9, de 10^8, du poids spot original et de 10^7 lorsque nous allons vers l'arrière de l'intrigue 3D.

	Spot A		Spot B		Spot C	
	iso-energy	iso-depth	iso-energy	iso-depth	iso-energy	iso-depth
Dose	2.47	1.60	6.40	4.07	6.40	6.40
PG emission	4.94	2.32	7.12	3.78	7.12	7.12
PG + PSF	3.20	2.18	4.94	4.51	4.94	6.25
μ_Δ MPS	3.13±1.16	2.18±0.89	4.82±1.00	3.67±0.96	4.72±1.17	5.77±1.05
μ_Δ KES	4.20±3.61	1.80±2.55	4.90±3.38	3.75±2.71	4.15±3.82	5.15±2.87
P_Δ MPS	96%	92%	99%	99%	99%	99%
P_Δ KES	34%	14%	54%	57%	99%	99%
Nr. protons	$0.84 \cdot 10^9$	$0.84 \cdot 10^9$	$1.07 \cdot 10^9$	$1.07 \cdot 10^9$	$1.07 \cdot 10^9$	$0.98 \cdot 10^9$

Table A.4: Vue d'ensemble de la comparaison iso-énergie versus iso-profondeur: la précision peut être obtenue en comparant les décalages, la probabilité de mesurer le changement est estimée avec P_Δ. Tous les quarts sont donnés en unités de millimètres. Le fait que les émissions de PG et les changements de dose pour Spot C soient exactement les mêmes avec le groupe iso-energy et iso-depth for purement fortuite.

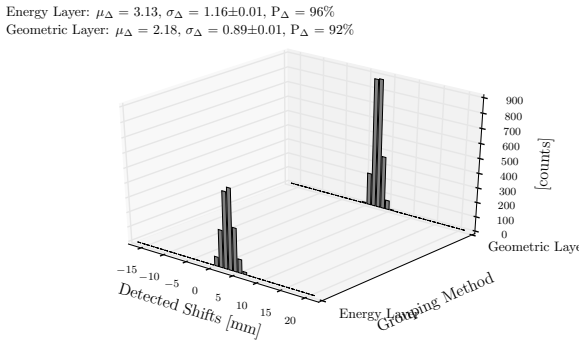
permis de constater qu'il y avait un changement significatif.

Spot groupage

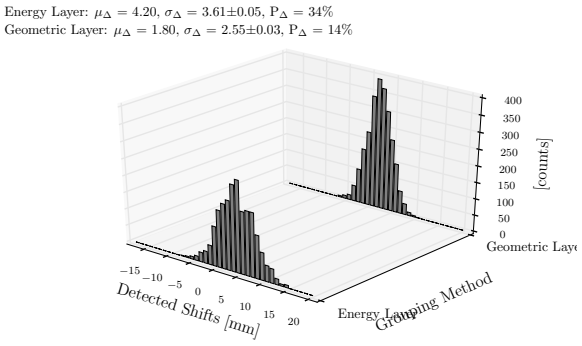
Étant donné que les mesures de décalage FOP Spot-by-Spot semblent être hors de la table, nous pouvons envisager de collecter des données lors de la livraison de plusieurs points. Sur la base de l'analyse par modulation ponctuelle, le nombre minimal de protons requis doit être d'au moins 10 9 primaires. Le point C fait commodément partie de la première couche d'iso-énergie (compté à partir de l'énergie la plus élevée) ayant au moins 10^9 de protons. Le groupe géométrique est construit en regroupant des points avec des FOP les plus proches de FOP de Spot C jusqu'à atteindre $10^9 \pm 5\%$ protons. Le nombre final est de $1.07 \cdot 10^9$ et $0.98 \cdot 10^9$ protons dans les couches iso-energy et iso-depth respectivement.

Si on effectue les groupes iso-profondeur et iso-énergétique basés sur les points 29 et 40, les parcelles de la figure A.25 sont obtenues. Maintenant, le nombre de primaires par groupe est presque identique entre les groupements (0.84 $cdot 10^9$ pour les groupes de points 20 et $1.07 \cdot 10^9$ pour les groupes Spot B). Le nombre inférieur de primaires se reflète dans P_Δ pour les deux caméras: les deux ne sont plus à 99% stables avec les groupes Spot C, mais moins. Le petit décalage de Spot A se traduit par de petits décalages dans les groupements, en particulier le groupement iso-depth. Ici, le changement FOP est bien dans la précision de la caméra KES, et il perd clairement son pouvoir de discriminer le signal CT pour le RPCT (P_Δ = 14 %). Le MPS montre, à l'exception du groupe iso-énergie sur Spot A, une légère amélioration sur μ_Δ par rapport à la place réglée à 10^9 protons. Pour chaque caméra et spot, la précision de la différence FOP (μ_Δ) s'améliore avec un regroupement iso-énergie sur iso-depth.

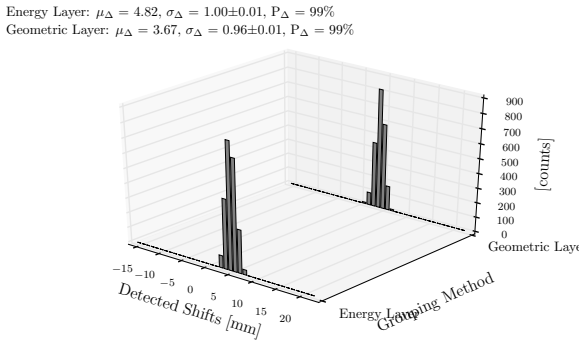
En observant tableau A.4, la meilleure précision FOP du KES pour le regroupement iso-depth est traduite par une probabilité plus élevée pour mesurer un changement significatif: μ_Δ amélioré de 38 % à 75 % . En regardant les groupements pour le point B, nous voyons que la convolution PG + PSF produit une différence FOP très similaire, alors que les émissions et les doses de FOP à PG sont très différentes.



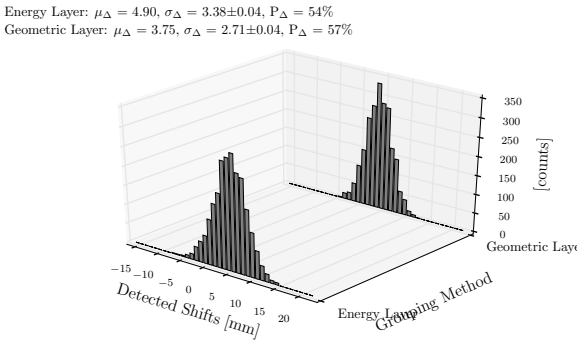
MPS, FOP shift distribution, Spot A



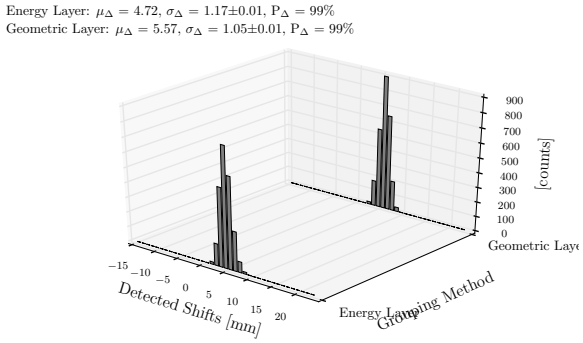
KES, FOP shift distribution, Spot A



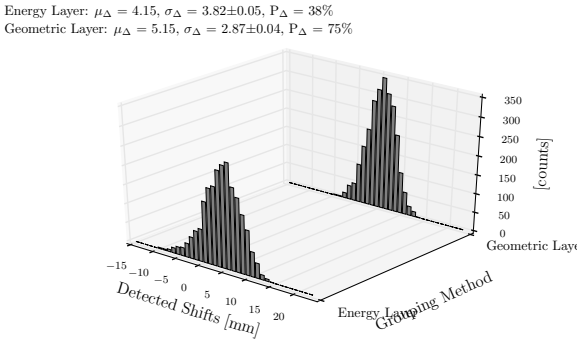
MPS, FOP shift distribution, Spot B



KES, FOP shift distribution, Spot B



MPS, FOP shift distribution, Spot C



KES, FOP shift distribution, Spot C

Figure A.25: Sur la rangée supérieure, les distributions de décalage FOP pour les deux méthodes de regroupement pour les groupes iso-energy et iso-depth (tous deux avec 0.84×10^9 primary) sur Spot A sont affichées. La ligne du milieu montre les résultats pour le groupage des points sur Spot B, avec des primaires de 1.07×10^9 pour les deux groupements. Sur la ligne inférieure, les groupes pour le point C sont affichés, avec respectivement 1.07×10^9 et 0.98×10^9 primaires pour les couches iso-energy et iso-depth respectivement. Les résultats dans la colonne de gauche ont été obtenus avec la caméra MPS et celles à droite avec la caméra KES.

A.5.4 Conclusion

Nous avons présenté les premières simulations cliniques de départ à la fin d'une caméra à fente multi-parallèle pour la détection de gamma rapide. En outre, nous avons présenté la première comparaison en tête-à-tête de deux caméras PG collimatées dans des conditions réalistes. En outre, nous présentons une nouvelle méthode de regroupement de points basée sur la notion d'iso-profondeur. Une petite étude sur les poids cliniques des taches a été effectuée. Enfin, une figure de mérite est présentée qui fournit une estimation de la probabilité d'une FOP mesurée tombant à l'intérieur ou à l'extérieur de l'écart type prévu de $\pm 2\times$.

Au poids nominal nominal, aucune caméra ne peut mesurer un résultat correct. Les poids typiques des taches sont étudiés et, avec une tendance à la baisse des poids ponctuels pour les plans de faible tolérance, il semble peu probable que les deux caméras PG envisagées soient en mesure de produire des résultats cliniquement pertinents sur le terrain, sans amélioration significative. Le groupage ponctuel est un moyen d'améliorer les statistiques et d'activer la précision millimétrique de l'estimation FOP. Deux méthodes de groupage ponctuel sont étudiées et, pour la caméra KES, la précision du changement FOP s'améliore en utilisant le nouveau groupe iso-depth qui mérite d'être approfondi. Une meilleure probabilité que deux fois sur trois pour obtenir une indication significative du décalage (P_Δ) est probable pour les deux caméras, lors de l'utilisation d'un groupe de points iso-depth au lieu de regrouper par une couche d'iso-énergie.

A.6 Conclusions

A.6.1 Résumé des réalisations

La méthode de réduction de la variance vpgTLE a été mise en œuvre et validée dans le cadre de ce programme de thèse. La méthode a été jugée compatible avec les inhomogénéités, flexible en termes de rendement et offre une accélération d'un facteur de 10^3. Nous avons également noté qu'aucun travail n'a été effectué pour valider les antécédents PG dans les simulations PG, ce qui est essentiel si les simulations sont utilisées pour prédire les résultats du monde réel. La méthode vpgTLE est publiquement disponible dans le cadre de la version 7.1 de la Gate et permet aux chercheurs du monde entier de bénéficier sans autre exigence. Certains chercheurs ont déjà montré de l'intérêt et j'attends avec impatience les résultats qui peuvent être obtenus avec cette méthode.

Deuxièmement, la première simulation de thérapie protonale entièrement clinique surveillée par des détecteurs collimatés PG a été présentée, dans laquelle une véritable déviation clinique a été étudiée. VpgTLE a permis de compléter cette étude dans un délai raisonnable et a permis une analyse spot par point. Une nouvelle méthode de regroupement de points a été présentée et démontrée, et nous pensons que l'étude déplacera la discussion dans le suivi de la PG sur la façon dont ces détecteurs peuvent être utilisés pour obtenir des résultats cliniquement pertinents. Sur la base d'une étude des plans de traitement, la vérification PG par point est peu probable avec l'état actuel des caméras collimées PG, mais nous avons montré que le groupement de points géométriquement proches ne modifiait pas le σ de la FOP et donc le regroupement des points Permet la vérification PG.

Le titre de la proposition de recherche de doctorat est «Vers un contrôle du traitement en temps réel de la protonthérapie en utilisant l'imagerie par rayonnement rapide: simulation et optimisation du système». Dans quelle mesure avons-nous progressé vers ce but? Je pense que nous l'avons fait, en produisant un outil plus rapide pour les simulations PG, et en montrant comment il peut être utilisé pour étudier divers aspects de la détection PG, et cela nous a permis d'identifier les domaines dans lesquels des améliorations sont nécessaires (traitement du profil PG et groupage des points).

A.6.2 Autres pensées

Si j'aurais encore 6 mois, je reprendrais la recherche du section 5 de manière plus systématique. À l'heure actuelle, nous avons dépendu d'une méthode de sélection de points qui a donné lieu à trois points, que nous avons étudiés en détail. Peut-être était-il nécessaire de commencer petit, car nous avons appris beaucoup même à partir de seulement trois points, mais en fin de compte, une approche plus statistique où nous obtenons des FOP, leurs précisions et les précisions de FOP-shift pour tous les points du plan nous auraient fait plus Confiant sur la question de savoir si nous avons choisi de choisir des valeurs aberrantes ou non. Actuellement, les principales limitations sont l'espace disque requis pour l'image intermédiaire (une source de PG est de 1,5 GB environ, donc si une source de PG pour chaque point est souhaitée, 1,5 TB d'espace disque est requis) et le temps de simulation de vpgTLE Etape 2 (la propagation des photons dans le détecteur). La première étape doit être exécutée une seule fois et prend environ 3 heures pour des protons primaires de 10^6 (équivalent à 10^9 protons avec MC analogique) sur un noyau d'un ordinateur portable de 4 ans (Intel Core i7-3740QM CPU à 2,70 GHz), tandis que la deuxième étape prend $\approx$ 14 heures, sur un seul noyau, pour atteindre les statistiques pour 10^9 protons. La réduction de la taille de l'image intermédiaire peut être possible en exploitant l'encombrement de l'image 4D, ainsi qu'en réduisant les tailles de voxel dans des dimensions qui ne sont pas mesurées et en réduisant la résolution d'échantillonnage des sections transversales de production de PG à l'étape 0. Un maître L'élève a été attiré par l'aide pour valider un composant de détection forcée fixe et de réponse angulaire récemment développé dans Gate en remplacement de l'étape 2. Cela permettra d'accélérer la partie la plus lente d'une simulation vpgTLE, ce qui permet une répétition plus rapide des simulations de recherche. Il aurait été intéressant de simuler les profils de détection par point, qui pourraient ensuite être additionnés au post-traitement comme souhaité, selon la méthode de groupage précise. Nous avons itéré sur la méthode de regroupement, ce qui nécessitait la génération de nouvelles sources et la simulation des 50 simulations CT et 50 RPCT Stage 2, nous ralentissant un peu en ce qui concerne la recherche d'une stratégie de regroupement optimale. En outre, il s'agissait d'un patient avec un champ d'un plan; Nous avons rassemblé plus d'images CT + RPCT mais n'avons pas eu le temps d'analyser les plans.

A.7 Utilisation clinique de la détection de Gamme rapide

Nous avons vu différents modèles de détecteurs PG, chacun révélant différentes propriétés de la PG. Les ions plus lourds produisent plus de PG: un carbone primaire a une probabilité de production de PG d'environ 40 % alors que pour les protons ce nombre est de 5 à 10%, ce qui doit être équilibré par le fait que moins de carbone est nécessaire pour atteindre une certaine dose (un Facteur 5 moins d'ions de carbone est utilisé dans une étude comparative avec des protons dans Pinto et al. (2015)). Peut-être parce que la première étape dans la conception d'un détecteur est de savoir s'il peut etre fait du tout, on n'a pas beaucoup pensé à la façon dont un clinicien pourrait utiliser ces sorties. Les cliniciens travaillant en radiothérapie et en radiologie sont habitués à travailler avec (CB) CT et occasionnellement des images IRM. L'ajout d'une modalité est un défi majeur, car des milliers de cliniciens devraient être formés à la compréhension. De plus, en ce moment, fournir des sorties de détection PG sous la forme d'images semble être lointain. Il peut également arriver que les statistiques spot-by-spot sont trop faibles et que la collecte de données sur des groupes de points plus importants est une nécessité. Quel genre de groupes de points ont le plus de sens clinique? On peut imaginer qu'en regroupant les points dans la proximité géométrique et dans les structures cliniquement liées, on peut donner une réponse plus ciblée. Peut-être qu'un plan de traitement pourrait-il produire une liste de points sensibles à la gamme qui s'arrêtent près des AOR ou des structures de croisement susceptibles de se déformer. Un critère géométrique semble être un bon moyen d'assurer

la cohérence entre les écarts des taches au sein d'un groupe, au lieu d'annuler lorsque des taches dans différentes structures souffrant de différentes déformations sont regroupées.

Si nous gardons à l'esprit les sorties des caméras collimatées et chronométriques, le point d'émission de la PG et, à partir de cette position de chute inférée et peut-être de la portée, l'information pourrait être présentée en termes de seuils ou de tolérances: combien de points ou de points Les groupes dépassent et dépassent? Peut-être que nous pouvons mesurer le profil PG complet, et avec un profil planifié (Schumann et al., 2016), nous pourrions effectuer un profil correspondant (Gueth et al., 2013) pour produire par point (groupe) une probabilité de changement et une probabilité d'estimation sans décalage. En allant encore plus loin, les profils peuvent être convertis en dose, comme cela a été démontré dans Schumann et al. (2016). Ensuite, la dose peut être additionnée sur tous les points (groupe) s, de différents champs, et recouvert sur une CT. Cela fournirait le type de données d'image que les cliniciens connaissent et peuvent travailler avec tout de suite. Un taux de dosage d'un indice gamma est une méthode établie pour comparer les dosemaps. Si un CBCT est présent dans la salle de traitement, ainsi que des techniques d'enregistrement déformables améliorées (Park et al., 2017; Yan, 2010), un CT du jour peut être produit pour une correspondance plus précise du dosemap dérivé de PG et des structures des patients. Complétez-le avec une thérapie de particules adaptative, également basée sur l'enregistrement déformable CT + CBCT (Veiga et al., 2015), des gains réels peuvent être accessibles en termes de réduction des marges de tolérance acceptables.

Radiomics est une méthode médicale à venir constituée sur l'idée que les machines peuvent désormais être formées pour en déduire [beaucoup plus] des données cliniques telles que CTs (Yip and Aerts, 2016). Sur la base des vastes volumes de données médicales que chaque hôpital recueille pendant les traitements des patients, une grande approche des données peut entraîner un logiciel qui peut apprendre à extraire plus des données disponibles que les cliniciens (humains) actuellement. L'apprentissage par machine proposé dans Gueth et al. (2013) cherche à utiliser de telles méthodes pour que PG se replie. Malheureusement, la quantité de données de formation disponibles pour la détection de PG est extrêmement limitée, mais c'est un début prometteur d'avoir des traitements de détecteur de détecteur PG en cours (Richter et al., 2016; Verburg and Bortfeld, 2016). Une approche similaire pourrait être envisagée pour la planification du traitement: le logiciel peut être beaucoup mieux au moment de la dose de peinture basée sur des données en direct d'un CT + CBCT déformable avant et PG détection pendant le traitement.

Appendix B

Note On Independency

Variance of the prompt-γ spectrum TLE estimator

Jean Michel Létang

July 30, 2015

Exact formula. The prompt-γ spectrum TLE estimator is

$$\widehat{\boldsymbol{S}}(v) = \frac{1}{n}\sum_{i=1}^{n}\sum_{g=1}^{G}\boldsymbol{\Gamma}_{m_v}(E_g)L_i(E_g,v) \tag{1}$$

where $L_i(E_g, v)$ corresponds to the total track length of protons of energy E_g in volume v within proton i history. The variance of this estimator can written without approximation from the distribution of the n scored quantities event-wise as

$$n\sigma^2\left[\widehat{S}_{\mathrm{TLE},E_\gamma}\right] = \frac{1}{n}\sum_{i=1}^{n}A_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n}A_i\right)^2 \tag{2}$$

with A_i the scored prompt-γ spectrum value in the energy bin E_γ for event i is

$$A_i = \sum_{g=1}^{G}\Gamma_{m_v E_\gamma}(E_g)L_i(E_g,v) \tag{3}$$

Therefore, to compute on-line the error on the prompt-γ TLE spectrum, two prompt-γ spectrum vectors must be scored over the events: $\sum_{i=1}^{n}A_i^2$ and $\sum_{i=1}^{n}A_i$. This summation process takes longer than the simple track-length sum based equation for the variance in equation 8, which is based on independency of the track-length distribution with respect to the proton energy.

Reformulation. Using $\bar{L}_n(E_g, v)$ and $\bar{L}_{2n}(E_g, v)$ the arithmetic mean of the n values of $L_i(E_g, v)$ and $L_i^2(E_g, v)$ respectively, we get

$$\begin{aligned} n\sigma^2\left[\widehat{S}_{\mathrm{TLE},E_\gamma}\right] &= \sum_{g=1}^{G}\Gamma_{m_v E_\gamma}^2(E_g)\bar{L}_{2n}(E_g,v) + \frac{2}{n}\sum_{i=1}^{n}\sum_{g<h}\Gamma_{m_v E_\gamma}(E_g)L_i(E_g,v)\Gamma_{m_v E_\gamma}(E_h)L_i(E_h,v) && (4)\\ &\quad - \left(\sum_{g=1}^{G}\Gamma_{m_v E_\gamma}(E_g)\bar{L}_n(E_g,v)\right)^2 && (5)\end{aligned}$$

We see from this formula that the triangular summation $\sum_{g<h} = \sum_{h=2}^{G}\sum_{g=1}^{h-1}$ over the proton energy cannot be computed on the sole knowledge of $\overline{L}_n(E_g, v)$ and $\overline{L}_{2n}(E_g, v)$. We now have two options: either to consider $L_i(E_g, v)$ as independent variables with respect to the proton energy, or to consider them fully correlated i.e. identical.

Independent track length distribution. Now, if all $L_i(E, v)$ variables are independent between all proton energies E for a given i event, the variance becomes

$$n\sigma^2\left[\widehat{S}_{\mathrm{TLE},E_\gamma}\right] = \left(\sum_{g=1}^{G}\Gamma^2_{m_v E_\gamma}(E_g)\overline{L}_{2n}(E_g, v) + 2\sum_{g<h}\Gamma_{m_v E_\gamma}(E_g)\Gamma_{m_v E_\gamma}(E_h)\overline{L}^2_n(E_g, v)\right) \tag{6}$$

$$- \left(\sum_{g=1}^{G}\Gamma_{m_v E_\gamma}(E_g)\overline{L}_n(E_g, v)\right)^2 \tag{7}$$

which is

$$n\sigma^2\left[\widehat{S}_{\mathrm{TLE},E_\gamma}\right] = \sum_{g=1}^{G}\Gamma^2_{m_v E_\gamma}(E_g)\left(\overline{L}_{2n}(E_g, v) - \overline{L}^2_n(E_g, v)\right) \tag{8}$$

and we get the random variance contribution that was given in the submitted article

$$\sigma^2\left[\widehat{S}_{\mathrm{TLE},E_\gamma}\right] = \sum_{g=1}^{G}\Gamma^2_{m_v E_\gamma}(E_g)\frac{\sigma^2[L(E_g, v)]}{n} \tag{9}$$

which gives the lower random variance boundary for the TLE estimator.

Identical track length distribution. If all $L_i(E, v)$ variables are identical (i.e. entirely correlated) between all proton energies E for a given i event, the variance becomes

$$n\sigma^2\left[\widehat{S}_{\mathrm{TLE},E_\gamma}\right] = \left(\sum_{g=1}^{G}\Gamma^2_{m_v E_\gamma}(E_g)\overline{L}_{2n}(E_g, v) + 2\sum_{g<h}\Gamma_{m_v E_\gamma}(E_g)\Gamma_{m_v E_\gamma}(E_h)\overline{L}_{2n}(E_g, v)\right) \tag{10}$$

$$- \left(\sum_{g=1}^{G}\Gamma_{m_v E_\gamma}(E_g)\overline{L}_n(E_g, v)\right)^2 \tag{11}$$

or

$$n\sigma^2\left[\widehat{S}_{\mathrm{TLE},E_\gamma}\right] = \sum_{g=1}^{G}\Gamma^2_{m_v E_\gamma}(E_g)\left(\overline{L}_{2n}(E_g, v) - \overline{L}^2_n(E_g, v)\right) \tag{12}$$

$$+2\sum_{g<h}\Gamma_{m_v E_\gamma}(E_g)\Gamma_{m_v E_\gamma}(E_h)\left(\overline{L}_{2n}(E_g, v) - \overline{L}^2_n(E_g, v)\right) \tag{13}$$

which gives

$$n\sigma^2\left[\widehat{S}_{\mathrm{TLE},E_\gamma}\right] = \sum_{g=1}^{G}\Gamma^2_{m_vE_\gamma}(E_g)\sigma^2[L(E_g,v)] + 2\sum_{g<h}\Gamma_{m_vE_\gamma}(E_g)\Gamma_{m_vE_\gamma}(E_h)\sigma^2[L(E_g,v)] \tag{14}$$

or finally

$$\sigma^2\left[\widehat{S}_{\mathrm{TLE},E_\gamma}\right] = \sum_{g=1}^{G}\Gamma^2_{m_vE_\gamma}(E_g)\frac{\sigma^2[L(E_g,v)]}{n} + 2\sum_{g<h}\Gamma_{m_vE_\gamma}(E_g)\Gamma_{m_vE_\gamma}(E_h)\frac{\sigma[L(E_g,v)]\sigma[L(E_h,v)]}{n} \tag{15}$$

which gives the upper random variance boundary for the TLE estimator.

Appendix C

Administrative: Écoles Doctorales

Département FEDORA – INSA Lyon - Ecoles Doctorales – Quinquennal 2016-2020

SIGLE	ECOLE DOCTORALE	NOM ET COORDONNEES DU RESPONSABLE
CHIMIE	**CHIMIE DE LYON** **http://www.edchimie-lyon.fr** Sec : Renée EL MELHEM Bat Blaise Pascal 3e etage secretariat@edchimie-lyon.fr Insa : R. GOURDON	M. Stéphane DANIELE Institut de Recherches sur la Catalyse et l'Environnement de Lyon IRCELYON-UMR 5256 Équipe CDFA 2 avenue Albert Einstein 69626 Villeurbanne cedex directeur@edchimie-lyon.fr
E.E.A.	**ELECTRONIQUE, ELECTROTECHNIQUE, AUTOMATIQUE** **http://edeea.ec-lyon.fr** Sec : M.C. HAVGOUDOUKIAN Ecole-Doctorale.eea@ec-lyon.fr	M. Gérard SCORLETTI Ecole Centrale de Lyon 36 avenue Guy de Collongue 69134 ECULLY Tél : 04.72.18 60.97 Fax : 04 78 43 37 17 Gerard.scorletti@ec-lyon.fr
E2M2	**EVOLUTION, ECOSYSTEME, MICROBIOLOGIE, MODELISATION** **http://e2m2.universite-lyon.fr** Sec : Sylvie ROBERJOT Bât Atrium - UCB Lyon 1 04.72.44.83.62 Insa : H. CHARLES secretariat.e2m2@univ-lyon1.fr	M. Fabrice CORDEY CNRS UMR 5276 Lab. de géologie de Lyon Université Claude Bernard Lyon 1 Bât Géode 2 rue Raphaël Dubois 69622 VILLEURBANNE Cédex Tél : 06.07.53.89.13 cordey@univ-lyon1.fr
EDISS	**INTERDISCIPLINAIRE SCIENCES-SANTE** **http://www.ediss-lyon.fr** Sec : Sylvie ROBERJOT Bât Atrium - UCB Lyon 1 04.72.44.83.62 Insa : M. LAGARDE secretariat.ediss@univ-lyon1.fr	Mme Emmanuelle CANET-SOULAS INSERM U1060, CarMeN lab, Univ. Lyon 1 Bâtiment IMBL 11 avenue Jean Capelle INSA de Lyon 696621 Villeurbanne Tél : 04.72.68.49.09 Fax :04 72 68 49 16 Emmanuelle.canet@univ-lyon1.fr
INFOMATHS	**INFORMATIQUE ET MATHEMATIQUES** **http://infomaths.univ-lyon1.fr** Sec :Renée EL MELHEM Bat Blaise Pascal 3e etage infomaths@univ-lyon1.fr	Mme Sylvie CALABRETTO LIRIS – INSA de Lyon Bat Blaise Pascal 7 avenue Jean Capelle 69622 VILLEURBANNE Cedex Tél : 04.72. 43. 80. 46 Fax 04 72 43 16 87 Sylvie.calabretto@insa-lyon.fr
Matériaux	**MATERIAUX DE LYON** **http://ed34.universite-lyon.fr** Sec : M. LABOUNE PM : 71.70 –Fax : 87.12 Bat. Direction Ed.materiaux@insa-lyon.fr	M. Jean-Yves BUFFIERE INSA de Lyon MATEIS Bâtiment Saint Exupéry 7 avenue Jean Capelle 69621 VILLEURBANNE Cedex Tél : 04.72.43 71.70 Fax 04 72 43 85 28 jean-yves.buffiere@insa-lyon.fr
MEGA	**MECANIQUE, ENERGETIQUE, GENIE CIVIL, ACOUSTIQUE** **http://mega.universite-lyon.fr** Sec : M. LABOUNE PM : 71.70 –Fax : 87.12 Bat. Direction mega@insa-lyon.fr	M. Philippe BOISSE INSA de Lyon Laboratoire LAMCOS Bâtiment Jacquard 25 bis avenue Jean Capelle 69621 VILLEURBANNE Cedex Tél : 04.72 .43.71.70 Fax : 04 72 43 72 37 Philippe.boisse@insa-lyon.fr
ScSo	**ScSo*** **http://recherche.univ-lyon2.fr/scso/** Sec : Viviane POLSINELLI Brigitte DUBOIS Insa : J.Y. TOUSSAINT Tél : 04 78 69 72 76 viviane.polsinelli@univ-lyon2.fr	M. Christian MONTES Université Lyon 2 86 rue Pasteur 69365 LYON Cedex 07 Christian.montes@univ-lyon2.fr

*ScSo : Histoire, Géographie, Aménagement, Urbanisme, Archéologie, Science politique, Sociologie, Anthropologie

Appendix D

Administrative: Folio

FOLIO ADMINISTRATIF

THESE DE L'UNIVERSITE DE LYON OPEREE AU SEIN DE L'INSA LYON

NOM : Huisman
(avec précision du nom de jeune fille, le cas échéant)

DATE de SOUTENANCE : 19/05/2017

Prénoms : Brent Fokke Bram

TITRE : Accelerated Clinical Prompt Gamma simulations for Proton Therapy

NATURE : Doctorat

Numéro d'ordre : AAAALYSEIXXXX

Ecole doctorale : N° 160 : ÉLECTRONIQUE, ÉLECTROTECHNIQUE, AUTOMATIQUE

Spécialité : Traitement du signal et de l'Image

RESUME :

The first part of this PhD program is the development, analysis and release of a variance reduction method for the simulation of prompt gammas (PGs) in clinical proton therapy simulations. The variance reduction method (named vpgTLE) is a two-stage track length estimation method developed to estimate the PG yield in voxelized volumes. As primary particles are propagated throughout the patient CT, the PG yields are computed at each step, resulting in a voxelized image of PG production yield. The second stage uses this intermediate image as a source to generate and propagate the number of PGs throughout the rest of the scene geometry, e.g. into a detection device, corresponding to the number of primaries desired. For both a geometrical heterogeneous phantom and a complete patient CT treatment plan with respect to analog MC, at a convergence level of 2\% relative uncertainty in the 90\% yield region, a gain of around 10^3 was achieved. The method agrees with reference analog MC simulations to within 10^{-4}, with negligible bias.

The second part of this PhD program is the study of PG fall-off position (FOP) estimation in clinical simulations. The number of protons (spot weight) required for a consistent FOP estimate was investigated for two PG cameras, a multi-parallel slit and a knife edge design, for a single spot of a fully clinical simulation of a patient treatment. By studying recent treatment plans from various proton clinics, we observe very few spots with weights over 10^8. We did not manage to detect the morphological change present, an approximately 13mm shift, between the (RP)CT with either PG camera, with such statistics. Only for 10^9 primaries, with one of the cameras, the change may be expected to be detected.A new spot-grouping method is proposed that combines better measurement statistics with fall-off preservation.

MOTS-CLÉS : Simulations, Monte Carlo, Prompt Gamma, Proton Therapy, Treatment Verification, Clinical Simulations

Laboratoire (s) de recherche : Creatis, IPNL

Directeur de thèse: David Sarrut, Étienne Testa

Président de jury :

Composition du jury : Peter Dendooven, Sébastien Jan, Magdalena Rafecas, Edmond Sterpin

Bibliography

U. Amaldi and S. Braccini. Present challenges in hadrontherapy techniques. *The European Physical Journal Plus*, 126(7):1–15, jul 2011. ISSN 2190-5444. doi: 10.1140/epjp/i2011-11070-4. URL http://www.springerlink.com/index/10.1140/epjp/i2011-11070-4.

Ugo Amaldi and Gerhard Kraft. Radiotherapy with beams of carbon ions. *Reports on Progress in Physics*, 68(8):1861–1882, aug 2005. ISSN 0034-4885. doi: 10.1088/0034-4885/68/8/R04. URL http://stacks.iop.org/0034-4885/68/i=8/a=R04?key=crossref.9a0f1e1c20da5f308fd7ce9a8adad161.

N Arbor, D Dauvergne, G Dedes, J M Létang, K Parodi, C T Quiñones, E Testa, and S Rit. Monte Carlo comparison of x-ray and proton CT for range calculations of proton therapy beams. *Physics in Medicine and Biology*, 60(19):7585, 2015. ISSN 0031-9155. doi: 10.1088/0031-9155/60/19/7585. URL http://stacks.iop.org/0031-9155/60/i=19/a=7585.

M. Benedikt and a. Wrulich. MedAustron—Project overview and status. *The European Physical Journal Plus*, 126(7), jul 2011. ISSN 2190-5444. doi: 10.1140/epjp/i2011-11069-9. URL http://www.springerlink.com/index/10.1140/epjp/i2011-11069-9.

M.J. Berger. *ESTAR, PSTAR and ASTAR: Computer Programs for Calculating Stopping Powers and Ranges for Electrons, Protons and Helium Ions*. National Institute of Standards and Technology, Gaithersburg, MD., 1995. URL http://www.iaea.org/inis/collection/NCLCollectionStore/_Public/26/063/26063376.pdf.

Aleksandra K Biegun, Enrica Seravalli, Patrícia Cambraia Lopes, Ilaria Rinaldi, Marco Pinto, David C Oxley, Peter Dendooven, Frank Verhaegen, Katia Parodi, Paulo Crespo, and Dennis R Schaart. Time-of-flight neutron rejection to improve prompt gamma imaging for proton range verification: a simulation study. *Physics in medicine and biology*, 57(20):6429–44, oct 2012. ISSN 1361-6560. doi: 10.1088/0031-9155/57/20/6429. URL http://www.ncbi.nlm.nih.gov/pubmed/22996154.

David J Brenner and Eric J Hall. Secondary neutrons in clinical proton radiotherapy: a charged issue. *Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology*, 86(2):165–70, feb 2008. ISSN 0167-8140. doi: 10.1016/j.radonc.2007.12.003. URL http://www.ncbi.nlm.nih.gov/pubmed/18192046.

Nichola Burridge, Ali Amer, Tom Marchant, Jonathan Sykes, Julie Stratford, Ann Henry, Catherine McBain, Pat Price, and Chris Moore. Online adaptive radiotherapy of the bladder: Small bowel irradiated-volume reduction. *International Journal of Radiation Oncology Biology Physics*, 66(3):892–897, 2006. ISSN 03603016. doi: 10.1016/j.ijrobp.2006.07.013.

J. Donald Chapman and Colin J. Gillespie. The power of radiation biophysics - Let's use it. *International Journal of Radiation Oncology Biology Physics*, 84(2):309–311, 2012. ISSN 03603016. doi: 10.1016/j.ijrobp.2012.04.020. URL http://dx.doi.org/10.1016/j.ijrobp.2012.04.020.

Benjamin M Clasie, Jacob B Flanz, and Hanne M Kooy. Interpolation of tabulated proton Bragg peaks. *Phys. Med. Biol. Phys. Med. Biol*, 57(57):405–409, 2012. ISSN 0031-9155. doi: 10.1088/0031-9155/57/21/N405. URL http://iopscience.iop.org/0031-9155/57/21/N405.

G. Coutrakon, J. Cortese, a. Ghebremedhin, J. Hubbard, J. Johanning, P. Koss, G. Maudsley, C. R. Slater, C. Zuccarelli, and J. Robertson. Microdosimetry spectra of the Loma Linda proton beam and relative biological effectiveness comparisons. *Medical Physics*, 24(9):1499, 1997. ISSN 00942405. doi: 10.1118/1.598038. URL http://scitation.aip.org/content/aapm/journal/medphys/24/9/10.1118/1.598038.

M Cunha, C Monini, E Testa, and M Beuve. NanOx, a new model to predict cell survival in the context of particle therapy. *Physics in Medicine and Biology*, 62(4):1248–1268, feb 2017. ISSN 0031-9155. doi: 10.1088/1361-6560/aa54c9. URL http://stacks.iop.org/0031-9155/62/i=4/a=1248?key=crossref.a946ff4dc79a733425207c4c9eb6ad52.

G Dedes, M Pinto, D Dauvergne, N Freud, J Krimmer, J M Létang, C Ray, and E Testa. Assessment and improvements of Geant4 hadronic models in the context of prompt-gamma hadrontherapy monitoring. *Physics in medicine and biology*, 59(7):1747–72, 2014. ISSN 1361-6560. doi: 10.1088/0031-9155/59/7/1747. URL http://www.ncbi.nlm.nih.gov/pubmed/24619152.

P Dendooven, H J T Buitenhuis, F Diblen, P N Heeres, A K Biegun, F Fiedler, M-J van Goethem, E R van der Graaf, and S Brandenburg. Short-lived positron emitters in beam-on PET imaging during proton therapy. *Physics in Medicine and Biology*, 60(23):8923–8947, 2015. ISSN 0031-9155. doi: 10.1088/0031-9155/60/23/8923. URL http://stacks.iop.org/0031-9155/60/i=23/a=8923?key=crossref.8976067304f2e5a0efda9c42424e628a.

Marco Durante and Harald Paganetti. Nuclear physics in particle therapy : a review. *Reports on Progress in Physics*, 00:1–59, 2016. ISSN 0034-4885. doi: 10.1088/0034-4885/79/9/096702. URL http://dx.doi.org/.

W El Kanawati, J M Létang, D Dauvergne, M Pinto, D Sarrut, É Testa, and N Freud. Monte Carlo simulation of prompt γ-ray emission in proton therapy using a specific track length estimator. *Physics in Medicine and Biology*, 60(20):8067–8086, 2015. ISSN 0031-9155. doi: 10.1088/0031-9155/60/20/8067. URL http://stacks.iop.org/0031-9155/60/i=20/a=8067?key=crossref.d38c03af2b13b06ae763a331d1d25dd2.

W. Enghardt, P. Crespo, F. Fiedler, R. Hinz, K. Parodi, J. Pawelke, and F. Pönisch. Charged hadron tumour therapy monitoring by means of PET. *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 525(1-2):284–288, 2004. ISSN 01689002. doi: 10.1016/j.nima.2004.03.128.

Samuel España and Harald Paganetti. The impact of uncertainties in the CT conversion algorithm when predicting proton beam ranges in patients from dose and PET-activity distributions. *Physics in medicine and biology*, 55(24):7557–7571, 2010. ISSN 0031-9155. doi: 10.1088/0031-9155/55/24/011.

P. Fossati, S. Molinelli, N. Matsufuji, M. Ciocca, A. Mirandola, A. Mairani, J. Mizoe, A. Hasegawa, R. Imai, T. Kamada, R. Orecchia, and H. Tsujii. Dose prescription in carbon ion radiotherapy: a planning study to compare NIRS and LEM approaches with a clinically-oriented strategy. *Physics in Medicine and Biology*, 57(22):7543–7554, 2012. ISSN 0031-9155. doi: 10.1088/0031-9155/57/22/7543.

N Freud, F ROELLINGHOFF, E Testa, M PINTO, and J Smeets. Apparatus for particle therapy verification comprising a collimator with multiple openings, 2015. URL https://www.google.ch/patents/US20150321025.

Hermann Fuchs, Markus Alber, Thomas Schreiner, and Dietmar Georg. Implementation of spot scanning dose optimization and dose calculation for helium ions in Hyperion. *Medical Physics*, 42(9): 5157–5166, 2015. ISSN 0094-2405. doi: 10.1118/1.4927789. URL http://www.ncbi.nlm.nih.gov/pubmed/26328967.

Christian Golnik, Fernando Hueso-González, Andreas Müller, Peter Dendooven, Wolfgang Enghardt, Fine Fiedler, Thomas Kormoll, Katja Roemer, Johannes Petzoldt, Andreas Wagner, and Guntram Pausch. Range assessment in particle therapy based on prompt γ-ray timing measurements. *Physics in Medicine and Biology*, 59(18):5399–5422, sep 2014. ISSN 0031-9155. doi: 10.1088/0031-9155/59/18/5399. URL http://stacks.iop.org/0031-9155/59/i=18/a=5399?key=crossref.5437fcd3059992135ec2113679c7dad6.

G Goodman, L Skarsgard, G Thompson, R Harrison, G Lam, and C Lugate. Pion therapy at TRIUMF. Treatment results for astrocytoma grades 3 and 4: a pilot study. *Radiotherapy and Oncology*, 17(1): 21–28, jan 1990. ISSN 01678140. doi: 10.1016/0167-8140(90)90045-X. URL http://linkinghub.elsevier.com/retrieve/pii/016781409090045X.

L Grevillot, M Stock, and S Vatnitsky. Evaluation of beam delivery and ripple filter design for non-isocentric proton and carbon ion therapy. *Physics in medicine and biology*, 60(20):7985–8005, 2015. ISSN 1361-6560. doi: 10.1088/0031-9155/60/20/7985. URL http://www.ncbi.nlm.nih.gov/pubmed/26418366.

P Gueth, D Dauvergne, N Freud, J M Létang, C Ray, E Testa, and D Sarrut. Machine learning-based patient specific prompt-gamma dose monitoring in proton therapy. *Physics in medicine and biology*, 58(13):4563–77, jul 2013. ISSN 1361-6560. doi: 10.1088/0031-9155/58/13/4563. URL http://www.ncbi.nlm.nih.gov/pubmed/23771015.

K Gunzert-Marx, H Iwase, D Schardt, and R S Simon. Secondary beam fragments produced by 200 MeV/u 12C ions in water and their dose contributions in carbon ion radiotherapy. *New Journal of Physics*, 10(7):075003, jul 2008. ISSN 1367-2630. doi: 10.1088/1367-2630/10/7/075003. URL http://stacks.iop.org/1367-2630/10/i=7/a=075003?key=crossref.d8c5fe2733dc0eb15571a822575b9130.

K M Hanson, J N Bradbury, T M Cannon, R L Hutson, D B Laubacher, R J Macek, M a Paciotti, and C a Taylor. Computed tomography using proton energy loss. *Physics in medicine and biology*, 26(6): 965–83, nov 1981. ISSN 0031-9155. URL http://www.ncbi.nlm.nih.gov/pubmed/6275424.

J Hartman, C Kontaxis, G H Bol, S J Frank, J J W Lagendijk, M van Vulpen, and B W Raaymakers. Dosimetric feasibility of intensity modulated proton therapy in a transverse magnetic field of 1.5 T. *Physics in Medicine and Biology*, 60(15):5955–5969, 2015. ISSN 0031-9155. doi: 10.1088/0031-9155/60/15/5955. URL http://iopscience.iop.org/0031-9155/60/15/5955/article/%5Cnhttp://stacks.iop.org/0031-9155/60/i=15/a=5955?key=crossref.133918021ca6932f1d7d19cea52617c3.

Kaspar Haume, Soraia Rosa, Sophie Grellet, Malgorzata Smialek, Karl Butterworth, Andrey Solov'yov, Kevin Prise, Jon Golding, and Nigel Mason. Gold nanoparticles for cancer radiotherapy: a review. *Cancer Nano*, 7(8), 2016. ISSN 1868-6958. doi: 10.1186/s12645-016-0021-x.

P Henriquet, E Testa, M Chevallier, D Dauvergne, G Dedes, N Freud, J Krimmer, J M Létang, C Ray, M-H Richard, and F Sauli. Interaction vertex imaging (IVI) for carbon ion therapy monitoring: a feasibility study. *Physics in medicine and biology*, 57(14):4655–69, jul 2012. ISSN 1361-6560. doi: 10.1088/0031-9155/57/14/4655. URL http://www.ncbi.nlm.nih.gov/pubmed/22750688.

J. Hopfgartner. Light Ion beam therapy, jan 2008. ISSN 18159346.

Nace Hudobivnik, Florian Schwarz, Thorsten Johnson, Linda Agolli, George Dedes, Thomas Tessonnier, Frank Verhaegen, Christian Thieke, Claus Belka, Wieland H. Sommer, Katia Parodi, and Guillaume Landry. Comparison of proton therapy treatment planning for head tumors with a pencil beam algorithm on dual and single energy CT images. *Medical Physics*, 43(1):495–504, 2016. ISSN 0094-2405. doi: 10.1118/1.4939106. URL http://scitation.aip.org/content/aapm/journal/medphys/43/1/10.1118/1.4939106.

Fernando Hueso-González, Fine Fiedler, Christian Golnik, Thomas Kormoll, Guntram Pausch, Johannes Petzoldt, Katja E. Römer, and Wolfgang Enghardt. Compton Camera and Prompt Gamma Ray Timing: Two Methods for In Vivo Range Assessment in Proton Therapy. *Frontiers in Oncology*, 6(April):1–13, 2016. ISSN 2234-943X. doi: 10.3389/fonc.2016.00080. URL http://journal.frontiersin.org/article/10.3389/fonc.2016.00080.

Brent F B Huisman, J M Létang, É Testa, and D Sarrut. Accelerated Prompt Gamma estimation for clinical Proton Therapy simulations. *Physics in Medicine and Biology*, 61:7725–7743, 2016. ISSN 0031-9155. doi: 10.1088/0031-9155/61/21/7725.

T Inaniwa and N Kanematsu. A trichrome beam model for biological dose calculation in scanned carbon-ion radiotherapy treatment planning. *Physics in medicine and biology*, 60(1):437–451, 2014. ISSN 1361-6560. doi: 10.1088/0031-9155/60/1/437. URL http://www.ncbi.nlm.nih.gov/pubmed/25517427.

The International Agency for Research on Cancer. *World Cancer Report 2014 (International Agency for Research on Cancer)*. World Health Organization, 1 edition, 2014. ISBN 9789283204299. URL http://amazon.com/o/ASIN/9283204298/.

F M F C Janssen, G Landry, P Cambraia Lopes, G Dedes, J Smeets, D R Schaart, K Parodi, and F Verhaegen. Factors influencing the accuracy of beam range estimation in proton therapy using prompt gamma emission. *Physics in medicine and biology*, 59(15):4427–41, aug 2014. ISSN 1361-6560. doi: 10.1088/0031-9155/59/15/4427. URL http://www.ncbi.nlm.nih.gov/pubmed/25049223.

Martin Jermann. Particle Therapy Statistics in 2014. *International Journal of Particle Therapy*, 2(1):50–54, 2015. doi: 10.14338/IJPT-15-00013. URL http://dx.doi.org/10.14338/IJPT-15-00013.

Masataka Karube, Naoyoshi Yamamoto, Mio Nakajima, Hideomi Yamashita, Keiichi Nakagawa, Tadaaki Miyamoto, Hiroshi Tsuji, Takehiko Fujisawa, and Tadashi Kamada. Single-Fraction Carbon-Ion Radiation Therapy for Patients 80 Years of Age and Older With Stage I Non-Small Cell Lung Cancer. *International journal of radiation oncology, biology, physics*, 95(1):542–548, 2016. ISSN 1879355X. doi: 10.1016/j.ijrobp.2015.11.034. URL http://dx.doi.org/10.1016/j.ijrobp.2015.11.034.

Soo Mee Kim, Hee Seo, Jin Hyung Park, Chan Hyeong Kim, Chun Sik Lee, Soo-Jin Lee, Dong Soo Lee, and Jae Sung Lee. Resolution recovery reconstruction for a Compton camera. *Physics in Medicine and Biology*, 58(9):2823, 2013. URL http://stacks.iop.org/0031-9155/58/i=9/a=2823.

Young-Hak Kim, Changyeon Yoon, and Wonho Lee. Monitoring 3D dose distributions in proton therapy by reconstruction using an iterative method. *Applied Radiation and Isotopes*, 114:33–39, 2016. ISSN 09698043. doi: 10.1016/j.apradiso.2016.05.005. URL http://linkinghub.elsevier.com/retrieve/pii/S0969804316301683.

Antje-Christin Knopf and Antony Lomax. In vivo proton range verification: a review. *Physics in medicine and biology*, 58(15):R131–60, aug 2013. ISSN 1361-6560. doi: 10.1088/0031-9155/58/15/R131. URL http://www.ncbi.nlm.nih.gov/pubmed/23863203.

A.M. Koehler. Proton radiography. *Science*, 160(3825):303–4, apr 1968. ISSN 0036-8075. doi: 10.1126/science.160.3825.303. URL http://www.ncbi.nlm.nih.gov/pubmed/17788234.

G Kraft. Tumor therapy with heavy charged particles. *Progress in Particle and Nuclear Physics*, 45, 2000. URL http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Progress+in+Particle+and+Nuclear+Physics#0http://www.sciencedirect.com/science/article/pii/S0146641000001125.

N Krah, M Testa, and I Rinaldi. Improved patient-specific optimization of the stopping power calibration for proton therapy planning using a single optimized proton radiography. In *ICTR-PHE 2016*, Genève, Switzerland, 2016. URL https://hal.archives-ouvertes.fr/hal-01276427.

Michael Krämer, Emanuele Scifoni, Christoph Schuy, Marta Rovituso, Walter Tinganelli, Andreas Maier, Robert Kaderka, Wilma Kraft-Weyrather, Stephan Brons, Thomas Tessonnier, Katia Parodi, and Marco Durante. Helium ions for radiotherapy? Physical and biological verifications of a novel treatment modality. *Medical Physics*, 43(4):1995–2004, 2016. ISSN 0094-2405. doi: 10.1118/1.4944593. URL http://scitation.aip.org/content/aapm/journal/medphys/43/4/10.1118/1.4944593.

Jochen Krimmer. Development of a Compton camera for medical applications based on silicon strip and scintillation detectors. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, pages 98–101, 2015.

J. J. W. Lagendijk, M. van Vulpen, and B. W. Raaymakers. The development of the MRI linac system for online MRI-guided radiotherapy: a clinical update. *Journal of Internal Medicine*, pages 1–6, 2016. ISSN 09546820. doi: 10.1111/joim.12516. URL http://doi.wiley.com/10.1111/joim.12516.

William R Leo. *Techniques for Nuclear and Particle Physics Experiments*. 2nd edition, 1994. ISBN 0-387-57280-5.

Ying Liang, Qibin Fu, Xudong Wang, Feng Liu, Gen Yang, Chunxiong Luo, Qi Ouyang, and Yugang Wang. Relative biological effectiveness for photons: implication of complex DNA double-strand breaks as critical lesions. *Physics in Medicine and Biology*, at press, 2017.

Hsin-hon Lin, Hao-ting Chang, Tsi-chian Chao, and Keh-shih Chuang. A comparison of two prompt gamma imaging techniques with collimator-based cameras for range verification in proton therapy. *Radiation Physics and Chemistry*, pages 1–7, 2016. ISSN 0969-806X. doi: 10.1016/j.radphyschem.2016.04.020. URL http://dx.doi.org/10.1016/j.radphyschem.2016.04.020.

Gabriela Llosá, Marco Trovato, John Barrio, Ane Etxebeste, Enrique Muñoz, Carlos Lacasta, Josep F. Oliver, Magdalena Rafecas, Carles Solaz, and Paola Solevi. First Images of a Three-Layer Compton Telescope Prototype for Treatment Monitoring in Hadron Therapy. *Frontiers in Oncology*, 6(February):14, 2016. ISSN 2234-943X. doi: 10.3389/fonc.2016.00014. URL http://journal.frontiersin.org/Article/10.3389/fonc.2016.00014/abstract.

Jay S Loeffler and Marco Durante. Charged particle therapy–optimization, challenges and future directions. *Nature reviews. Clinical oncology*, 10(7):411–424, 2013. ISSN 1759-4782. doi: 10.1038/nrclinonc.2013.79. URL http://www.ncbi.nlm.nih.gov/pubmed/23689752.

Patricia Cambraia Lopes, Enrico Clementel, Paulo Crespo, Sebastien Henrotin, Jan Huizenga, Guillaume Janssens, Katia Parodi, Damien Prieels, Frauke Roellinghoff, Julien Smeets, Frederic Stichelbaut, and Dennis R Schaart. Time-resolved imaging of prompt-gamma rays for proton range verification using a knife-edge slit camera based on digital photon counters. *Physics in Medicine and Biology*, 60(15):6063, 2015. ISSN 0031-9155. doi: 10.1088/0031-9155/60/15/6063. URL http://iopscience.iop.org/0031-9155/60/15/6063/article/.

P. Mayles, a. Nahum, J. C. Rosenwald, and Niko Papanikolaou. *Handbook of Radiotherapy Physics: Theory and Practice*, volume 35. 2008. ISBN 0750308605. doi: 10.1118/1.2969650.

S E McGowan, F Albertini, S J Thomas, and a J Lomax. Defining robustness protocols: a method to include and evaluate robustness in clinical plans. *Physics in medicine and biology*, 60(7):2671–84, 2015. ISSN 1361-6560. doi: 10.1088/0031-9155/60/7/2671. URL http://www.ncbi.nlm.nih.gov/pubmed/25768095.

Chul Hee Min, Han Rim Lee, Chan Hyeong Kim, and Se Byeong Lee. Development of array-type prompt gamma measurement system for in vivo range verification in proton therapy. *Medical physics*, 39(4):2100–7, apr 2012. ISSN 0094-2405. doi: 10.1118/1.3694098. URL http://www.ncbi.nlm.nih.gov/pubmed/22482631.

Silvia Molinelli, Giuseppe Magro, Andrea Mairani, Naruhiro Matsufuji, Nobuyuki Kanematsu, Taku Inaniwa, Alfredo Mirandola, Stefania Russo, Edoardo Mastella, Azusa Hasegawa, Hiroshi Tsuji, Shigeru Yamada, Barbara Vischioni, Viviana Vitolo, Alfredo Ferrari, Mario Ciocca, Tadashi Kamada, Hirohiko Tsujii, Roberto Orecchia, and Piero Fossati. Dose prescription in carbon ion radiotherapy: How to compare two different RBE-weighted dose calculation systems. *Radiotherapy and Oncology*, 2016. ISSN 18790887. doi: 10.1016/j.radonc.2016.05.031. URL http://dx.doi.org/10.1016/j.radonc.2016.05.031.

M Moteabbed, S España, and H Paganetti. Monte Carlo patient study on the comparison of prompt gamma and PET imaging for range verification in proton therapy. *Physics in medicine and biology*, 56(4):1063–82, feb 2011. ISSN 1361-6560. doi: 10.1088/0031-9155/56/4/012. URL http://www.ncbi.nlm.nih.gov/pubmed/21263174.

P G Ortega, I Torres-Espallardo, F Cerutti, a Ferrari, J E Gillam, C Lacasta, G Llosá, J F Oliver, P R Sala, P Solevi, and M Rafecas. Noise evaluation of Compton camera imaging for proton therapy. *Physics in medicine and biology*, 60(5):1845–63, 2015. ISSN 1361-6560. doi: 10.1088/0031-9155/60/5/1845. URL http://www.ncbi.nlm.nih.gov/pubmed/25658644.

Harald Paganetti. Range uncertainties in proton therapy and the role of Monte Carlo simulations. *Physics in Medicine and Biology*, 57(11):R99–R117, 2012. ISSN 0031-9155. doi: 10.1088/0031-9155/57/11/R99. URL http://stacks.iop.org/0031-9155/57/i=11/a=R99?key=crossref.4b6c83cb5125b4a360ead936994fceb4.

Harald Paganetti. *Proton Beam Therapy*. IOP Publishing, 2016. ISBN 978-0-7503-1370-4. doi: 10.1088/978-0-7503-1370-4. URL http://iopscience.iop.org/book/978-0-7503-1370-4.

Harald Paganetti, Andrzej Niemierko, Marek Ancukiewicz, Leo E Gerweck, Michael Goitein, Jay S Loeffler, and Herman D Suit. Relative biological effectiveness (RBE) values for proton beam therapy. *International Journal of Radiation Oncology*Biology*Physics*, 53(2):407–421, 2002. ISSN 03603016. doi: 10.1016/S0360-3016(02)02754-2.

Seyoun Park, William Plishker, Harry Quon, John Wong, Raj Shekhar, and Junghoon Lee. Deformable registration of CT and cone-beam CT with local intensity matching. *Physics in Medicine and Biology*, 62(3):927–947, feb 2017. ISSN 0031-9155. doi: 10.1088/1361-6560/aa4f6d. URL http://stacks.iop.org/0031-9155/62/i=3/a=927?key=crossref.6d51a84e851cc257d939edf081764a29.

Katia Parodi, Falk Pönisch, and Wolfgang Enghardt. Experimental study on the feasibility of in-beam PET for accurate monitoring of proton therapy. *IEEE Transactions on Nuclear Science*, 52(C):778–786, 2005. ISSN 00189499. doi: 10.1109/TNS.2005.850950.

I Perali, a Celani, L Bombelli, C Fiorini, F Camera, E Clementel, S Henrotin, G Janssens, D Prieels, F Roellinghoff, J Smeets, F Stichelbaut, and F Vander Stappen. Prompt gamma imaging of proton pencil beams at clinical dose rate. *Physics in Medicine and Biology*, 59(19):5849–5871, oct 2014. ISSN 0031-9155. doi: 10.1088/0031-9155/59/19/5849. URL http://stacks.iop.org/0031-9155/59/i=19/a=5849?key=crossref.d1c598721e0b970b36f1c8a6ad1dd1a1.

M Pinto, D Dauvergne, N Freud, J Krimmer, J M Letang, C Ray, F Roellinghoff, and E Testa. Design optimisation of a TOF-based collimated camera prototype for online hadrontherapy monitoring. *Physics in medicine and biology*, 59(24):7653–7674, 2014. ISSN 1361-6560. doi: 10.1088/0031-9155/59/24/7653. URL http://www.ncbi.nlm.nih.gov/pubmed/25415207.

M Pinto, M Bajard, S Brons, M Chevallier, D Dauvergne, G Dedes, M De Rydt, N Freud, J Krimmer, C La Tessa, J M Létang, K Parodi, R Pleskač, D Prieels, C Ray, I Rinaldi, F Roellinghoff, D Schardt, E Testa, and M Testa. Absolute prompt-gamma yield measurements for ion beam therapy monitoring. *Physics in Medicine and Biology*, 60(2):565–594, 2015. ISSN 0031-9155. doi: 10.1088/0031-9155/60/2/565. URL http://stacks.iop.org/0031-9155/60/i=2/a=565?key=crossref.c4760640602341dfb0ce09b91884d4e4.

Marco Pinto, Denis Dauvergne, Nicolas Freud, Jochen Krimmer, Jean M. Létang, and Etienne Testa. Assessment of Geant4 Prompt-Gamma Emission Yields in the Context of Proton Therapy Monitoring. *Frontiers in Oncology*, 6(January):1–7, 2016. ISSN 2234-943X. doi: 10.3389/fonc.2016.00010. URL http://journal.frontiersin.org/Article/10.3389/fonc.2016.00010/abstract.

E. B. Podgorsak and Kristofer Kainz. Radiation Oncology Physics: A Handbook for Teachers and Students. *Medical Physics*, 33(6):1920, 2006. ISSN 00942405. doi: 10.1118/1.2201870. URL http://scitation.aip.org/content/aapm/journal/medphys/33/6/10.1118/1.2201870.

Jerimy C Polf, Stephen Avery, Dennis S Mackin, and Sam Beddar. Imaging of prompt gamma rays emitted during delivery of clinical proton beams with a Compton camera: feasibility studies for range verification. *Phys Med Biol*, 60(18):7085–7099, 2015. ISSN 0031-9155. doi: 10.1088/0031-9155/60/18/7085. URL http://dx.doi.org/10.1088/0031-9155/60/18/7085.

J. P. Pouget and S. J. Mather. General aspects of the cellular response to low- and high-LET radiation. *European Journal of Nuclear Medicine*, 28(4):541–561, 2001. ISSN 03406997. doi: 10.1007/s002590100484.

M Priegnitz, S Helmbrecht, G Janssens, and I Perali. Measurement of prompt gamma profiles in inhomogeneous targets with a slit camera. *Physics in Medicine and Biology*, 4849:4849, 2015. ISSN 0031-9155. doi: 10.1088/0031-9155/60/12/4849. URL http://dx.doi.org/10.1088/0031-9155/60/12/4849.

C T Quiñones, J M Létang, and S Rit. Filtered back-projection reconstruction for attenuation proton CT along most likely paths. *Physics in Medicine and Biology*, 61(9):3258–3278, 2016. ISSN 0031-9155. doi: 10.1088/0031-9155/61/9/3258. URL http://stacks.iop.org/0031-9155/61/i=9/a=3258?key=crossref.000ce6a57c97a1107276b548832c2813.

Paul Retif, Sophie Pinel, Magali Toussaint, Céline Frochot, Rima Chouikrat, Thierry Bastogne, and Muriel Barberi-Heyob. Nanoparticles for radiation therapy enhancement: The key parameters. *Theranostics*, 5(9):1030–1044, 2015. ISSN 18387640. doi: 10.7150/thno.11642.

M.-H. Richard, M. Chevallier, D. Dauvergne, N. Freud, P. Henriquet, F. Le Foulher, J. M. Létang, G. Montarou, C. Ray, F. Roellinghoff, E. Testa, M. Testa, and A. H. Walenta. Design Guidelines for a Double Scattering Compton Camera for Prompt-γ Imaging During Ion Beam Therapy: A Monte Carlo Simulation Study. *IEEE Transactions on Nuclear Science*, 58(1):87–94, 2011. ISSN 0018-9499. doi: 10.1109/TNS.2010.2076303.

M. H. Richard, M. Dahoumane, D. Dauvergne, M. De Rydt, G. Dedes, N. Freud, J. Krimmer, J. M. Letang, X. Lojacono, V. Maxim, G. Montarou, C. Ray, F. Roellinghoff, E. Testa, and A. H. Walenta. Design study of the absorber detector of a compton camera for on-line control in ion beam therapy. *IEEE Transactions on Nuclear Science*, 59(5 PART 1):1850–1855, 2012. ISSN 00189499. doi: 10.1109/TNS.2012.2206053.

Christian Richter, Guntram Pausch, Steffen Barczyk, Marlen Priegnitz, Isabell Keitz, Julia Thiele, Julien Smeets, Francois Vander Stappen, Luca Bombelli, Carlo Fiorini, Lucian Hotoiu, Irene Perali, Damien Prieels, Wolfgang Enghardt, and Michael Baumann. First clinical application of a prompt gamma based in vivo proton range verification system. *Radiotherapy and Oncology*, 118 (2):232–237, 2016. ISSN 18790887. doi: 10.1016/j.radonc.2016.01.004. URL http://linkinghub.elsevier.com/retrieve/pii/S0167814016000074.

Eike Rietzel, Dieter Schardt, and Thomas Haberer. Range accuracy in carbon ion treatment planning based on CT-calibration with real tissue samples. *Radiation oncology (London, England)*, 2:14, jan 2007. ISSN 1748-717X. doi: 10.1186/1748-717X-2-14. URL http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1858700&tool=pmcentrez&rendertype=abstract.

C Robert, G Dedes, G Battistoni, T T Böhlen, I Buvat, F Cerutti, M P W Chin, a Ferrari, P Gueth, C Kurz, L Lestand, a Mairani, G Montarou, R Nicolini, P G Ortega, K Parodi, Y Prezado, P R Sala, D Sarrut, and E Testa. Distributions of secondary particles in proton and carbon-ion therapy: a comparison between GATE/Geant4 and FLUKA Monte Carlo codes. *Physics in medicine and biology*, 58(9):2879–99, may 2013. ISSN 1361-6560. doi: 10.1088/0031-9155/58/9/2879. URL http://www.ncbi.nlm.nih.gov/pubmed/23571094.

F Roehlinghoff. *Design and Implementation of a Prompt-Gamma Camera for Real-Time Monitoring of Ion Beam Therapy*. PhD thesis, 2014.

F. Roellinghoff, M. H. Richard, M. Chevallier, J. Constanzo, D. Dauvergne, N. Freud, P. Henriquet, F. Le Foulher, J. M. Létang, G. Montarou, C. Ray, E. Testa, M. Testa, and a. H. Walenta. Design of a Compton camera for 3D prompt-γ imaging during ion beam therapy. *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 648(SUPPL. 1):S20–S23, aug 2011. ISSN 01689002. doi: 10.1016/j.nima.2011.01.069. URL http://linkinghub.elsevier.com/retrieve/pii/S0168900211001471.

F Roellinghoff, a Benilov, D Dauvergne, G Dedes, N Freud, G Janssens, J Krimmer, J M Létang, M Pinto, D Prieels, C Ray, J Smeets, F Stichelbaut, and E Testa. Real-time proton beam range monitoring by means of prompt-gamma detection with a collimated camera. *Physics in medicine and biology*, 59

(5):1327–38, 2014. ISSN 1361-6560. doi: 10.1088/0031-9155/59/5/1327. URL http://www.ncbi.nlm.nih.gov/pubmed/24556873.

Harald H. Rossi. Specification of Radiation Quality. *Radiation Research*, 10(5):522, may 1959. ISSN 00337587. doi: 10.2307/3570787. URL http://www.jstor.org/stable/3570787?origin=crossref.

S. Rossi. The status of CNAO. *The European Physical Journal Plus*, 126(8), aug 2011. ISSN 2190-5444. doi: 10.1140/epjp/i2011-11078-8. URL http://www.springerlink.com/index/10.1140/epjp/i2011-11078-8.

Roel A Rozendaal, Ben J Mijnheer, Olga Hamming-vrieze, Anton Mans, and Marcel Van Herk. Impact of daily anatomical changes on EPID-based in vivo dosimetry of VMAT treatments of head-and-neck cancer. *RADIOTHERAPY AND ONCOLOGY*, 2015. ISSN 0167-8140. doi: 10.1016/j.radonc.2015.05.020. URL http://dx.doi.org/10.1016/j.radonc.2015.05.020.

Hyungjoon Ryu, Eunsuk Song, Jaeki Lee, and Jongwon Kim. Density and spatial resolutions of proton radiography using a range modulation technique. *Physics in medicine and biology*, 53(19):5461–8, oct 2008. ISSN 0031-9155. doi: 10.1088/0031-9155/53/19/012. URL http://www.ncbi.nlm.nih.gov/pubmed/18765893.

David Sarrut, Manuel Bardiès, Nicolas Boussion, Nicolas Freud, Sébastien Jan, Jean-Michel Létang, George Loudos, Lydia Maigne, Sara Marcatili, Thibault Mauxion, Panagiotis Papadimitroulas, Yann Perrot, Uwe Pietrzyk, Charlotte Robert, Dennis R Schaart, Dimitris Visvikis, and Irène Buvat. A review of the use and potential of the GATE Monte Carlo simulation code for radiation therapy and dosimetry applications. *Medical Physics*, 41(6), 2014. doi: http://dx.doi.org/10.1118/1.4871617. URL http://scitation.aip.org/content/aapm/journal/medphys/41/6/10.1118/1.4871617.

Dieter Schardt and Thilo Elsässer. Heavy-ion tumor therapy: Physical and radiobiological benefits. *Reviews of Modern Physics*, 82(1):383–425, feb 2010. ISSN 0034-6861. doi: 10.1103/RevModPhys.82.383. URL http://link.aps.org/doi/10.1103/RevModPhys.82.383.

U Schneider, E. Pedroni, and Antony Lomax. The calibration of CT Hounsfield units for radiotherapy treatment planning. *Physics in medicine and biology*, 41(1):111–24, 1996. ISSN 0031-9155. doi: 10.1088/0031-9155/41/1/009. URL http://www.ncbi.nlm.nih.gov/pubmed/8685250.

A Schumann, M Priegnitz, S Schoene, W Enghardt, H Rohling, and F Fiedler. From prompt gamma distribution to dose: a novel approach combining an evolutionary algorithm and filtering based on Gaussian-powerlaw convolutions. *Physics in medicine and biology*, page 6919, 2016. ISSN 0031-9155. doi: 10.1088/0031-9155/61/19/6919. URL http://dx.doi.org/10.1088/0031-9155/61/19/6919.

S. M. Seltzer, D. T. Bartlett, D. T. Burns, G. Dietze, H.-G. Menzel, H. G. Paretzke, A. Wambersie, and J. Tada. Fundamental Quantities And Units For Ionizing Radiation (Revised) ICRU-report No 85. *Journal of the ICRU*, 11(1):1–35, 2011. doi: 10.1093/jicru/ndr012.

Eric Sheldon and Douglas M Van Patter. Compound Inelastic Nucleon and Gamma-Ray Angular Distributions for Even- and Odd-Mass Nuclei. *Rev. Mod. Phys.*, 38(1):143–186, jan 1966. doi: 10.1103/RevModPhys.38.143. URL http://link.aps.org/doi/10.1103/RevModPhys.38.143.

Xiaorong Shi, Xue Meng, Xindong Sun, Ligang Xing, and Jinming Yu. PET/CT imaging-guided dose painting in radiation therapy. *Cancer Letters*, 355(2):169–175, 2014. ISSN 18727980. doi: 10.1016/j.canlet.2014.07.042. URL http://dx.doi.org/10.1016/j.canlet.2014.07.042.

J. P. Seuntjens Shortt, W. Strydom, and K. R. Chapter 2 Dosimetric Principles , Quantities and Units. In *Radiation Oncology Physics handbook - IAEA*, volume 1, pages 1–66. 2006. ISBN 9201073046.

J Smeets, F Roellinghoff, D Prieels, F Stichelbaut, A Benilov, P Busca, C Fiorini, R Peloso, M Basilavecchia, T Frizzi, J C Dehaes, and A Dubus. Prompt gamma imaging with a slit camera for real-time range control in proton therapy. *Physics in medicine and biology*, 57(11):3371–405, 2012. ISSN 1361-6560. doi: 10.1088/0031-9155/57/11/3371. URL http://www.ncbi.nlm.nih.gov/pubmed/22572603.

Julien Smeets, Damien Prieels, Eugen-lucian Hotoiu, and Edmond Sterpin. Sensitivity study of prompt gamma imaging of scanned 2 beam proton therapy in heterogeneous anatomies. *RADIOTHERAPY AND ONCOLOGY*, pages 1–6, 2015. ISSN 01678140. doi: 10.1016/j.radonc.2015.11.002. URL http://dx.doi.org/10.1016/j.radonc.2015.11.002.

Paola Solevi, Enrique Muñoz, Carles Solaz, Marco Trovato, Peter Dendooven, John E Gillam, Carlos Lacasta, Josep F Oliver, Magdalena Rafecas, and Irene Torres-Espallardo. Performance of MACACO Compton telescope for ion-beam therapy monitoring: first test with proton beams. *Physics in Medicine and Biology*, 61(14):5149–5165, 2016. ISSN 0031-9155. doi: 10.1088/0031-9155/61/14/5149. URL http://stacks.iop.org/0031-9155/61/i=14/a=5149?key=crossref.6bf836448c34a4bc8348edeba28d35c4.

G Sportelli, N Belcari, N Camarlinghi, G a P Cirrone, G Cuttone, S Ferretti, a Kraan, J E Ortuño, F Romano, a Santos, K Straub, a Tramontana, a Del Guerra, and V Rosso. First full-beam PET acquisitions in proton therapy with a modular dual-head dedicated system. *Physics in medicine and biology*, 59(1):43–60, dec 2013. ISSN 1361-6560. doi: 10.1088/0031-9155/59/1/43. URL http://www.ncbi.nlm.nih.gov/pubmed/24321855.

E Sterpin, G Janssens, J Smeets, François Vander Stappen, D Prieels, Marlen Priegnitz, Irene Perali, and S Vynckier. Analytical computation of prompt gamma ray emission and detection for proton range verification. *Physics in Medicine and Biology*, 60(12):4915–4946, 2015. ISSN 0031-9155. doi: 10.1088/0031-9155/60/12/4915. URL http://stacks.iop.org/0031-9155/60/i=12/a=4915?key=crossref.aabd8815e135401a22d165e343a7bac4.

Mauro Testa, Joost M Verburg, Mark Rose, Chul Hee Min, Shikui Tang, El Hassane Bentefour, Harald Paganetti, and Hsiao-Ming Lu. Proton radiography and proton computed tomography based on time-resolved dose measurements. *Physics in Medicine and Biology*, 58(22):8215–8233, nov 2013. ISSN 0031-9155. doi: 10.1088/0031-9155/58/22/8215. URL http://stacks.iop.org/0031-9155/58/i=22/a=8215?key=crossref.9766466f3cf307946adae72478a2726d.

Francesco Tommasino and Marco Durante. Proton radiobiology. *Cancers*, 7(1):353–381, 2015. ISSN 20726694. doi: 10.3390/cancers7010353.

S.M. Valle, G. Battistoni, V. Patera, D. Pinci, A. Sarti, A. Sciubba, E. Spiriti, and M. Marafini. The MONDO project: A secondary neutron tracker detector for particle therapy. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 621(1-3):322–325, 2016. ISSN 01689002. doi: 10.1016/j.nima.2016.05.001. URL http://linkinghub.elsevier.com/retrieve/pii/S016890021630359X.

Wouter van Elmpt, Leah McDermott, Sebastiaan Nijsten, Markus Wendling, Philippe Lambin, and Ben Mijnheer. A literature review of electronic portal imaging for radiotherapy dosimetry. *Radiotherapy and Oncology*, 88(3):289–309, 2008. ISSN 01678140. doi: 10.1016/j.radonc.2008.07.008.

Catarina Veiga, Jailan Alshaikhi, Richard Amos, Ana Mónica Lourenço, Marc Modat, Sebastien Ourselin, Gary Royle, and Jamie R. McClelland. Cone-Beam Computed Tomography and Deformable Registration-Based "Dose of the Day" Calculations for Adaptive Proton Therapy. *International Journal of Particle Therapy*, 2(2):404–414, sep 2015. ISSN 2331-5180. doi: 10.14338/IJPT-14-00024.1. URL http://theijpt.org/doi/10.14338/IJPT-14-00024.1.

J Verburg and T Bortfeld. TU-FG-BRB-06: A Prompt Gamma-Ray Spectroscopy System for Clinical Studies of in Vivo Proton Range Verification. *Medical Physics*, 43(6):3757–3757, jun 2016. ISSN 00942405. doi: 10.1118/1.4957546. URL http://doi.wiley.com/10.1118/1.4957546.

J. M. Verburg, C. Grassberger, S. Dowdell, J. Schuemann, J. Seco, and H. Paganetti. Automated Monte Carlo Simulation of Proton Therapy Treatment Plans. *Technology in Cancer Research and Treatment*, 2015. ISSN 1533-0346. doi: 10.1177/1533034615614139. URL http://tct.sagepub.com/lookup/doi/10.1177/1533034615614139.

Joost M Verburg. *Reducing range uncertainty in proton therapy*. 2015. ISBN 9789038638492.

Joost M Verburg and Joao Seco. Proton range verification through prompt gamma-ray spectroscopy. *Physics in Medicine and Biology*, 59(23):7089–7106, 2014. ISSN 0031-9155. doi: 10.1088/0031-9155/59/23/7089. URL http://stacks.iop.org/0031-9155/59/i=23/a=7089?key=crossref.2c380dcfa8360cb6e6defe62dd49f24a.

Joost M Verburg, Helen a Shih, and Joao Seco. Simulation of prompt gamma-ray emission during proton radiotherapy. *Physics in Medicine and Biology*, 57(17):5459–5472, sep 2012. ISSN 0031-9155. doi: 10.1088/0031-9155/57/17/5459. URL http://www.ncbi.nlm.nih.gov/pubmed/22864267.

G. Vilches-Freixas, J.M. Létang, N. Ducros, and S. Rit. Dual-energy CT spectra optimization for proton treatment planning. *Proc. 4th Intl. Mtg. on image formation in X-ray CT*, 7:585–588, 2016.

R.R. Wilson. Radiological Use of Fast Protons. *Radiology*, 47:487, 1946. URL http://phys4.harvard.edu/$\sim$wilson/cyclotron/bobwilsn.html.

L Wisser. Pion treatment of prostate carcinoma at Paul Scherrer Institute (formerly Swiss Institute for Nuclear Research (SIN)) from 1983 to 1992. *Cancer radiothérapie : journal de la Société française de radiothérapie oncologique*, 8(2):88–94, apr 2004. ISSN 1278-3218. doi: 10.1016/j.canrad.2003.12.004. URL http://www.ncbi.nlm.nih.gov/pubmed/15063876.

Anne-Sophie Wozny, Gersende Alphonse, Priscillia Battiston-Montagne, Stéphanie Simonet, Delphine Poncet, Etienne Testa, Jean-Baptiste Guy, Chloé Rancoule, Nicolas Magné, Michael Beuve, and Claire Rodriguez-Lafrasse. Influence of Dose Rate on the Cellular Response to Low- and High-LET Radiations. *Frontiers in Oncology*, 6(March):1–7, 2016. ISSN 2234-943X. doi: 10.3389/fonc.2016.00058. URL http://journal.frontiersin.org/article/10.3389/fonc.2016.00058.

Di Yan. Adaptive Radiotherapy: Merging Principle Into Clinical Practice. *Seminars in Radiation Oncology*, 20(2):79–83, apr 2010. ISSN 10534296. doi: 10.1016/j.semradonc.2009.11.001. URL http://www.ncbi.nlm.nih.gov/pubmed/20219545.

M Yang, G Virshup, J Clayton, X R Zhu, R Mohan, and L Dong. Theoretical variance analysis of single- and dual-energy computed tomography methods for calculating proton stopping power ratios of biological tissues. *Physics in medicine and biology*, 55(5):1343–1362, 2010. ISSN 0031-9155. doi: 10.1088/0031-9155/55/5/006.

Pavel S. Yarmolenko, Eui Jung Moon, Chelsea Landon, Ashley Manzoor, Daryl W. Hochman, Benjamin L. Viglianti, and Mark W. Dewhirst. Thresholds for thermal damage to normal tissues: An update. *International Journal of Hyperthermia*, 27(4):320–343, jun 2011. ISSN 0265-6736. doi: 10.3109/02656736.2010.534527. URL http://linkinghub.elsevier.com/retrieve/pii/S0022202X15370834http://www.tandfonline.com/doi/full/10.3109/02656736.2010.534527.

Stephen S. F. Yip and Hugo J. W. L. Aerts. Applications and limitations of radiomics. *Physics in medicine and biology*, 61(13):R150–66, jul 2016. ISSN 1361-6560. doi: 10.1088/0031-9155/61/13/R150. URL http://stacks.iop.org/0031-9155/61/i=13/a=R150?key=crossref.134478778713970aff90f16abe110608%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/27269645%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4927328.