Algorithms and software tools for the reconstruction of nuclear data in hadrontherapy

Alexandre Sécher

Université de Strasbourg

Université de Strasbourg

École doctorale
Physique, chimie-physique
ED 182
Université de Strasbourg

École Doctorale de Physique et Chimie Physique
Institut Pluridisciplinaire Hubert Curient (IPHC), UMR 7178

Thèse

présentée par :

Alexandre Sécher

soutenue le : **13 Juin 2022**

pour obtenir le grade de: **Docteur de l'Université de Strasbourg**
Discipline/Spécialité : Physique nucléaire appliquée

Outils logiciels et algorithmiques de reconstruction pour les données nucléaires en hadronthérapie

---

Algorithms and software tools for the reconstruction of nuclear data in hadrontherapy

Thèse dirigée par :

M. Christian Finck	Institut Pluridisciplinaire Hubert Curien

Rapporteurs :

Mme Chiara La Tessa	Université de Trento
M. Christophe Schuy	GSI

Autres membres du jury :

M. Alessio Sarti	Université de Rome
M. Samuel Salvador	Laboratoire de Physique Corpusculaire
M. Jérôme Baudot	Institut Pluridisciplinaire Hubert Curien
Mme Marie Vanstalle	Institut Pluridisciplinaire Hubert Curien

Contents

Acknowledgments

What a ride it was. A few months short of four years. A thesis shadowed by the global pandemic, which ended with a potential worldwide war on the background. And here I am, writing the acknowledgments for the manuscript. First things first, I would like to apologize to anyone I will not mention here; it is hard to remember and thank appropriately every person who had an impact in my work or life during those long years.

The firsts of my thanks goes to Christian and Marie, my two supervisors, the boss, and the real boss. I know everything has not been pink all the way down, but I would like to remember the good moments we shared together, abroad and in the lab. Thanks to the two of you for your patience and dedication, for sharing some of the load of the thesis with me, for being there when I needed it. Christian, I would like to thank you for the freedom you allowed me to have during the thesis. Indeed, having the liberty to explore ideas and develop solutions as I envisioned them was priceless. Marie, I would like to thank you for being a good friend during those demanding times. Feeling accepted and welcome in a new place is not always obvious; you made that part easy, and I knew I could count on you when I needed it. Thanks for your humanity. Anyhow, both of you, please take care of yourself and each other, and also of the ones that will follow in my footsteps.

Next, I would like to thank deeply the members of the jury. First, Jerôme, always present, always a pleasure to be in your company. Christophe, with whom I would have liked to go into even more technical details; know that your input on the manuscript and defense were really appreciated. Chiara, it was a great pleasure to work with you, and to see people with a refreshing look on the field, trying to follow their vision. Samuel, of course, for your dedication and the questioning you raised all along this thesis, notably during the two mid-term defenses. And finally, the one and only, Alessio. I am very glad to have the opportunity to work with you during this thesis. FOOT will be (already is ?) missing your full presence at its heart.

I would also like to express a huge thank to Marc Rousseau, without whom I would probably not have gone through this thesis. I met you at a time where I was heavily questioning what I would do in the future, which branch of physics I would like to explore, and you made the decision quite easy.

I would also like to express special thanks to Nicolas Arbor, who was a member of the preliminary jury, during my two mid-term defenses. Thank you for your kindness and your good words, always knowing how to boost someone back into confidence zone.

Of course, I would like to thank all of the people I shared time with on a day-to-say basis, coming either from building 25 or 35 (or from elsewhere in the lab). So in no particular order: Virgile, Séverine, Nicolas C., Laurent, Antoine, Cédric, Ali, Michel, Quentin, Estelle, Eric, Jacky, Addil, Momo, Nicolas S., Stéphane, The-Duc.

Huge thanks to the FOOT collaboration as well, and notably to the people I mostly interacted with: Ilaria, Marco and Yun.

Needles to say, during this thesis, I also met quite a few people that made the same curious choice as I did: other PhD students. With some of them I shared an office, for quite some time, and I want to thank them for putting up with me everyday: Kajal, Truong and Florian. Then came the house move from building 25 to 35, and there, Clément had the chance to share his workspace with me. Good luck for the few months you have left, and know that your always kind words will be deeply missed.

During this thesis, I was lucky to have someone akin to a big sister, Claire-Anne. Huge thanks to you.

A huge thanks to my family as well, who helped me get my head above the water during tough times. A particular thanks to my brother and sisters, Ben, Romane and M'elissa.

Thanks to Susie also, without whom Jessica would probably not have been completed.

I would like to thank a few people with whom I grew close or closer during the span of the thesis. Dylan first: know that game nights (and midday also), were always good times, and helped a lot with the mental sides of things. Robin, good friend since the early years of master, always there when needed.

Finally, I would like to thank the members of Afterglow. Without them and the emotional support they provided during tough times, I am pretty sure none of this would have been possible. Loris, always mysterious, I hope you find something you can enjoy and strive in, in the future. And if that thing is spending time with Leonard, all the better ! Julie, what a journey ! From PhD student in physics, to law, to acting, to dealing (cards). The

apparent easiness with which you make decisions, your absence of fear at the idea of trying new things will always amaze me. Guillaume, the source of so much fascination and incomprehension at the same time. How you work inside will probably always be a mystery. And that is probably the reason why it works so well. None of our outings would have been the same without you there, even though you often do not remember much. Clément, friend from before, friend until after. I am glad I got to share this time with you, and I promise, it is almost over for you too. And Douja, obviously. Maybe you were the reason why I had to come to Strasbourg to do this thesis, I do not know. But I am oh so glad we found each other. You can now call me doctor, and that definitely would not have been possible without you there.

Introduction

Cancer is, together with cardiovascular diseases, one of the leading cause of premature death worldwide. From 12.7 million new cancer cases for 7.6 million deaths in 2008, almost 20 million new cases, for 10 millions deaths have been estimated around the world, in 2020. Predictions for 2040 put the number of new cases to 28.4 million [1]. Looking at those numbers, cancer is undeniably a global health concern.

Because of the large variations in possible cancer sites, or the progression of the disease in any patient, several treatments modalities have been developed through the years. Surgery, chemotherapy and radiotherapy are the principal ones. The first involves an invasive removal of the tumoral volumes found within the patient, the second corresponds to the administration of drugs aiming to target the tumors and stop their proliferation, and the latter aims to destroy the replication potential of the cancerous cells by exposing them to ionizing radiations.

So-called classical radiotherapy makes use of photons in order to irradiate the tumoral volume. R. Wilson proposed the alternative use of charged particles in 1946 [2], which lead to the development of the charged particle therapy field. Nowadays, more than a hundred centers around the world can deliver such treatments; most offer the use of protons, but carbon ions are also available, and the use of helium or oxygen ions is under development.

The rise of this treatment modality can be explained by the advantageous properties charged particle bring over photons: their energy deposition profile is peaked at the end of the path of the particle in the medium. Based on the energy of the beam, the depth of the peak can be modified, and therefore used to target tumoral volumes. This results in an increase in energy deposition in the cancerous tissues, while sparing healthy ones, with respect to conventional radiotherapy.

However, the relative limited spread of this technology can be explained by several factors: first, the installations required to produce such beams are expensive and require dedicated personnel, which results in costly treatments; second, charged particle are subject to nuclear reactions with the

nuclei from the medium they traverse; third, range monitoring of such a technique is difficult, because the primary particles from the beam do not escape the patient.

Nuclear reactions can lead to the production of fragments from either the target or the projectile. In both cases, the production of those fragments results in an alteration of the energy deposition anticipated for the treatment; it is therefore crucial to able to take into account such effects while planning treatment, a role performed in radiotherapy by a dedicated treatment planning software (TPS).

And indeed, an observable desire to move from analytical treatment plans to Monte-Carlo based ones, in order to handle the production of secondary fragments, has been flourishing over the last decade. However, the hadronic models used in those simulation codes are known to be inaccurate at the therapeutic energy range. Therefore, in order to improve those models, experimental results, under the form of fragmentation cross-sections, need to pursued.

To this day, an important lack of experimental data remains at energies of interest for hadrontherapy. As a consequence, the FOOT experiment was proposed in order to extend and complete the projectile fragmentation cross sections for helium, carbon and oxygen beams, and also to provide measurements for target fragmentation. The international collaboration fixed two goals concerning cross-sections: isotopic double differential cross-sections for projectile fragmentation should be achieved with an accuracy better than 5% and differential cross-section with respect to kinetic energy for target fragmentation maximum uncertainty should not exceed 10%.

During treatment, several sources can lead to treatment delivery errors: an inaccurate positioning of the patient, anatomical variations, movement of the patient, the resolution of imaging techniques, conversion between imaging modality and treatment planning input are all factors that results in uncertainties in the delivery of the beam. While those effects are usually taken into account in treatment planning through the use of safety margins, *in vivo* range verification is a important tool to complete treatment planning. However, because in charged particle therapy the beam does not exit the patient, monitoring can only be based on the secondary particles produced through nuclear reactions of the primary ions.

One of the principal method of range verification developed in the charged particle therapy context is based on the detection of the prompt gammas emitted by the beam while crossing the irradiated tissues. Indeed, it has been shown that the emission profile of those photons is closely related to the energy deposition profile of the beam. However, this method

requires a comparison with an expected signal in order to be used as a range verification method, and in consequence relies also on Monte-Carlo simulations, and more specifically on their hadronic models.

A better evaluation of the impact of secondary particles on treatment and proper monitoring of said treatment are two of the most crucial issues the charged particle therapy field faces, because without both of those, the benefits of the modality cannot be fully exploited.

This work is divided in four chapters. The first introduces charged particle therapy, through an overview of the state of the modality, as well as a summary of the underlying physical notions. The biological effects of charged particles are also reviewed. The chapter ends on the necessity to extend the available experimental measurements in order to improve the field further.

The second chapter is a description of the FOOT experiment, its rationale and setup, as well as opening towards the need of a global reconstruction algorithm in order to produce cross-section measurements. The third chapter concerns the work performed for the collaboration, under the form of the TOE algorithm.

The final chapter is focused on an experiment around prompt gammas statistical issue. It explores the use of unfolding algorithms to retrieve additional information from the acquired spectra, and as a mean to evaluate Monte-Carlo simulations performances.

Le cancer, ainsi que les maladies cardiovasculaires, est une des causes principales de mort prématurée au monde. En 2008, 12.7 millions de nouveaux cas et 7.6 millions de morts ont été estimés, un nombre qui s'élève en 2020 à presque 20 millions de nouveaux cas pour 10 millions de morts. Les prédictions pour 2040 portent le nombre de nouveaux cas à 28.4 millions. Ces chiffres appuient le fardeau que fais porter le cancer sur la santé publique à travers le monde.

À cause des multiples possibles types de cancers, ainsi que des différents stades que la maladie peut atteindre, plusieurs modalités de traitement ont été développées au fur et à mesure des années. La chirurgie, la chimiothérapie et la radiothérapie sont les principales d'entre elles. La première implique une procédure invasive visant à retirer les tumeurs du corps du patient ; la deuxième consiste en l'administration de médicaments ciblant la tumeur

afin de stopper la prolifération des cellules qui la compose ; finalement, la dernière vise à détruire les tissus cancéreux en les exposant a des radiations ionisantes.

La radiothérapie classique utilise des photons pour irradier le volume tumoral. R. Wilson, en 1946, proposa l'idée d'utiliser des faisceaux d'ions à la place, ce qui conduisit au développement du domaine de la hadronthérapie. Aujourd'hui, plus d'une centaine de centre autour du monde peuvent délivrer des traitements basés sur des faisceaux d'ions : la plupart utilise des faisceaux de protons, mais les ions carbones sont aussi disponibles, et l'utilisation d'ions héliums ou oxygènes est en développement.

L'essor de cette modalité de traitement peut s'expliquer grâce aux propriétés avantageuses dont les particules chargées disposent par rapport aux photons : le profil de leur dépôt d'énergie est maximal est maximal à la fin du parcours de la particule dans le milieu. En fonction de l'énergie du faisceau, la profondeur du pic est modifiée, et peut donc être utilisé pour cibler les volumes tumoraux. Il en résulte une augmentation du dépôt d'énergie dans les tissus cancéreux, tout en épargnant les tissus sains, par rapport à la radiothérapie conventionnelle.

Cependant, force est de constater que la diffusion de cette technologie est relativement limitée. Ceci peut s'expliquer par plusieurs facteurs : premièrement, les installations nécessaires pour produire de tels faisceaux sont coûteuses et nécessitent du personnel spécialisé, ce qui entraîne des traitements onéreux ; deuxièmement, les particules chargées du faisceau sont sujettes à des réactions nucléaires avec les noyaux du milieu qu'elles traversent ; troisièmement, le contrôle du traitement est difficile, car les particules primaires du faisceau ne s'échappent pas du patient.

Les réactions nucléaires peuvent conduire à la production de fragments, provenant de la cible ou du projectile. Dans les deux cas, la production de ces fragments entraîne une modification du dépôt d'énergie prévu pour le traitement ; il est donc crucial de pouvoir tenir compte de ces effets lors de la planification du traitement, un rôle joué en radiothérapie par un logiciel dédié de planification du traitement (TPS).

Au cours de la dernière décennie, le domaine a vu naître une volonté de s'affranchir des plans de traitement analytiques au profit de plans basés sur des simulations Monte-Carlo, afin d'accomplir cet objectif. Cependant, les modèles hadroniques utilisés dans ces codes de simulation sont connus pour être imprécis dans la gamme d'énergie thérapeutique. Afin d'améliorer ces modèles, des mesures expérimentales de sections efficaces de fragmentation, doivent être obtenues.

À ce jour, un manque important de données expérimentales subsiste

aux énergies d'intérêt pour l'hadronthérapie. En conséquence, l'expérience FOOT a été proposée afin d'étendre et de compléter les sections efficaces de fragmentation des projectiles pour les faisceaux d'hélium, de carbone et d'oxygène, et de fournir des mesures de fragmentation des cibles. La collaboration internationale a fixé deux objectifs concernant ces mesures : les sections efficaces doublement différentielles isotopiques pour la fragmentation des projectiles doivent être obtenues avec une précision supérieure à 5 % et les sections efficaces simplement différentielles, par rapport à l'énergie cinétique, pour la fragmentation de la cible, ne doit pas dépasser une incertitude maximale de 10 %.

Pendant le traitement, plusieurs sources peuvent entraîner des erreurs dans la délivrance du traitement : un positionnement inexact du patient, des variations anatomiques, le mouvement du patient, la résolution des techniques d'imagerie, ou même la conversion entre les unités utilisées en imagerie vers celles importantes pour la planification, sont autant de facteurs qui entraînent des incertitudes dans la délivrance du faisceau. Bien que ces effets soient généralement pris en compte lors de la planification du traitement par l'utilisation de marges autour des volumes ciblés, le contrôle du traitement in vivo est un outil important pour compléter la planification. Cependant, étant donné qu'en thérapie par particules chargées, les particules primaires du faisceau ne sortent pas du patient, le contrôle ne peut être basé que sur les particules secondaires produites par les réactions nucléaires des ions primaires.

L'une des principales méthodes de contrôle développée dans le contexte de hadronthérapie est basée sur la détection des gammas prompts émis par le faisceau lors de la traversée des tissus irradiés. En effet, il a été démontré que le profil d'émission de ces photons est étroitement lié au profil de dépôt d'énergie du faisceau. Toutefois, cette méthode nécessite une comparaison avec un signal attendu afin d'être utilisée comme méthode de contrôle, et par conséquent repose également sur des simulations Monte-Carlo, et donc sur leurs modèles hadroniques.

Une meilleure évaluation de l'impact des particules secondaires sur le traitement et un suivi adéquat de ce traitement sont deux des questions les plus cruciales auxquelles le domaine de la thérapie par particules chargées est aujourd'hui confronté, car sans ces deux éléments, les avantages de cette modalité ne peuvent être pleinement exploités.

Cet ouvrage est divisé en quatre chapitres. Le premier présente la thérapie par particules chargées, ainsi qu'un résumé des notions physiques sous-jacentes. Les effets biologiques des particules chargées sont également

passés en revue. Le chapitre se termine sur la nécessité d'étendre les mesures expérimentales disponibles afin d'améliorer encore le domaine.

Le deuxième chapitre est une description de l'expérience FOOT. Le besoin d'utiliser un algorithme de reconstruction globale afin de produire des mesures de sections efficaces y est finalement évoqué. Le troisième chapitre concerne le travail effectué pour la collaboration, sous la forme de l'algorithme TOE.

Le dernier chapitre est centré sur une expérience autour des gammas prompts. L'utilisation d'algorithmes de déconvolution pour extraire des informations supplémentaires à partir des spectres acquis, et comme moyen d'évaluer les performances des simulations de Monte-Carlo, y est exploré.

CHAPTER 1

Introduction to hadrontherapy

Contents

1.1 Cancer and hadrontherapy

During the past century, improvements in public hygiene and the development of vaccines and antibiotics have lead to a sharp decrease in mortality from infectious diseases worldwide. In consequence, the prevalence of cancer as global health burden is now apparent. Indeed, it is one of the leading cause of premature death in most countries, together with cardiovascular diseases [3]. Worldwide maps of estimated age-standardized incidence rate and corresponding mortality rate can be found in figure 1.1. In 2020, there were an estimated 19.3 millions new cases of cancer overall, and around 10 million deaths due to it [1]. In table 1.1, cancer incidence and number of death per site is reported. According to [1] 28.4 million new cases are to be expected by 2040.

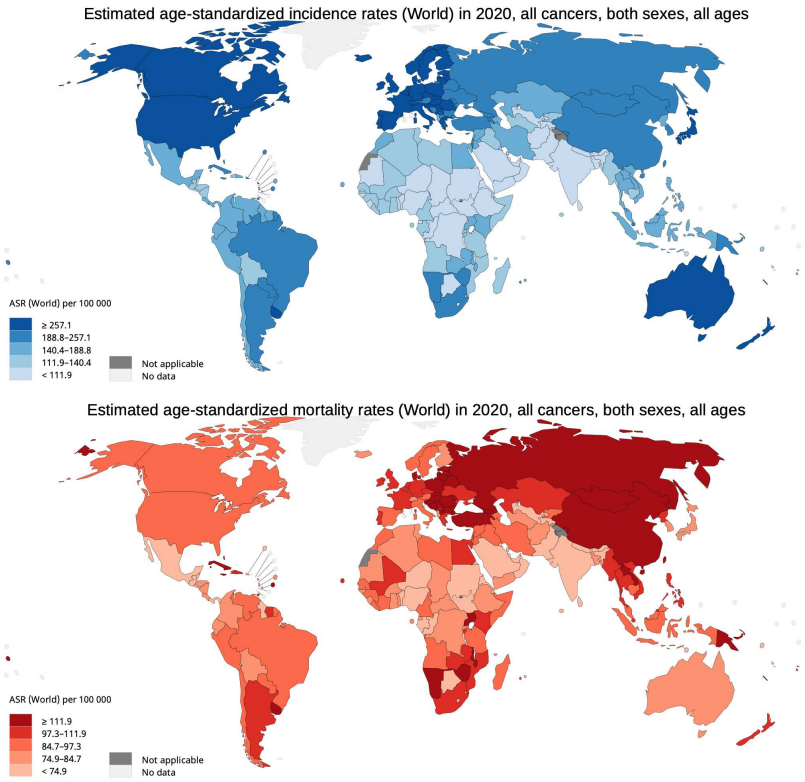

Figure 1.1: Cancer incidence and mortality by country reported as age standardized rates per 100000 inhabitants. Produced by [4].

The strain cancer place on the world can only be relieved through proper prevention and treatment. While the former is a major challenge for the

Cancer site	Cancer incidence (% of all sites)		Cancer death (% of all sites)	
Female breast	2261419	(11.7)	684996	(6.9)
Lung	2206771	(11.4)	1796144	(18.0)
Prostate	1414259	(7.3)	375304	(3.8)
Nonmelanoma of skin	1198073	(6.2)	63731	(0.6)
Colon	1148515	(6.0)	576858	(5.8)
Stomach	1089109	(5.6)	768793	(7.7)
Liver	905677	(4.7)	830180	(8.3)
Rectum	732210	(3.8)	339022	(3.4)
Cervix uteri	604127	(3.1)	341831	(3.4)
Esophagus	604100	(3.1)	544076	(5.5)
Thyroid	586202	(3.0)	43646	(0.4)
Bladder	573278	(3.0)	212536	(2.1)
Non-Hodgkin lymphoma	544352	(2.8)	259793	(2.6)
Pancreas	495773	(2.6)	466.003	(4.7)
Leukemia	474519	(2.5)	311594	(3.1)
Kidney	431288	(2.2)	179368	(1.8)
Corpus uteri	417367	(2.2)	97.370	(1.0)
Lip, oral cavity	377713	(2.0)	177757	(1.8)
Melanoma of skin	324635	(1.7)	57.043	(0.6)
Ovary	313959	(1.6)	207252	(2.1)
Brain, nervous system	308102	(1.6)	251329	(2.5)
Larynx	184615	(1.0)	99.840	(1.0)
Multiple myeloma	176404	(0.9)	117077	(1.2)
Nasopharynx	133354	(0.7)	80008	(0.8)
Gallbladder	115949	(0.6)	84695	(0.9)
Oropharynx	98412	(0.5)	48143	(0.5)
Hypopharynx	84254	(0.4)	38599	(0.4)
Hodgkin lymphoma	83087	(0.4)	23376	(0.2)
Testis	74458	(0.4)	9334	(0.1)
Salivary glands	53583	(0.3)	22778	(0.2)
Anus	50865	(0.3)	19293	(0.2)
Vulva	45240	(0.2)	17427	(0.2)
Penis	36068	(0.2)	13211	(0.1)
Kaposi sarcoma	34270	(0.2)	15086	(0.2)
Mesothelioma	30870	(0.2)	26278	(0.3)
Vagina	17908	(0.1)	7995	(0.1)
All sites	19292789		9958133	

Table 1.1: Cancer incidence and mortality per sites for the year 2020, extracted from [1].

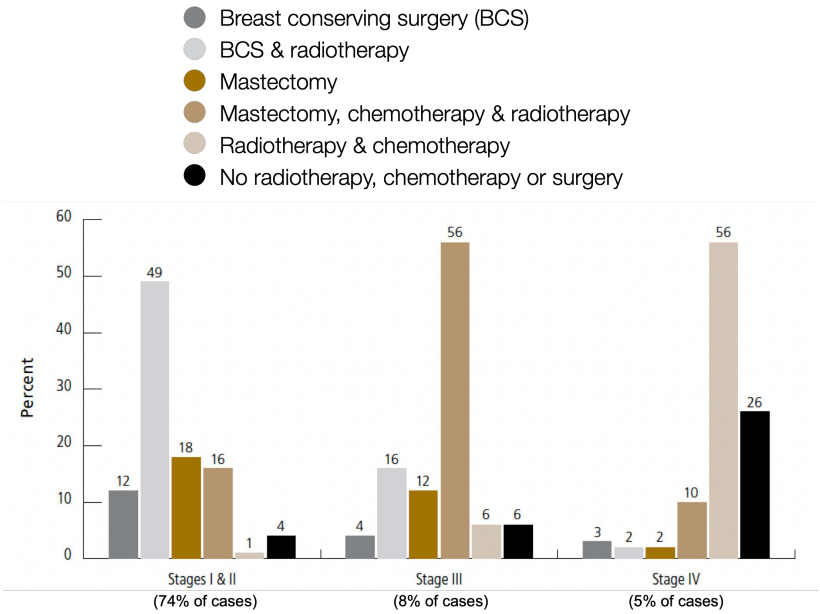

Figure 1.2: Treatment distribution in percentage of cases as a function of stage for breast cancer in 2016 in the U.S. Stage I corresponds to early invasive cancer, while stage IV is the most advanced stage. Percentage of all cases is indicated below each stage. Adapted from [5].

years ahead, a specific portion of the latter is the core of the work presented in this thesis, and will therefore be developed from here on.

Nowadays, cancer treatment usually relies on three main axes, often used in combination: surgery, chemotherapy and radiotherapy. Additionally, immunotherapy, which leverages the host immune system to remove tumoral cells, has seen important developments in the recent years. How a treatment modality is selected is highly dependent on the particular type of cancer as well as the stage it is currently in. For example, the distribution of treatment combinations for female breast cancer in the U.S. in 2016 are reported as a function of the cancer stage in figure 1.2, together with the corresponding stage distribution, in percentage of cases [5].

The work presented here is part of the field of radiotherapy, and more specifically, hadrontherapy. While classical radiotherapy is based on the use of photons (x-rays), hadrontherapy relies on charged particles as a mean to irradiate tumors, mainly using protons and carbon ions. In order to understand the advantages charged particles can bring over conventional radiotherapy, two clinical quantities resulting from an irradiation should be considered: the probability to sterilize a localized tumor volume or tumor control probability (TCP) and normal tissue complication probability

(NTCP). NTCP corresponds to the probability of inducing early or late morbidity because of the treatment, and is correlated to the exposition of healthy tissues by the irradiation.

Charged particle therapy, due to its advantageous ballistic properties section 1.2, should in principle be able to produce one of two outcomes, compared to classical radiotherapy: either a decrease of NTCP while conserving a similar TCP, or an increased TCP leaving the NTCP unchanged. By the end of 2020, a total of almost 300 000 patients worldwide have been treated using charged particle therapy [6]. While this number is growing by the year (almost 40 000 patients where treated in 2020), the overall fraction of patients concerned by this treatment modality is quite low. This is mainly due to the cost of hadrontherapy facilities: they require large accelerators such as cyclotrons or synchrotrons and dedicated personnel to handle those machines. Therefore, the cost per treatment about 2 to 3 times higher than the one of conventional therapy [7].

Still, as a consequence of the benefits of hadrontherapy, the numbers of clinical centers around the world is steadily increasing. As of today, 12 centers are able to deliver carbon-ion beam as a treatment modality, while 100 can deliver protons [6]. Others complementary ions, such as helium or oxygen, are currently in development phases in several centers.

Recommended indications for the use of charged particle therapy

Charged particle therapy can be considered a medical necessity in a few different cases: tumor with complex geometries, often close to organ at risks (OAR), organs that should be spared to avoid detrimental effect to the patient; pediatric tumors and radiosensitive patients for which the exposition of healthy tissue should be limited; and finally for heavier-ions than protons, radio-resistant tumor sites. The recommended indications for several countries can be found in table 1.2.

Furthermore, it was pointed out that charged particle therapy could be a viable alternative for conventional radiotherapy in a wide variety of cancer cases.

- Prostate patients, who benefit from the sparing of bowel and bladder while conserving the same TCP [8].

- Breast and lung cancer patients, who are prone to cardiovascular disease following irradiation of the heart while treating the tumoral sites, which could be avoided using charged particle therapy [9] .

Country	Medical necessity	Potential indications
U.S.	Eye tumors Chordoma and chondrosarcoma Spine tumors Hepatocellular carcinoma Pediatric tumors Patients with radiosensitive syndromes	Head and neck cancers Thoracis malignancies Abdominal cancers Pelvic cancer All other solid tumors
U.K.	Skull base and spinal chordoma Skull base chondrosarcoma Spine and paraspinal soft-tissue sarcomas Pediatric tumors	
Italy	Skull base and spine chordomas and chondrosarcomas Adenoid cystic carcinoma of the salivary glands Mucosal malignant melanoma Ocular melanoma Osteosarcomas Pediatric tumors	
Denmark	Chordoma and chondrosarcoma Ependymoma Primitive neuroectodermal tumors Pituitary adenoma Acoustic neuroma Arterovenous malformations Germinoma Eye tumors Lymphomas Selected sarcomas Nasopharyngeal cancer recurrence Pediatric tumors	
The Netherlands	Skull base and spine chordomas and chondrosarcomas Meningionma Pediatric tumors	Re-irradiations Nasopharyngeal carcinoma Retroperitoneal sarcoma Paranasal simus tumors
Canada	Chordomas and chondrosarcomas Ocular melanomas Pediatric tumors	Benign tumors of the central nervous system Paranasal sinus and nasal cavity tumors

Table 1.2: Recommended indications for charged particle therapy in several countries. Taken from [7].

- Liver and pancreatic cancers, associated with low survival rates (15% and 6% at 5-year respectively), were treated by carbon-ion beams in NIRS with promising results [10].

Finally, it should be noted that one of the major contraindication for charged particle therapy concerns tumor sites localized in extensively moving regions: a treatment would indeed require effective tracking methods in order to be fully accurate. While a few different techniques have been developed to this day, it is still a major obstacle the field has to overcome [11].

A point on cancer biology

Cancer is a family of diseases defined by unrestrained cell proliferation, leading to invasion and disruption of the function of normal tissues and organs. In essence, cancer arises from deregulation of various cellular mechanisms. Because of its nature, cancer can only be depicted through opposition: knowledge about usual cell operation is required to explain how tumors are formed.

Cell and DNA: The main purpose of the cell is to replicate itself or rather its genetic information. The genetic information carrier, in every multicellular organisms, is stored on a double-stranded helix, the DNA, or deoxyribonucleic acid. This molecule is tightly arranged into sets of chromosomes inside the nucleus of the cell. Ensuring that it is exempted from alteration and is properly copied at each replication is of key importance to properly achieve this function. Therefore, a large part of the cell resources and mechanisms are dedicated to the monitoring and restoration of the integrity of the genetic information.

Repair mechanisms: Even under normal cell conditions, DNA is prone to various form of damages, notably through the reactive oxygen species (ROS) produced by aerobic metabolism of the cell. Its double-helical structure is particularly suited to allow repairs: if one of the strands is damaged, the other can be used to fix the nucleotide sequence. In consequence, most repair mechanisms make use of excision of the damaged DNA and synthesis of new DNA can then occur. Base-excision repair, nucleotide-excision repair and mismatch repair fall into this category. A problem arises in case of double strand breaks, where the complementarity of information is lost. Nonetheless, two distinct repair mechanisms for double strand breaks exist. The first one, nonhomologous end joining, combines the two broken ends

of the chromosomes and generally leads to the loss of nucleotides at the junction. The second one, homologous recombination requires an available duplicated chromosome to be used as a template to perform the reconstruction.

Damage supervision and cell death: Controlling the integrity of the genetic information, and acting accordingly, is a tremendous task. This role is ensured in each cell by a dedicated component: the protein p53, which acts as the cell supervisor. By receiving signals informing it that genetic damage or metabolic disorder are present within a cell, p53 may stop the advance of the cell through its cycle, and launch the repair program. If those damages are too important, it can initiate cell death through the activation of the apoptosis pathway.

The development of tumors, or tumorigenesis, is a complex process requiring multiple steps in order to turn normal cells into malignant ones. Those different steps correspond to various alterations of molecular mechanisms regulating cell growth, proliferation, or death. Six predominant capabilities shared by most cancers were identified by Hanahan and Weinberg ; they were revised a decade later to include two more capabilities [12, 13].

Apoptosis deregulation: As mentioned above, programmed cell death is one of the necessary mechanisms for the cells to control the rise of malignant growth. As such, it is one of the first obstacles cancerous cells must remove in order to freely proliferate. Usually, it is achieved through the loss of functionality of the aforementioned p53 protein, central to the damage sensor circuitry of the cell.

Hypoxic workaround: Unlimited proliferation implies the need for a substantial amount of resources, be it oxygen or nutrients. The growth resulting from this proliferation may result in a cutoff from vasculature, removing access to those critical supplies. This is indeed the reason why a large number of solid tumors are hypoxic, which corresponds to a state of oxygen deprivation. Thus, malignant cells, if they are to strive, should develop ways to either incite the expansion of vascular infrastructure (known as angiogenesis) or reconfigure their energy metabolism to circumvent resource shortage [14].

Immunosurveillance evasion: Overall, tumorigenesis is a complicated process involving more than just the malignant cells: not only the surround-

ing environment is involved, but the host immune system plays a role as well. Indeed, the interaction between immune surveillance and the tumoral mass follows a delicate balance, and can lead to either side winning. It is therefore not surprising that cancer can lay dormant in the host organism for several years.

After this global overview of cancer, the physical properties of particles as well as their radiobiological effects need to be reviewed in order to illustrate the benefits of hadrontherapy.

1.2 Physical aspects of radiotherapy

In order to understand how radiotherapy works, knowledge about the interactions particle can undergo while traveling through matter is necessary. The interactions relevant to conventional therapy are reviewed first.

Photon interactions

Photons interact through isolated processes. Among them, three are of particular interest for this discussion: the photoelectric effect, Compton scattering and pair production. The energy domain in which they are dominant is reported in figure 1.3.

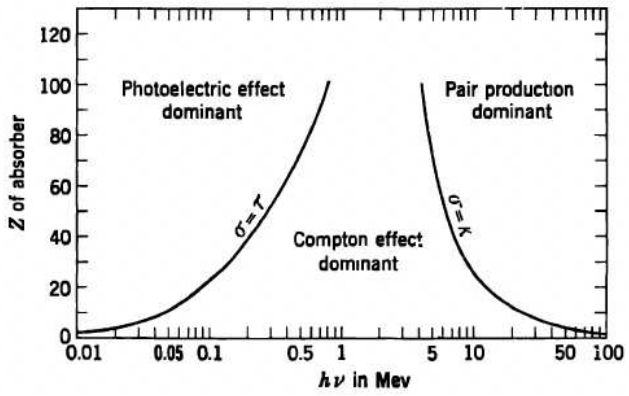

Figure 1.3: Relative importance of the three types of photon interactions evoked here. Taken from [15].

Photoelectric effect If the photon undergoes such an interaction, it is absorbed by one of the atoms of the medium. An electron from one of

the higher electronic shells of the atom is ejected in its place. The energy of the electron is therefore $E_{e-} = h\nu - E_b$. The photoelectric process is the interaction mode that dominates at low energy (below 1 MeV), and its probability is affected by the charge number of the absorber atoms: the more electrons available, the more probable this process.

Compton scattering This interaction involves an incoming photon and an electron from the medium: the photon is deflected from its course, and some of its energy is transferred to the electron according to the equation:

$$h\nu' = \frac{h\nu}{1 + \frac{h\nu}{m_e c^2}(1 - \cos\theta)} \tag{1.1}$$

This effectively means that the resulting electron has a wide range of possible energy, since scattering can occur in any angle. The liberated electron can have an energy up to:

$$E_{e-}|_{\theta=\pi} = h\nu \left(\frac{2h\nu/m_e c^2}{1 + 2h\nu/m_e c^2} \right) \tag{1.2}$$

Therefore the initial photon always conserves some of its energy: $E_c = h\nu - E_{e-} = h\nu/(1+2h\nu/m_e c^2)$. With higher energies of the incident photon, the conserved energy tends towards a constant value $E_c \approx m_e c^2/2$. Again, the probability of such a process to occur increases with the available orbital electrons from the medium.

Pair production This process requires an energy of the incident photon superior to twice the rest-mass of an electron. If such a consideration occurs, a pair of electron and positron can be generated, while the photon disappear. The exceeding energy is shared between the pair as kinetic energy. The probability of such a process to occur increases with the available energy.

$$E_{e-} + E_{e+} = h\nu - 2m_e c^2 \tag{1.3}$$

Interactions of charged particles with electrons and associated notions

The dominant mode of interaction for charged particles traversing a medium is inelastic collisions with the orbital electrons of the material. As will be exposed in this section, it is indeed the physical justification behind the concept of hadrontherapy.

Stopping power

While traversing matter, the incident particle exert an attractive force on the orbital electrons in its proximity. The force applied may be enough to fully ionize the electrons, based on the closeness of the encounter. The energy transferred to the electron comes from the charged particle, whose kinetic energy decreases accordingly.

Given E and $m_{projectile}$ the energy and mass of the incoming particle, m_e the mass of the electron and θ the angle at which the electron is scattered, the energy transferred can be expressed by the following equation:

$$T = E\frac{4m_e}{m_{projectile} + m_e}\cos^2\theta \tag{1.4}$$

For a particle heavier than an electron, the maximum transferred energy is $T = 4Em_e/m_{projectile}$, which represent a low fraction of the total energy of the particle. In order to loose the entirety of its energy in the medium, a charged particle must therefore experience many of those collisions. Inelastic scattering from electrons of the medium is described by the Bethe-Bloch formula [16], which theorizes the energy loss for the incident particle, also called stopping power, and can be derived in its classical form through the following considerations.

Let a particle with charge ze and velocity v, interacting in a medium with N atoms per cm^3 of atomic number Z_{medium}. Since the mass of the incident particle is several time larger than the one of the electrons, no scattering can occur and the trajectory is not modified through interactions. The impact parameter b is the principal value of interest: it corresponds to the distance between the center of the two interacting particles in the perpendicular plane to the direction of motion. Interactions occurs on a length $\mathrm{d}x$ during $\mathrm{d}t$ such that $v = \mathrm{d}x/\mathrm{d}t$.

The momentum transferred from the incoming particle to one electron of the medium is therefore:

$$\Delta\overrightarrow{p} = \int \overrightarrow{F}\mathrm{d}t = \int e\overrightarrow{E}\mathrm{d}t \tag{1.5}$$

Symmetry of the position of electrons around the incoming ion implies the simplification:

$$\Delta\overrightarrow{p} = \int e\overrightarrow{E_\perp}\mathrm{d}t = \frac{e}{v}\int \overrightarrow{E_\perp}\mathrm{d}x. \tag{1.6}$$

According to Gauss's law, the electric flux through a surface S enclosing a volume V, with Q the total charge contained within that volume is expressed by $\Phi_E = Q/\epsilon_0$. The electric flux can also be written as

$\Phi_E = \oiint_S \overrightarrow{E}\,\mathrm{d}\overrightarrow{s}$. Considering the cylinder defined by the height $\mathrm{d}x$ and the impact parameter b, of lateral surface $\mathrm{d}s = 2\pi b\mathrm{d}x$ we get:

$$\Delta\overrightarrow{p} = \frac{e}{2\pi bv}\int \overrightarrow{E_\perp}\mathrm{d}s = \frac{e}{2\pi bv}\frac{ze}{\epsilon_0} \tag{1.7}$$

Thus, the energy given to any one the electrons is:

$$\Delta E(b) = \frac{1}{2}\frac{(\Delta\overrightarrow{p})^2}{m_e} = \left(\frac{1}{4\pi\epsilon_0}\right)^2 \frac{2z^2e^4}{b^2v^2m_e} \tag{1.8}$$

From there, the energy lost by the incident particle to the orbital electrons included between the impact parameter b and $b + \mathrm{d}b$ can be deduced:

$$-\mathrm{d}E(b) = NZ_{target}\Delta E(b)\mathrm{d}V = NZ_{target}\Delta E(b)\ 2\pi\mathrm{d}x\mathrm{d}b \tag{1.9}$$

Applying the value for $\Delta E(b)$ found above, and re-arranging the equation, we get:

$$-\frac{\mathrm{d}E(b)}{\mathrm{d}x} = NZ_{target}\ \frac{1}{4\pi\epsilon_0^2}\ \frac{z^2e^4\mathrm{d}b}{v^2m_eb} \tag{1.10}$$

Integrating over the impact parameter finally leaves us with the stopping power:

$$-\frac{\mathrm{d}E}{\mathrm{d}x} = \int_{b_{min}}^{b_{max}} NZ_{target}\ \frac{1}{4\pi\epsilon_0^2}\ \frac{z^2e^4}{v^2m_eb}\ \mathrm{d}b = NZ_{target}\ \frac{1}{4\pi\epsilon_0^2}\ \frac{z^2e^4}{v^2m_e}\ln\frac{b_{max}}{b_{min}} \tag{1.11}$$

The lower limit for the integration, b_{min} corresponds to a frontal collision, where the transferred energy is maximum. Taking into account relativistic limits, it takes the form:

$$b_{min} = \frac{1}{4\pi\epsilon_0}\frac{ze^2}{\gamma m_ev^2} \tag{1.12}$$

The upper limit of the impact parameter, b_{max} correspond to the point where the transferred energy is not enough to free the electron: this is represented by the average ionization potential of the absorber medium, $\langle I\rangle$.

$$b_{max} = \frac{1}{4\pi\epsilon_0}\frac{ze^2}{v}\left(\frac{2}{m_e\langle I\rangle}\right)^{\frac{1}{2}} \tag{1.13}$$

We are finally left with the stopping power S :

$$S = -\frac{\mathrm{d}E}{\mathrm{d}x} = NZ_{target}\ \frac{1}{4\pi\epsilon_0^2}\ \frac{z^2e^4}{v^2m_e}\ln\left(\gamma m_ev\left(\frac{2}{m_e\langle I\rangle}\right)^{\frac{1}{2}}\right) \tag{1.14}$$

To finally achieve the formula now in use in the radiotherapy world, multiple corrections comings from relativistic considerations, as well as quantum ones, on top of density effects and shell consideration are needed. If all of this is done, the formula then takes the form :

$$S = \frac{4\pi z^2}{m_e v^2} N Z_{target} \left(\frac{e^2}{4\pi\epsilon_0}\right)^2 \left[\ln\frac{2m_e v^2}{\langle I\rangle} - \ln(1-\beta^2) - \beta^2 - \frac{C}{Z_{target}} - \frac{\delta}{2}\right] \tag{1.15}$$

with $\beta = v/c$, C the correction factor for the shell model screening effect at low energies and δ the density correction term at higher energies.

This equation is considered valid down to energies where the velocity of the incident particle is comparable to the one of the orbital electrons, region where charge exchanges become a prevalent effect. When looking at the higher part of the energy spectrum, the equation is viable up to a point where radiative effects, through Bremsstrahlung interaction, become dominant.

Three conclusions emerges from the equation 1.15:

- Because of the term in z^2 in the numerator, particles with greater charge will loose more energy over the same traveled length
- The product NZ_{target} represents the electron density of the medium: high-density and high atomic number materials will result in a greater loss of energy
- For non-relativistic particles, the term in $\langle I\rangle$ is dominant: the energy loss is governed by its $1/v^2$ dependency, in other words, the stopping power varies inversely with particle energy. Therefore, the slower the incident particle, the higher its energy loss will be.

Range and straggling

The length of the path of the particle inside of a medium could in theory directly be found by integrating the stopping power formula:

$$R(E_{particle}) = \int_0^{E_{particle}} \left(\frac{\mathrm{d}E}{\mathrm{d}x}\right)^{-1} \mathrm{d}E \tag{1.16}$$

Small variations of energy loss due to statistical fluctuations in the number of collisions, also called energy-straggling, will therefore lead to different path length for particles of the same energy.

In consequence, the range is an averaged quantity and is defined for a beam, through the continuous slowing down approximation (CSDA).

R_{CSDA} is therefore the path length any particle composing the beam would traverse until it stops, if its rate of energy loss was equal to the one given by the stopping power formula.

Since it is directly dependent on the Bethe-Bloch formula, it directly follows that an heavier particle will need much more energy to reach the same depth as a lighter one as illustrated in figure 1.4. This is of particular relevance for hadrontherapy, because of the need to irradiate deep-seated tumor, which was estimated to be around 30 cm in water. It corresponds to an energy of 220 MeV/u for protons and helium ions, and an energy of 425 MeV/u for carbon ions.

The Bragg-Kleeman rule give us a practical formula relating the range and the incident energy of the projectiles [17]:

$$R_{CSDA} \approx \alpha E^p \tag{1.17}$$

Here, α is a material-dependent constant and p is a dimensionless factor taking into account the particle velocity.

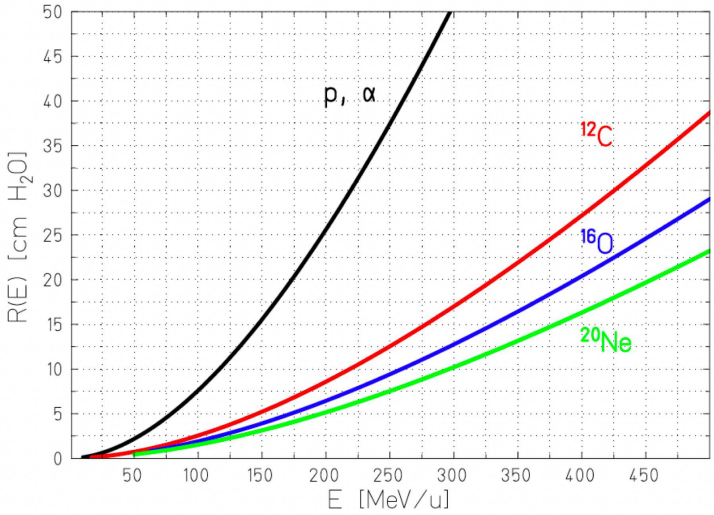

Figure 1.4: Range in water as a function of energy for different ions. The range of protons is shown in black, the one for ^{12}C ions in red, the one for ^{16}O in blue and finally the one for ^{20}Ne in green. The higher the atomic number of the ion, the lowest the range in water for the same energy per nucleon. Taken from [18].

Comparing the range of different particles of the same velocity in the

same medium can be done through the scaling formula:

$$\frac{R_{(2)}}{R_{(1)}} = \frac{M_{(2)}}{M_{(1)}} \frac{z_{(1)}^2}{z_{(2)}^2} \tag{1.18}$$

For example, protons and helium ions beams will have the same range if they have the same incident energy per nucleon, and they will travel three times the distance covered by carbon-ions.

As explained above, the energy straggling caused by fluctuations in the number of collisions experienced by individual particles while traversing a medium results in a range straggling of the beam. Another scaling law allow us to compare the range straggling for different particle with the same range:

$$\frac{\sigma_{R,(1)}}{\sigma_{R,(2)}} = \sqrt{\frac{M_{(2)}}{M_{(1)}}} \tag{1.19}$$

In consequence, the straggling of lighter particles is far more important than the one of heavier particles for the same range.

Bragg curve

Looking at the pattern of energy deposition as a function of depth of a particle in any given medium leads to a curve known as the Bragg curve. Indeed, by comparing the different depth dose profiles of several types of particle in water, one of the main advantage toward the use of charged particle becomes evident, as can be seen in figure 1.5.

While photons present an initial build-up of the dose, which then decreases exponentially with depth, charged particles present a narrow peak at the end of their path, the so-called Bragg peak, where most of the energy of the particle is deposited. Before that, the dose distribution is relatively flat, forming a plateau. This behavior is in direct link with the Bethe-Bloch formula: the energy loss for charged particles increases as the velocity of the incoming particle decreases. Additionally, the effect of range straggling explained in the last section can be visualized: while all three ion beams have approximately the same range, carbon ions present the narrowest peak and protons the broadest.

The rationale behind hadrontherapy is here apparent: by leveraging the property of the energy deposition pattern of charged particles, healthy tissues can be partly spared from irradiation while most of the energy is delivered to the tumor. It is the basic argument made by Wilson [2] which lead to the emergence of protontherapy.

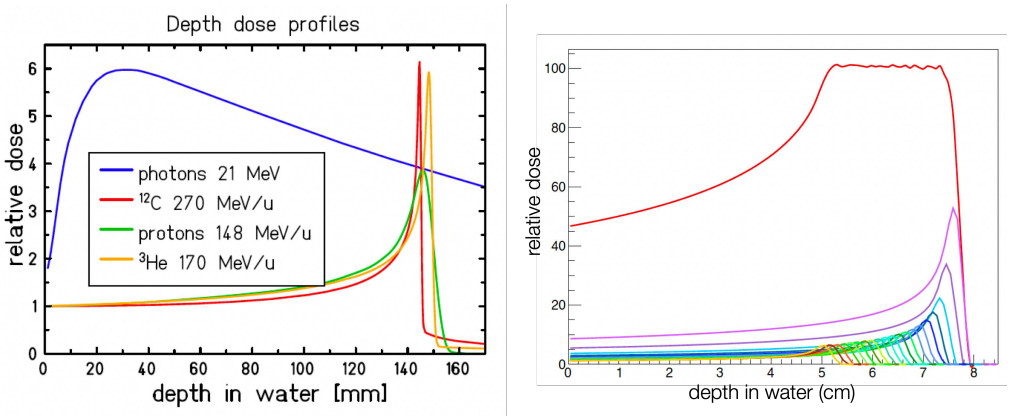

Figure 1.5: Left, depth dose profiles as a function of depth in water for 4 different types of irradiations. Photons corresponds to the blue line, ^{12}C ions to the red one, ^{3}He to the orange one and oprotons to the green one. The respective energies have been selected so that the range in water for the various beams is approximately the same. Taken from [19]. Right, SOBP for a 100 MeV proton beam in red, established from 20 attenuation planes. The relative influence of each beam energy is also shown. Taken from [20].

It should be noted that in reality, the comparison between charged particle and classical radiotherapy is much more nuanced. The narrow Bragg peak formed by a mono-energetic beam of charged particles need to be extended because of the entirety of the targeted tumoral volume has to be covered for the treatment to be effective. This is known as the spread-out Bragg peak (SOBP) and requires several beam energies to be setup. An example of such a SOBP can be found in figure 1.5 for a 100 MeV proton beam in water. As a consequence, tissues in the path of the beam can be exposed to a non-negligeable amount of energy deposition.

However, in conventional therapy, a large amount of beam incidences is required in order to achieve conformation to the tumoral volume and, but results in an overall exposition of more healthy tissues. Charged particle therapy can achieve conformation to the targeted volume with only a few beam incidence, allowing for more sparing of healthy tissues, which is of particular importance when OAR are close by, as can be seen in figure 1.6, which shows 3 different treatment plans for the same patient: the first with conventional radiotherapy, the second with carbon ions and the last one with protons. Since the tumor is located at the base of the skull, both the spine and the eyes (OAR) are in its proximity. The two charged particle plans result in a total sparing of the spine, while the eyes are only slightly irradiated. The photon plan results in a dose delivered to must of the skull,

including the two OARs.

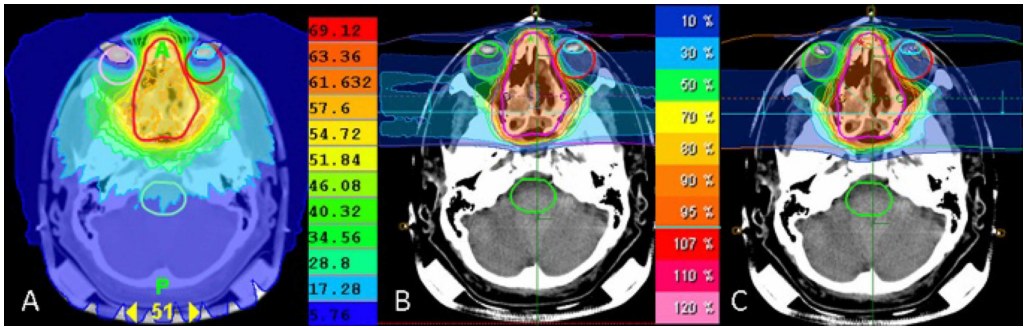

Figure 1.6: Dose distribution for 3 different plans, applied to a skull based tumor. Targeted volume is delimited in red. Several organ at risks are present and should be avoided: the eyes and the spine of the patient. (a) corresponds to photon intensity modulated radiotherapy (IMRT), (b) to carbon ion beam and (c) to a proton beam. Both (b) and (c) make use of two different incidences of the beam. Taken from [21].

Dose

One of the most important quantity in radiotherapy is the dose deposited by the incident particle. Indeed, radiotherapy prescriptions are usually made in terms of total dose to the targeted volume. It corresponds to the mean energy $\mathrm{d}E$ deposited in a mass element $\mathrm{d}m$, as defined in [22]:

$$D = \frac{\mathrm{d}E}{\mathrm{d}m} \qquad (\mathrm{Gy}) \tag{1.20}$$

Dose can be linked back to the average stopping power of the beam, by introducing the density of the medium and the particle fluence, the number of particles per surface area at a given depth.

Linear Energy Transfer

The linear energy transfer of particle inside of a given medium is defined by [23]. It corresponds to the ratio between the average energy locally imparted $\mathrm{d}E_\Delta$ to the medium and the length $\mathrm{d}l$ the particle went through:

$$L_\Delta = \frac{\mathrm{d}E_\Delta}{\mathrm{d}l} \qquad (\mathrm{keV}/\mu m) \tag{1.21}$$

The locally imparted part corresponds to the symbol Δ, which may refer to either: a maximum distance from the track of the incident particle, or

a maximum value of energy loss above which the energy released in the medium is not considered to be locally delivered. Thus, if this limit is taken as infinity, it follows that the corresponding linear energy transfer is the same as the stopping power for the particle $L_\infty = S$.

Usually, linear energy transfer is presented as a dose averaged quantity, for all the species and energies of particle presented at given depth. The absorbed dose at a depth x can be expressed as follows, with $\phi(E, x)$ the fluence for the particle type at a given energy E and a depth x:

$$D(x) = \frac{1}{\rho} \int_0^\infty S(x)\phi(E, x)dE \tag{1.22}$$

From there the dose averaged linear energy transfer, over all N particles and energies E, is:

$$\langle L_\Delta(x) \rangle = \frac{\sum_{i=1}^{N} \int_0^\infty \phi(E, x)S(E, x)L_\Delta(x)\mathrm{d}E}{\sum_{i=1}^{N} \int_0^\infty \phi(E, x)S(E, x)dE} \tag{1.23}$$

LET is a crucial notion when looking at the effect of charged particles during treatment. Indeed, since LET is a representation of how the energy is deposited locally, it can be correlated to the structure of damages done to DNA by those radiations.

Effectively, charged particles present a LET distribution that increases with depth, as the form taken by their stopping power implies. The depth distribution of the LET is maximal after the end of the Bragg peak, as can be seen in figure 1.7, which presents both the dose depth distribution and the dose averaged LET distribution for 110 MeV/u helium beam in water.

In current clinical practice, dose average LET distributions range from 1-3 keV/μm and 10-15 keV/μm in the entrance channel up to 5-10 keV/μm and 40-90 keV/μm in the SOBP, for protons and carbon ions respectively [24].

Interactions of charged particles with nuclei

Elastic interactions

Elastic coulomb scattering on nuclei from the medium is another process that can take place while a particle moves through matter. Those interactions result in a deviation from the initial trajectory of the particle. While it can be considered particle by particle, in the context of hadrontherapy and thus the use of an ion beam, it makes more sense to look at the statistical distribution of the scattering angle θ, after traveling a given length

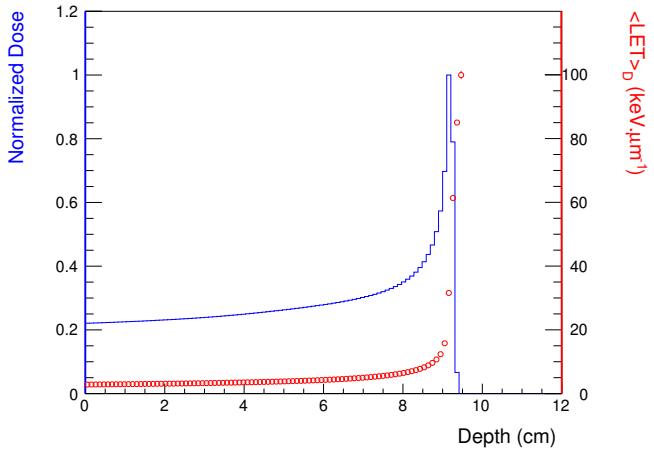

Figure 1.7: Normalized dose and corresponding dose averaged LET distribution for a 110 MeV/u α beam in water.

in the medium of interest. This process, called multiple coulomb scattering, is well described by Molière's theory, which can be approximated by a gaussian function for small angles, with a standard deviation given by [25]:

$$\sigma_\theta = \frac{13.6}{\beta pc} z \sqrt{\frac{d}{L_{rad}}} \left[1 + 0.088 \log_{10} \left(\frac{d}{L_{rad}} \right) \right] \tag{1.24}$$

The absorber is characterized by its thickness d and its radiation length L_{rad}, which itself depends inversely on the square of the atomic number of the medium. Three conclusions immediately arises:

- Thick target containing heavy elements will cause a larger angular spread
- For the same particle, scattering decreases with higher energies because of the term $1/\beta pc$.
- Because of the term $z/\beta pc$ heavier particles scatter less at the same range

Inelastic reactions

As low as the probabilities for inelastic nuclear interactions are between an incident particle and a nuclei from the medium, they should also be considered. Inelastic reactions are characterized by a the total kinetic energy of the system that is not conserved; in some reactions, the compositions of the nuclei involved are modified.

The kinetic energy loss from the incident particle may result in several repercussions: one or both of the nuclei involved will be left in an excited state, therefore emitting photons in order to reach their ground state; nucleons can be exchanged; one or both of the nuclei can fragment into lighter particles; or even, if the energy of the incident particle is included in the Gamow window, fusion between the two nuclei can occur.

It should be emphasized that protons, as they are only composed of one nucleon, cannot fragment themselves. However they can still cause the target nuclei dissolution, or the production of secondary protons and neutrons. Their peculiar status exclude them from the following arguments, which will focus on heavier ions, unless the target fragmentation is considered.

Given the energy range of therapeutic beams, the nuclei involved in the reaction can be approximated by absorbing disks. The radius of any of those disks represents the radius of the corresponding nucleus, and is given by the following formula, with r_0 the nucleon radius, c a correction factor which is nucleus dependent and A the mass number of the nucleus:

$$R = r_0 A^{1/3} - c \tag{1.25}$$

In this framework, the violence of the interaction depends on the overlapping portion of the two disks, which can be represented by the impact parameter of the reaction b. The probability of a collision is therefore proportional to the square of the impact parameter, making central collisions particularly unlikely.

Thus, peripheral collisions are the most likely to occur. This kind of collision results mainly in the creation of fragments from both the projectile and the target. It is described by the abrasion-ablation model, which subdivides the process into two steps: nucleons in the reaction zone from both the projectile and target nuclei form a so-called fireball, while the remaining ones are marginally impacted; de-excitation occurs afterwards for all three components of the reaction: projectile, target and fireball [26, 27]. An illustration of the model can be seen in figure 1.8.

The probabilities for any kind of reaction is represented by the notion of cross-section, which is akin to the surface area of the disks mentioned above and is usually expressed in barns, with $1b = 10^{-24} cm^2$. Cross-sections are an invaluable tool in order to understand the physical mechanisms behind inelastic reactions, and will be formally defined in section 1.2. Furthermore, their knowledge is required in order to take into account the effects of fragmentation on treatments.

One of the basic prediction of the abrasion-ablation model is the property called weak factorization: the idea is that the branching ratio of the

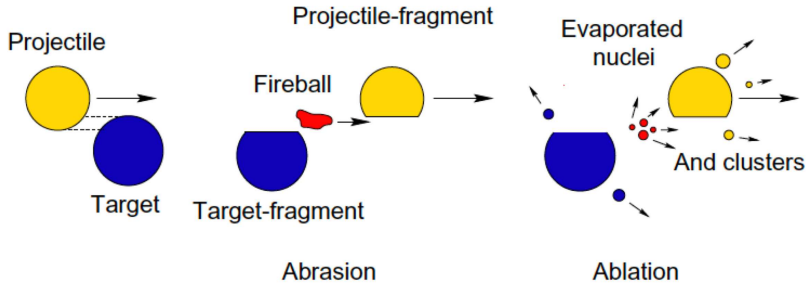

Figure 1.8: Illustration of the two steps abrasion-ablation model. The overlapping portion of the two nuclei involved in the reaction forms a fireball in the abrasion step, while all participants de-excite in the ablation phase, producing various fragments. Taken from [28].

fragmentation channels (i.e. the various particle productions) is independent of the target [29, 30]. According to this property, the cross-section for a given reaction can be expressed as :

$$\sigma(P, F, T) = \sigma_P^F \gamma_{PT} \tag{1.26}$$

Here, P, F and T stand for projectile, fragment and target respectively. σ_P^F is a factor depending upon the projectile and fragment, while γ_{PT} is a factor depending on both projectile and target only.

Under this assumption, the term γ_{PT} can be written in term of the radii of the projectile and target nuclei. Furthermore, it can be used to deduce the fragmentation cross-sections for other targets:

$$\sigma(P, T', F) = \sigma(P, T, F) \frac{\gamma_{PT}}{\gamma_{PT'}} \tag{1.27}$$

Weak factorization was proven experimentally in [31], using a wide variety of projectiles and targets at the therapeutic energy range, therefore validating the use of the abrasion-ablation model in this configuration. Hydrogen targets were also used in this study, and protons were found not to follow the weak factorization rule therefore excluding them from the abrasion-ablation model.

Inelastic reactions will produce drastic consequences for the beam in treatment: in the best case, the range of the incident particle will be reduced while in the worst case they can result in either partial fragmentation or total disintegration of the projectile and target nuclei, therefore creating a mixed field of radiation.

The loss of primary ions, or beam consumption, is an outcome that becomes increasingly prevalent as the beam travel through the medium. As the beam is consumed, a corresponding build-up of fragments can be observed in the medium. For example, experimental measurements of the beam consumption and fragments build up for a carbon ion beam in water may be found in figure 1.9.

Fragments originate from both the target and projectile. Their properties are however vastly different: fragments coming from the former are usually heavy and have a low amount of kinetic energy, therefore depositing it in the vicinity of the interaction; fragments coming from the latter are emitted mainly towards the forward direction, with a velocity close to the one of the beam.

Projectile fragments angular and energy distribution are charge dependent: both distributions get narrower with increasing charge [32]. This can be understood conceptually by looking at the impact parameters required to produce various fragments: a high impact parameter leaves the projectile relatively undisturbed by the target nuclei, and results in a bigger fragment which will retain most of the incident particle characteristics; a small impact parameter leads to a more violent collision, producing smaller fragments with a wider angular and energy range.

The production of lighter fragments is one major drawback for charged particle therapy, and notably for the use of carbon-ions: the light fragments travel further in the medium than the primary ion, causing a dose tail after the Bragg peak of the beam, as can be seen in figure 1.5.

Cross-sections

The formula to obtain a cross-section can be derived considering a beam of N_{beam} particles, impinging on a material of thickness e, with density ρ and A its molar mass. Given a reaction of interest $P + F \rightarrow F + X$, with P the projectile, T the target nuclei, F the fragment and X representing all the other possible outgoing particles, then the total inclusive cross-section, denoted as σ, can then be considered to be the effective area for the reaction of interest. If $N_{fragments}$ of type F are produced after the beam has traversed the material, then the cross-section can be expressed as:

$$\sigma = \frac{N_{fragments}}{N_{beam}} \frac{A}{\rho e \mathcal{N}_{\mathcal{A}}} \tag{1.28}$$

Charge-changing $\sigma_{\Delta Z}$ and mass-changing $\sigma_{\Delta A}$ cross sections are two important cross-sections variations used extensively by experimentalists:

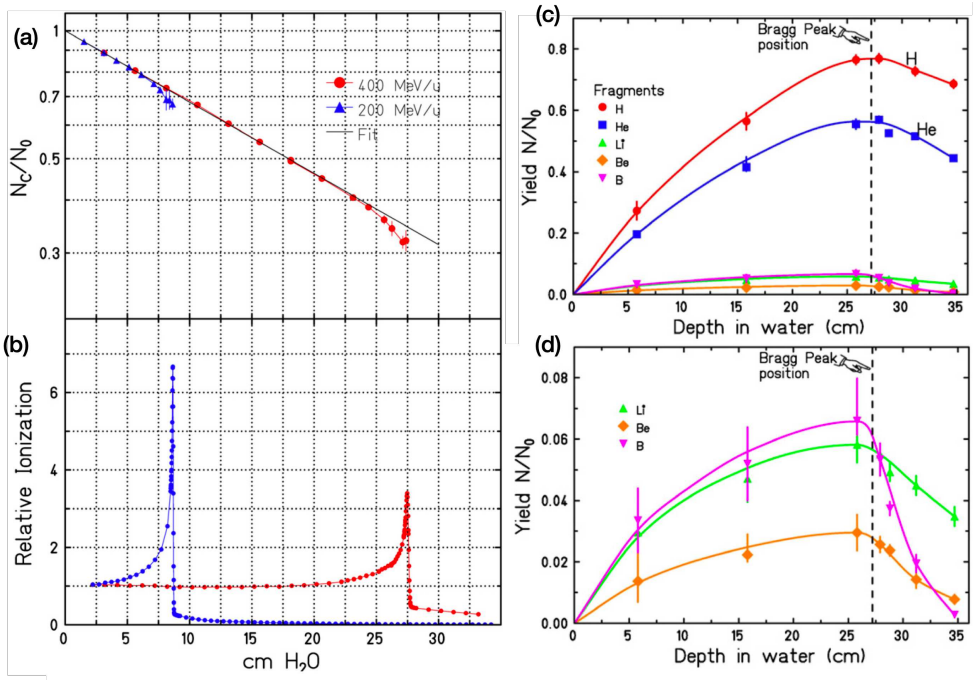

Figure 1.9: Beam consumption and fragments build-up over path in water for carbon-ion beams. (a) shows the ratio of primary ions remaining as a function of depth in water. (b) is the corresponding Bragg curve. For both of them, two different beam energies are shown: 400 MeV/u in red and 200 MeV/u in blue. (c) and (d) are showing the ratio between the number of fragments produced and the initial number of primary ions, for the 400 MeV/u beam. As the proportion of heavier fragments are lower, the corresponding fragment build-up can be seen in (d), zoomed up by a factor ten from (c). Taken from [32].

because they include a large variety of reactions they are often a good approximation to total cross-sections. The first corresponds to the probability for the projectile to have its atomic number modified after the interaction; the second matches the probability for mass number modification of the projectile, therefore including neutrons removal. While charge identification is relatively straightforward with a simple detection apparatus, mass identification represents an experimental challenge and therefore requires advanced setups to be performed.

These two types of cross-sections can be sub-divided further into elemental cross-sections, which focus on the production of a fragment with given atomic number Z, or even isotopic cross-sections, for which both atomic and mass number are identified. The former can be deduced by summing

over all of the isotopic cross-sections for a given atomic number. They are also both known under the terms partial charge-changing and partial mass-changing cross-sections, respectively.

The quantities considered so far are not taking into account for angular of energy information: they are integral or total cross-sections. However, cross-sections measurements as a function of one or both of these variables are inherently more informative on the processes that occurred. Furthermore, they can be integrated to recover the values for the corresponding total cross-sections.

Measurements of angular single differential cross-section are often an experimental constraint: detectors have a given size and therefore can only cover some of the emission angle for the fragments that are produced. Double differential cross-section requires significant counts over a wide range of energy and angles in order not to be dominated by statistical errors. As argued in [33], isotopic double differential cross-sections present a particular interest for the field: an important lack of experimental data remains at energies of interest for hadrontherapy.

Through the years, several cross-sections measurements were performed by various teams around the world: table 1.3 and 1.4 summarize the different experiments performed on thick and thin targets, respectively. Targets can be considered thin in case the incident particle can at most interact with a few target nuclei, with an negligible energy loss.

This overview of those experiments focuses on measurements made in the charged particle therapy context. Therefore, only experiments using carbon, oxygen or helium ion beams, impinging on water, carbon, polyethylene, polycarbonate or acrylic targets at the therapeutic energy range are considered.

To this day, most experiments used thick targets, which allows for charge-changing and elemental cross-sections to be established. Isotopic cross-sections represents a difficulty in such conditions: the fragments energy and angular resolutions are degraded inside the target, making their mass identification more difficult. For the same reasons, differential cross-sections are hard to achieve, even though angular and energy distributions for the fragments can still be retrieved. Thin targets, in the other hand, permit an easier measurement of isotopic double differential cross-sections, but require additional irradiation time in order to produce significant statistics.

Paper	Site	Beam	Energy (MeV/u)	Target	Measurements
[34]	LBL (Berkley)	^{12}C	250	C	$\partial\sigma/\partial p_T$, $\partial\sigma/\partial\Omega$, $\sigma_{\Delta A\ (\Delta A\leq 6)}$
[35]		^{16}O	215	H_2O	$\sigma_{\Delta Z\ (\Delta Z\leq 3)}$
[36]		^{12}C, ^{16}O	300-1700	CH_2, C, He	$\sigma_{\Delta Z}$, $\sigma_{\Delta A}$
[37, 38]		^{12}C, ^{16}O	200-670	H_2O, CH_2, PMMA	$\sigma_{\Delta Z}$, $\sigma_{\Delta Z\ (\Delta Z\leq 3)}$
[39]	GSI (Darmstadt)	^{12}C	270	H_2O	$\sigma_{\Delta Z\ (\Delta Z\leq 5)}$, $\partial\sigma/\partial\Omega$
[28]		^{12}C	200	H_2O	$\partial y/\partial\Omega$, $\partial^2 y/\partial\Omega\partial E$
[32]		^{12}C	200, 400	H_2O	$\partial y/\partial\Omega$, $\partial^2 y/\partial\Omega\partial E$
[40]		^{12}C	500	H_2O, PMMA, C_2H_4, C	$\sigma_{\Delta Z}$, $\sigma_{\Delta Z\ (\Delta Z\leq 5)}$
[41]		^{12}C	110-250	H_2O, C, CH_2	$\sigma_{\Delta Z}$, $\sigma_{\Delta Z\ (\Delta Z\leq 2)}$
[42]	HIMAC (Chiba)	^{12}C	200-400	H_2O, $C_{16}H_{14}O_3$	$\sigma_{\Delta Z}$, $\sigma_{\Delta Z\ (\Delta Z\leq 3)}$
[43]		^{12}C	313	$C_{16}H_{14}O_3$	$\sigma_{\Delta Z}$, $\sigma_{\Delta Z\ (\Delta Z\leq 4)}$
[44, 45]		^{12}C, ^{16}O	290-600	H, C	$\sigma_{\Delta Z}$, $\sigma_{\Delta Z\ (\Delta Z\leq 5)}$, $\partial\sigma/\partial\Omega$
[46]		^{12}C	100-400	C	$\sigma_{\Delta A\ (\Delta A\leq 6)}$
[47, 48, 49]	HIT (Heidelberg)	^{4}He	120-200	H_2O, PMMA	$\partial y/\partial\Omega$, $\partial^2 y/\partial\Omega\partial E$
[50, 51]		^{4}He	70-220	H_2O, C, CH_2	$\sigma_{\Delta Z}$, $\sigma_{\Delta A}$

Table 1.3: Cross-sections experiments on thick targets. Total charge and mass changing cross-sections are denoted as $\sigma_{\Delta Z}$, $\sigma_{\Delta A}$ respectively. Elemental and isotopic cross-sections are referred as $\sigma_{\Delta Z\ (Z\leq N_Z)}$ and $\sigma_{\Delta A\ (\Delta A\leq N_A)}$ with N_Z and N_A referring to the maximum charge and nucleon removal observed in the corresponding experiment. Single and double differential yields $\partial y/\partial\Omega$, $\partial^2 y/\partial\Omega\partial E$ and cross-sections $\partial\sigma/\partial\Omega$, $\partial\sigma/\partial p_T$, $\partial^2\sigma/\partial\Omega\partial E$ have been measured in some experiments. Differentials quantities are reported per isotope.

Paper	Site	Beam	Energy (MeV/u)	Target	Measurements
[52]	CERN (Geneva)	^{12}C	86	C	$\sigma_{\Delta A\ (\Delta A \leq 6)}$, $\partial^2\sigma/\partial\Omega\partial E$
[53]	LPS (Catania)	^{12}C	62	C	$\partial\sigma/\partial\Omega$, $\partial^2\sigma/\partial\Omega\partial E$
[54, 55]	GANIL (Caen)	^{12}C	95, 50	C, CH_2, PMMA	$\sigma_{\Delta A\ (\Delta A \leq 11)}$, $\partial\sigma/\partial\Omega$, $\partial^2\sigma/\partial\Omega\partial E$

Table 1.4: Cross-sections experiments on thin targets. Measurements are denoted as explained in table 1.3.

A side note on Monte-Carlo simulations

Monte-Carlo simulations in the context of hadrontherapy refers to tools aimed to solve particle transport and interaction in matter. The method in itself is a mathematical tool intended to solve problems using random sampling. Several Monte-Carlo packages are commonly used in the field: Geant4 [56], GATE [57], TOPAS [58], FLUKA [59], PHITS [60], MCNPX [61] or SHIELD-HIT[62].

In essence, a Monte-Carlo simulation of a particle going through matter is done according to the following steps [63]:

- Selection of the distance towards the next interaction point, based on random sampling of the different possible interactions according to their respective cross-sections in the current material
- Transport of the particle to the interaction point, taking into account electromagnetic fields and applying energy loss
- Application of the interaction type, production of secondary particles with properties sampled according to the available energy and angular distributions

These steps are applied repetitively until the particle and all the secondary particles it generated either have reached the minimum amount of energy or escape the geometry defined by the user.

The main benefit of Monte-Carlo methods as a tool to simulate the transport of particles through matter lies in the reproduction of nuclear reactions. It is generally performed through several steps, reminding the

abrasion-ablation model evoked in section 1.2: first, a dynamic phase takes place, where excited and compound nuclei are formed by the use of intranuclear cascades; afterwards, a pre-equilibrium phase intervenes, where those nuclei go back to equilibrium through the emission of light fragment; finally, products are brought to their ground state in the evaporation phase. The last phase is the one responsible for the production of heavier fragments.

Overall, the aforementioned steps rely heavily in theoretical models, that need to be benchmarked with experimental data in order to verify the accuracy of their results. A few studies have looked at various models and packages for a few reactions in conditions relevant to therapy [64, 65, 66, 67, 68, 69].

While significant progress was made to improve Monte-Carlo methods through the years, inconsistencies can still be found, and especially when looking at differential quantities for the secondary particles. On top that, theoretical models often disagree with each other, as can be seen in figure 1.10, which shows both single differential cross-sections for the production of an ^{4}He fragment and the double differential cross-sections for the same fragment at 17°. The cross-sections for three Geant4 models are shown on top of the experimental data: the agreement is correct at best for the single differential quantity, and large discrepancies can be observed for the double differential cross-section.

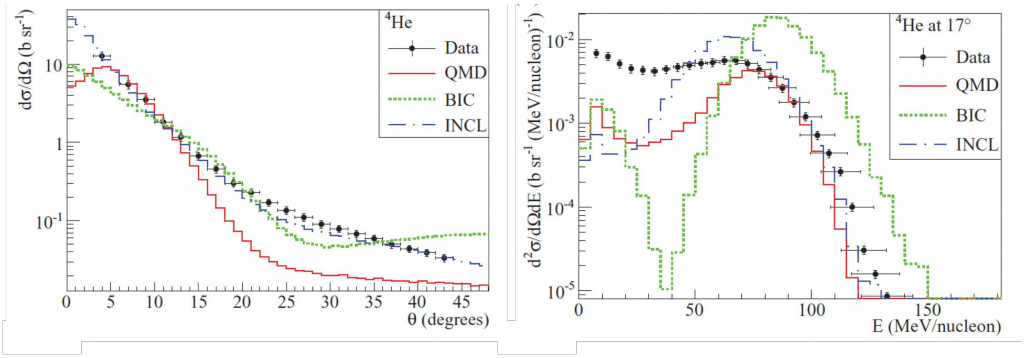

Figure 1.10: Left, single differential cross-sections for the production of an ^{4}He fragment as a function of the emission angle, for carbon ions interacting in a carbon target. Right, double differential cross-sections for the same fragment at 17°. Measured data corresponds to the black circles. Predicted cross sections from three Geant4 models are shown: QMD in red, BIC in green and INCL in blue. Taken from [67].

Now that the physical properties of particles going through matter have been established, the next step in order to understand how radiotherapy

works is to explain their radiobiological impact on tissues, healthy and malignant.

1.3 Radiobiological aspects of radiotherapy

While going through matter, both photons and charged particles transfer their energy to orbital electrons of the medium, causing mostly ionization. The energy deposited by the incident particle will break chemical bonds, weakening the overall structure of the molecules in their path. Furthermore, those ionized electrons are now free to deposit their additional energy to other bound electrons while traveling through the medium, therefore increasing the disturbance to the neighboring environment. Because of their specific energy deposition patterns, photons are considered as sparsely-ionizing radiations while charged particles are densely ionizing.

Direct breaks of chemical bonds, known are direct damages, is not the only way incident particles can produce structural disturbance to the medium: ionization of water molecules will produce free radicals or ROS, that are highly reactive and will further add to the disruption of the medium, therefore causing indirect damages [70].

When irradiating living matter, any part or component of a cell can be exposed to such lesions, but for most molecules, multiple copies are present inside of a cell, and the turnover can be important. However, DNA is present in two copies at most. On top of that, it is the largest molecule present and therefore a big target for the incoming radiation.

Cells have developed various mechanisms to repair lesions to DNA in order to preserve genomic integrity, as evoked in section 1.1. The volume of the lesions produced by ROS on a daily basis is far higher than the ones caused by a classical irradiation. Therefore, understanding why the DNA damages caused by irradiation relative to the one from daily exposure to ROS leads to significant cell death is crucial [71].

The lesions to DNA aforementioned include single strand break, alteration or loss of bases, cross-links formed between complementary strands and finally double strand breaks. Several studies demonstrated that the main asset of charged particles, regarding cell death, is the complexity of the damages done to DNA: the several possible lesions enumerated above are produced in very close proximity to each other, forming DNA damage clusters, a pattern that increases with more densely-ionizing radiations [72, 73].

The clustered sites also lead to an overall slower repair kinematics: to restore properly the genetic information, a high coordination between repair

pathways is required, coordination that might not always be achievable. Thus repair of such lesions can trigger a snowball effect: further double strand breaks can be produced while the primary damages are processed.

Relative Biological Effectiveness

Because conventional radiotherapy has been and still is the predominant choice for treatment by radiations, it also benefits from the knowledge accumulated so far in its clinical use. Therefore, charged particle therapy prescriptions are usually issued in term of photon-equivalent dose, or biological dose. This is illustrated by the crucial notion of relative biological effectiveness (RBE). This quantity allows for convenient comparisons between radiation qualities.

In the context of radiotherapy, TCP and NTCP are two of the main goals: RBE is therefore historically based on cell survival curves, which can be related to both.

Those curves relies on clonogenic assay, where the capacity of cells to form a sizable colony of descendant is tested. In order to produce meaningful results, control cultures are required. Cells of interest are extracted from the tumor, placed in a growth environment, and after some pre-defined amount of time, the colonies are scored.

Cell survival curves show the fraction of surviving cells against the dose necessary to produce this effect. In order to be established, several assays at various doses are necessary. Survival curves are usually presented in logarithmic scale. One of the most used model for cell survival curves is the so-called linear-quadratic model: a second-order polynomial is fitted to the experimental points, with a zero constant term to make sure that survival is 100% when no dose is applied. The probability of survival given a dose D as input is therefore:

$$p = \exp(-\alpha D - \beta D^2) \tag{1.29}$$

Using this model, the bend of the curve is then defined by the α/β ratio, which corresponds to the dose for which the linear and quadratic contributions to the damages done to the cell are equal. It should be noted that while this model works well for photons irradiation, the linear part of the model is often enough to describe charged particles response.

Cell survival curves for various radiation types can be found in figure 1.11. Comparing two of those curves leads to the notion of RBE: in order to produce the same cell survival outcome, usually 10%, less dose is needed overall for ions with respect to photon irradiation. Indeed in the example

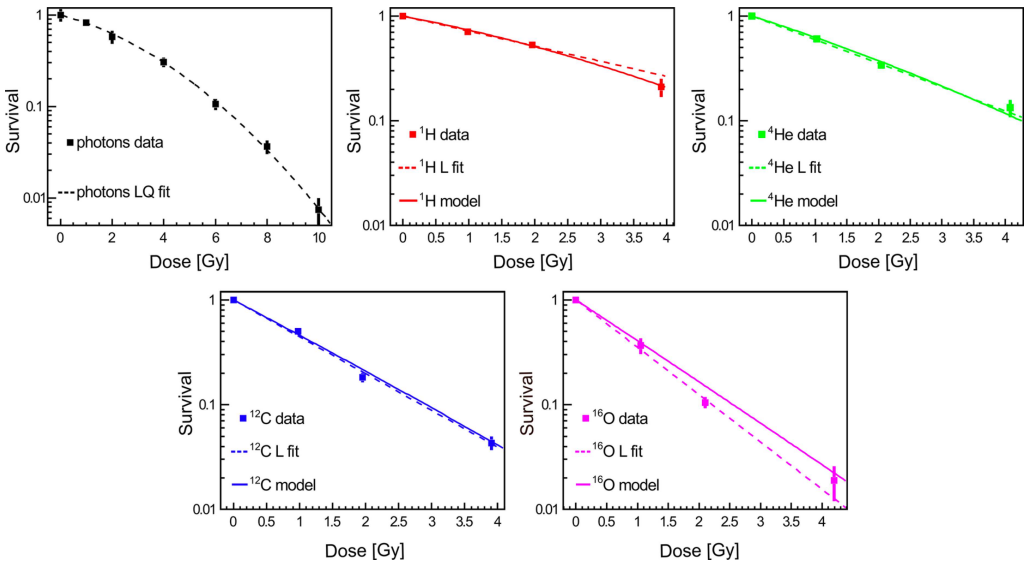

Figure 1.11: Cell surival curves for various types of particles, on A549 cells. Experimental points are fitted with LQ model and prediction from the LEM section 1.4. Taken from [74].

given here, 6 Gy are necessary for photons to produce an outcome of 10% survival, while it is around the same value for protons, 4 Gy for ^{4}He, 3 Gy for ^{12}C and 2 Gy for ^{16}O. The RBE can then be defined as the ratio between D_{ion} the dose from the particle type for which a comparison is necessary and $D_{reference}$ the dose from the reference particle.

$$\mathrm{RBE}_{endpoint} = \frac{D_{reference}}{D_{ion}} \tag{1.30}$$

Finally, it should be noted that RBE is a complex notion depending on a wide variety of parameters such as dose, fractionation, LET, particle type, cell line, endpoint, oxygen concentration, cell cycle phase, radiosensitivity. The dependence of RBE on LET is illustrated in figure 1.12, for protons, helium, carbon, neon and heavier ions. All particles follow a similar trend; RBE_{10} first rises with LET up to a maximum around 100 - 200 keV.μm^{-1} and then drops down because of the overkilling effect; at this point, cells crossed by the beam are already inactivated, increasing the complexity of the lesions to DNA will not provide any additional effects.

While the increased RBE of charged particles is widely accepted in the community, it is not the only biological advantage of this modality. Indeed, heavier particles than protons have been shown to be particularly useful when used to treat tumors that are resistant to conventional radiotherapy.

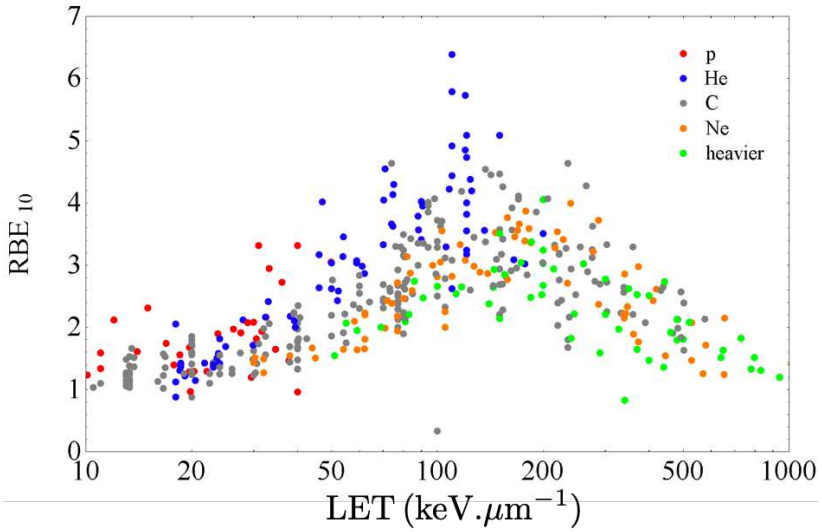

Figure 1.12: RBE_{10} as a function of LET, for protons in red, helium ions in blue, carbon ions in gray, neon ions in orange and heavier ions in green. Data from the database PIDE [75].

Overcoming radiation-resistant cell lines

To illustrate how and why charged particles are able to kill cells that are deemed radio-resistant, this section will rely on a study by *Amornwichet et al.* which focused on a comparison between x-rays and carbon ions [76]. Figure 1.14 recapitulates their findings.

Radioresistance can be achieved by cells through the alteration of the TP53 gene, which is responsible for the regulation of the p53 protein, and therefore plays a major role in the activation of the apoptosis pathway, as earlier mentioned in section 1.1. The quoted work reported on two cell lines (HCT166): the first had an unaltered gene while for the second the TP53 gene was inactivated.

Cell survival curves for both particle types as well as both cell lines are represented in figure 1.13. For the non-radioresistant cell line, the dose needed to produce a survival of 10% of the cell irradiated was 3.8 Gy for x-rays while it was 1.7 Gy for carbon ions. For the altered cell line, the dose needed for x-rays to produce the same outcome was 6.8 Gy while it was only 1.9 Gy for carbon ions. This shows that indeed, the altered cell line present a resistance to conventional radiotherapy, while the effect of carbon-ions is independent of the cell line status.

The corresponding RBE values are approximately 2.2 and 3.6 for the non-radioresistant and radioresistant cell lines, respectively. Those values

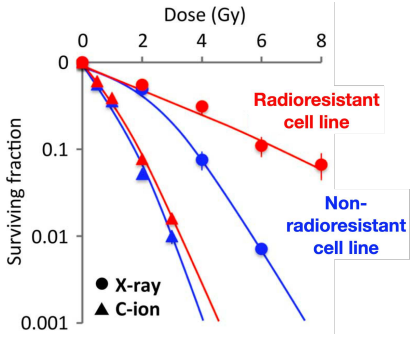

Figure 1.13: Cell survival curves for the non-radioresistant and radioresistant HCT116 cell lines, in blue and red respectively, fitted by the LQ model. Irradiations with x-rays correspond to circles, while carbon-ion ones correspond to upper triangles. Taken from [76].

demonstrate that carbon ions are killing radioresistant cells more effectively. In order to understand why, the type of cell death induced by each radiation type need to be surveyed.

Figure 1.14.(a) and figure 1.14.(b) report the proportions of the principal modes of cell death at various time after x-rays irradiation on the non-radioresistant and radioresistant cell lines, respectively, while figure 1.14.(c) and figure 1.14.(d) corresponds to carbon ions. Three death modes were reported: apoptosis, mitotic catastrophe and senescence.

Apoptosis corresponds to programmed cell death, mitotic catastrophe to an end of cell replication following aberrant mitosis and senescence to an irreversible cell cycle arrest, induced because replication can no longer take place without damaging the DNA.

For all cases considered, senescence only played a secondary role. For non-radioresistant tumoral cells, apoptosis was clearly dominant for both irradiation type even though some of the cells died through mitotic catastrophe. On the contrary, radioresistant cells presented very few cases of apoptosis. For those kind of cells, carbon-ion caused far more death by mitotic catastrophe then x-rays, thus explaining the origin of the increased efficiency of the charged particles. The study went on and look at the effect of those two types of irradiations on several other cell lines, confirming the results aforementioned.

It should be noted that the term mitotic catastrophe covers all type of death that follows from aberrant mitosis. The type of death, in itself, can vary: while it can simply corresponds to an inability to replicate furthermore, other mechanisms can be provoked. Notably, necrosis is an usual outcome. While the particularities of this method of decay are not the fo-

cus here, the fact that it usually lead to local inflammation is noteworthy in the context of radiation therapy [77].

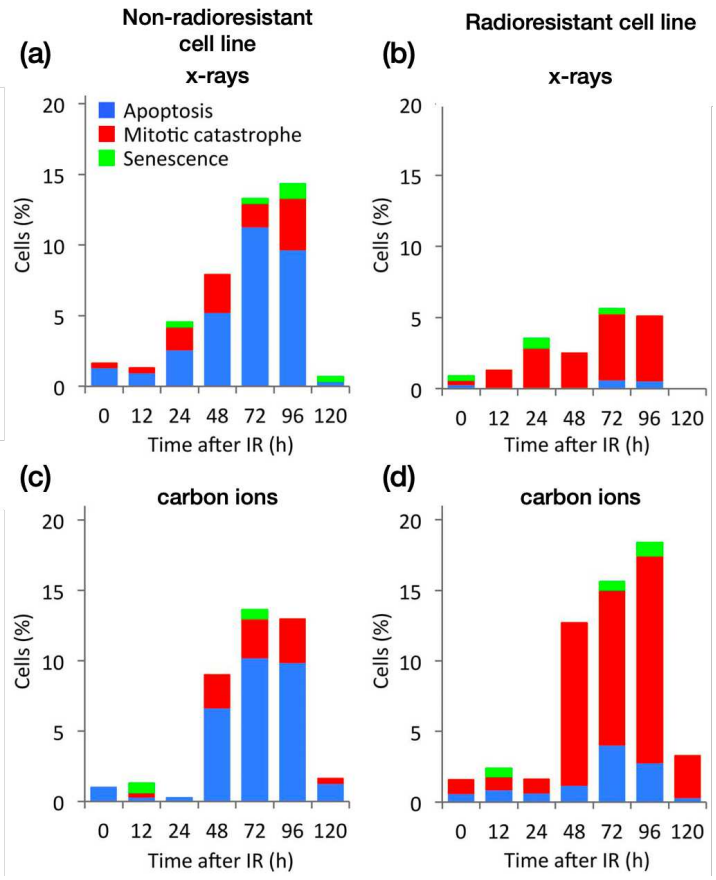

Figure 1.14: Percentage of cell death for each mode studied, in function of the time after irradiation. Apoptosis has been marked in blue, mitotic catastrophe in red and senescence in green. (a) corresponds to x-rays irradiation of non-radioresistant cells, (b) to x-rays irradiation of radioresistant cells, (c) to carbon-ion irradiation of non-radioresistant cells and finally (d) to carbon-ion irradiation of radioresistant cells. Taken from [76].

Hypoxia and oxygen enhancement ratio

The hypoxic conditions in which tumors grow is also a factor that leads to resistance to conventional radiotherapy. Indeed, when photons are used to kill tumoral cells, the impact of indirect damages to further increases the complexity of the lesions cannot be neglected. Thus, if oxygen is scarcely

present in the environment, the efficiency of the irradiation is decreased accordingly. [78].

The increased radioresistance is quantified by the oxygen enhancement ratio (OER) for a given endpoint:

$$OER(pO_2)|_{endpoint} = \left.\frac{D(pO_2)}{D_{normoxic}}\right|_{endpoint} \tag{1.31}$$

Here, $D_{normoxic}$ is the dose necessary to produce the endpoint in tissue with normal tissue oxygenation, while pO_2 defines the targeted level of hypoxia, and $D(pO_2)$ the corresponding dose necessary to reach the endpoint. Several in vitro studies have shown that the OER for x-rays is approximately 3. This is one of the main reasons why fractionation is used in radiotherapy: delivering part of the prescribed dose allows for reoxygenation of the tumoral tissues between irradiation, and therefore reduces the impact of hypoxia.

OER is known to decreases with increasing LET. It is related to the energy deposition pattern of the particle: an increased LET leads to more clustered DNA damages, that will result in cell death more effectively, even in the absence of indirect damages. This property is illustrated in figure 1.15.

In order to reach an OER of 1 in typical tumoral oxygen concentration, the LET of the incident charged particle need to be close to 200 keV.μm^{-1}. This means that the LET of carbion ions typically found in clinical situations is not enough to overcome hypoxia. It is therefore not surprising that multi-ion centers are looking at heavier ions such as oxygen as a mean to treat highly hypoxic tumors.

The relevant physical and radiobiological aspects of radiotherapy have now been established. In the following section, the need for cross-sections measurements in the context of hadrontherapy is addressed: nuclear reactions, leading to the consumption of the primary particles of the beam and to the production of secondary fragments should be taken into account while panning treatment, and can be used as a mean to monitor them.

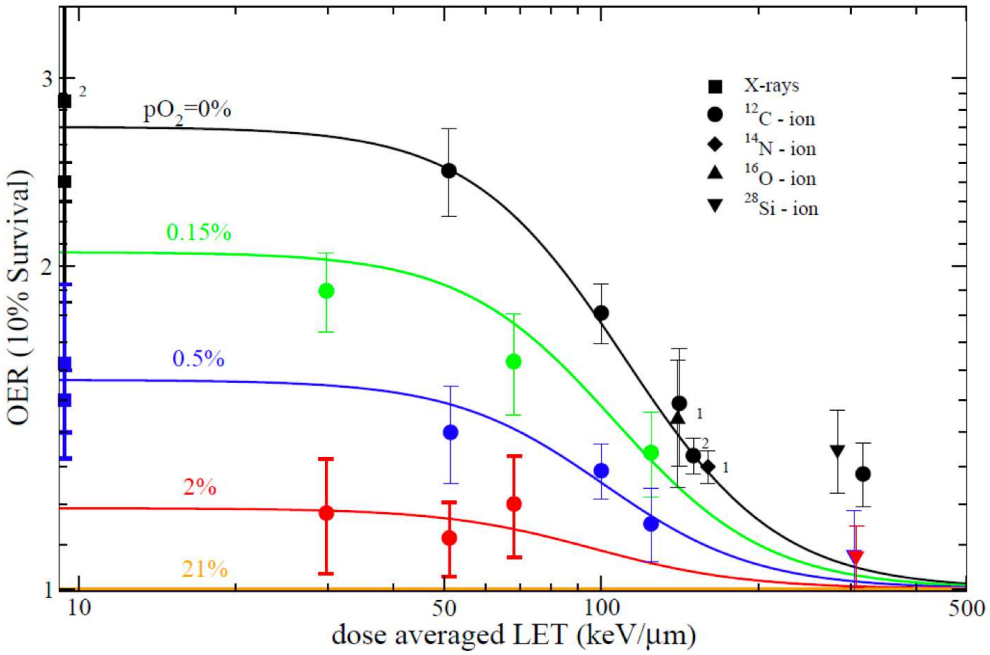

Figure 1.15: OER at 10% survival as a function of dose averaged LET. OER curves for several oxygenation levels are shown as well as experimental points coming from several studies. An oxygenation level of 21% corresponds to normal cellular conditions. Taken from [79].

1.4 The need for cross-sections in hadrontherapy

Treatment planning

The goal of treatment planning is to deduce beam energies, intensities and incidences needed to cover the tumoral volume with an optimal dose, or biological dose, while sparing healthy tissues as much as possible. The first step in this process is to delimit the targeted volumes based on imaging techniques, i.e. computed tomography (CT) scans, in combination with positron emission tomography (PET) or magnetic resonance imaging (MRI). In order to account for uncertainties from various sources, safety margins are added to the volume definition.

Volumes definitions and safety margins

First, the gross tumor volume (GTV) is delimited. It corresponds to the visible location of the tumor, and as such, should receive an adequate dose in order to obtain local tumor control. Second, the clinical target volume (CTV) is delineated: it takes into account microscopic extensions of the tumor around the GTV, at a certain probability level. Finally, the planning target volume (PTV) needs to be defined as well: this volume includes the CTV and safety margins needed to compensate for the various uncertainties related to treatment planning and delivery.

The overall setup of the treatment, be it patient or beam delivery, anatomical variations following previous irradiation, imaging techniques resolution and conversion to relative stopping power, dose calculation as well as biological considerations will results in uncertainties in the delivery of the beam.

Taking into account those various form of uncertainties is crucial in hadrontherapy: an improper quantification of safety margins can not only lead to under-dosage of the tumor but in dire circumstances part of the tumor could be left without any dose deposition. Moreover, decrease in NTCP can be achieved in part through the reduction of treatment volumes, and therefore of safety margins.

Treatment delivery

Treatment planning systems (TPS) should also take into account the way the beam is delivered to the patient, in order to adjust dose optimizations accordingly. Two principal modalities exist today in order to modulate the energy of the beam: passive beam modulation or active systems. The former essentially makes use of the stopping power of charged particles to modulate the energy of the beam: several absorber plates can be used to shift the depth of the SOBP; the latter actively modify the beam energy by varying the electric fields applied in the accelerator. Furthermore, how the beam is shaped in order to achieve lateral conformance to the tumoral volume should also be taken into account; nowadays, scanning, i.e. using magnets to bend the path of the beam, is most commonly used.

Dose optimizaton and biological models

Once the tumoral volume is defined, dose optimization can be performed. Usually for protons, only absorded dose is considered and an RBE of 1.0 to 1.1 is used [80]. For ion beams, it is the biological effective dose that should be optimized. Given the complexity of the radiation field and its

dependency on RBE, the task is far from easy: it requires knowledge of ion interactions with tissues, as well as energy spectra and dose distributions for primary and secondary ions.

The optimization of biological effective dose is the most time consuming part of the process, and is based on theoretical models that make use of *in vitro* and *in vivo* experimental data: historically, european platforms rely on the local effect model (LEM) [81], while japanese ones make use of the microdosimetric kinetic model (MKM) [82].

Traditionally, treatment planning softwares dedicated to those tasks were based on analytical methods, which are known to be fast, an important criterion in the clinical world. Nowadays, a shift towards the use of Monte-Carlo methods is apparent [83, 84].

The use of Monte-Carlo in treatment planning

The desire to move from analytical treatment plans to Monte-Carlo based ones has been flourishing in the recent years, even in the protontherapy field [85]. In order for this transition to be effective, without regard for the computational time involved with current Monte-Carlo methods, the impact of fragmentation modeling inaccuracies on dose distribution should be put under the spotlight. Indeed, several studies aimed toward this particular point [86, 87, 64, 88, 89].

Lühr et al. implemented an in-between treatment plan for carbon ion beams, where particle transport is simulated with the SHIELD-HIT package, while dose planning and radiobiology are computed by the TRiP software[90]. The idea was to produce a relative assessment of the effects of inaccurate fragmentation modeling on the plan. They computed several data sets to be used as input for the analytical treatment planning software, where they purposefully altered the hadronic models behavior. Six different cases were established:

- (A): classical description of nuclear reactions for the package
- (B): shut off of any hadronic interaction
- (C): inelastic nuclear cross-sections decreased by twenty percent
- (D): inelastic nuclear cross-sections increased by twenty percent
- (E): alteration of evaporation phase, resulting in larger and fewer fragments produced
- (F): alteration of evaporation phase, resulting in smaller and more fragments produced

In order to be as conclusive as possible, the study looked at both dose and relative dose optimization for a single fraction, for two different SOBPs, a large deep-seated one and a smaller one close to the surface. To be able to compare the results from the several Monte-Carlo data sets, dose and biological effective dose optimizations were first performed using the set corresponding to case A. From there, the resulting fluence distribution was used to calculate the dose or biological dose distribution for the other various sets, highlighting the impact of the different modeling of hadronic interaction, as can be seen in figure 1.16.

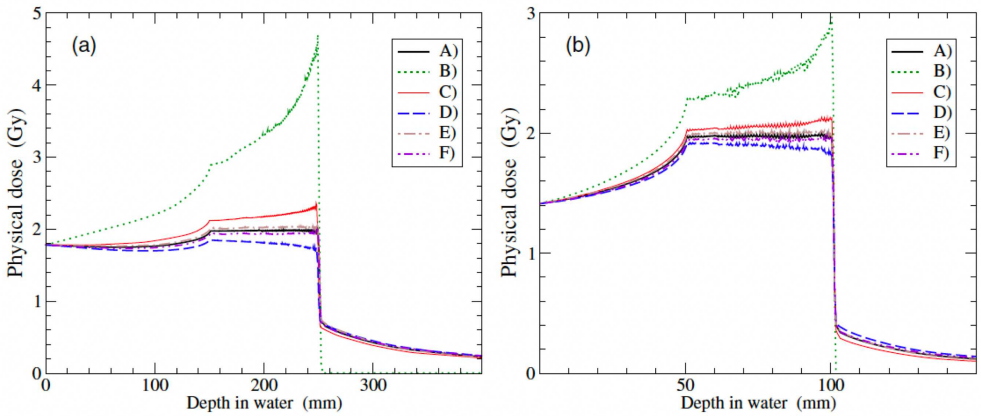

Figure 1.16: Dose distribution for a carbon ion beam in water, optimized to 2Gy according to case A. Two different SOBPs have been considered: (a) is a fairly large SOBP aiming for a deep-seated tumor, while (b) is a narrow SOBP for a targeted volume close to the skin of the patient, representing two extreme possibilities. The different curves show the impact of modified nuclear reaction models on dose distribution. Taken from [89].

The main finding of the study is that absorbed dose distribution can be highly dependent on the description of nuclear reactions. Indeed, in the worst case, i.e. the large deep-seated tumor, a difference as high as 15 percent higher (or lower) dose deposition was observed at the end of the SOBP. In all cases considered, while the effect is not as high, it cannot be neglected: it could lead to a severe under-dosage of the tumoral volume. As the authors stated in the paper, the values reported here are unlikely, because simple dose measurements in water could be use to partially correct them. Nonetheless, current medical prescriptions state that the total delivered dose in the PTV should be no more than 7% and no less than 5% of the prescribed dose [91].

Additionally, it should be pointed out that a simple geometry was used

in the study, with water representing the patient, which is not necessarily representative of actual treatment conditions. Finally, while the impact on biological effective dose is not as obvious, it should still be considered: the TRiP software makes use of the LEM, and as pointed out by *Fossati et al.*, there is no guarantees that the results would be the same if the microdosimetric kinetic model (MKM) was used[92] .

In order to improve the results from the Monte-Carlo simulations, and by doing so improving the treatment plans that would make use of it, more experimental measurements of cross-section is needed. Indeed, it would allow for a better constraining of the hadronic models handling the production of fragments.

Treatment monitoring

In order to reduce treatment delivery errors, *in vivo* range verification is a desirable property. However, it is not easily achievable in hadrontherapy, in contrast with conventional radiotherapy. Indeed, the beam does not exit the patient, and therefore, monitoring can only be based on the secondary particles produced through nuclear reactions of the primary ions.

Several techniques were indeed proposed over the years. Positron emission tomography (PET) and the detection of prompt gamma are two of major ones. The first relies on the production of positron emitting isotopes, such as ^{11}C and ^{15}O, while the second makes use of the gamma rays emitted by excited nuclei left in the beam track. For both of theses, the cross-sections for the production of the particles of interest were shown to be maximal close to the Bragg peak [93, 94].

However, both of those technologies require a comparison of the measurements with an expected signal to be used as a range verification method. Typically, this expected signal is computed through Monte-Carlo simulations. As explained above, and in this context as well, Monte-Carlo nuclear models currently fail to accurately reproduce the distribution of positron emitters [95] or the emission of prompt gammas [96, 97, 98, 99]. Again, cross-sections measurements are needed in order to further constrain the various hadronic models used in Monte-Carlo simulations.

1.5 Conclusion

After stating the global health burden cancer represents, we went on to understand what is the rationale behind the use of charged particles as a treatment modality. Their advantageous energy deposition pattern can

indeed be leveraged to increase dose deposited in tumoral volumes while sparing healthy tissues with respect to conventional radiotherapy. However, the ratio between cost and benefit is a major hurdle for this approach to be more generally used.

Therefore, hadrontherapy is currently only advisable for a few distinct types of cancer, which can make use of its properties. Pediatric cancers, cancer with complex geometries or radio-resistant cancers are among them.

Fragmentation incurred by nuclear reactions between particle of the beam and nuclei from the medium is a process of particular importance in hadrontherapy, which finds no equivalence in classical radiotherapy.

Indeed, nuclear reactions experienced by incident particles generate a mixed field of particles which need to be taken into account while planning treatment. In consequence, the use of Monte-Carlo simulations in treatment plans is the foreseeable direction the charged particle therapy field would like to take. Indeed, it would in theory improve the handling of the production of secondary particles in comparison with the current analytical methods.

Nuclear reactions can also be used as a mean to monitor treatment delivery, through technologies such as PET imaging or gamma prompt detection. However, both of these methods also rely on Monte-Carlo simulations and the hadronic models therein to be effective.

In consequence, both treatment planning and treatment monitoring will benefit from an accurate modeling of nuclear reactions from Monte-Carlo simulations. The measurement of cross-sections for reactions of interest is the only way to demonstrate the viability of the models, and, if need be, properly constrain them.

Therefore, measurements of double differential cross-sections for various reactions of interest, in the therapy energy range, is a crucial goal for the discipline. It is one of the main aim of the FragmentatiOn Of Target (FOOT) experiment.

Après avoir rappelé le fardeau que le cancer fait peser sur la santé publique, nous avons cherché à expliciter la raison d'être de l'utilisation des particules chargées comme modalité de traitement. Les caractéristiques de leur dépôt d'énergie peuvent en effet être exploitées pour augmenter la dose déposée dans les volumes tumoraux tout en épargnant les tissus sains par rapport à la radiothérapie conventionnelle. Toutefois, le rapport entre

le coût inhérent de cette modalité et les avantages apporté par cette technique constitue un obstacle majeur à une utilisation plus générale de cette approche.

Par conséquent, l'hadronthérapie n'est actuellement conseillée que pour quelques types de cancer distincts, qui bénéficient de ses propriétés. Les cancers pédiatriques, les cancers à géométrie complexe ou les cancers radiorésistants en font partie.

La fragmentation induite par les réactions nucléaires entre les particules du faisceau et les noyaux du milieu est un processus particulièrement important en hadronthérapie, qui ne trouve pas d'équivalent en radiothérapie classique.

En effet, les réactions nucléaires subies par les particules incidentes génèrent un champ mixte de particules qui doit être pris en compte lors de la planification du traitement. Par conséquent, l'utilisation de simulations de Monte-Carlo dans les plans de traitement est la direction prévisible que le domaine de la thérapie par particules chargées souhaiterait prendre. En effet, elle permettrait en théorie d'améliorer le traitement de la production de particules secondaires par rapport aux méthodes analytiques actuelles.

Les réactions nucléaires peuvent également être utilisées comme un moyen de contrôle du traitement, grâce à des technologies telles que l'imagerie TEP ou la détection de gammas prompt. Cependant, ces deux méthodes reposent également sur les simulations de Monte-Carlo et les modèles hadroniques qu'elles utilisent.

Par conséquent, la planification et le suivi du traitement bénéficieront d'une modélisation précise des réactions nucléaires à partir des simulations de Monte-Carlo. La mesure des sections efficaces des réactions d'intérêt est le seul moyen de démontrer la viabilité des modèles et, si nécessaire, de les contraindre correctement.

Par conséquent, la mesure des sections efficaces différentielles doubles pour diverses réactions d'intérêt, dans la gamme d'énergie thérapeutique, est un objectif crucial pour la discipline. C'est l'un des principaux objectifs de l'expérience FOOT (FragmentatiOn Of Target).

CHAPTER 2

The FOOT experiment

Contents

2.1 Rationale for the experiment

As mentioned in the earlier chapter, in section 1.2, the available cross-sections measurements, in the context of hadrontherapy, present a large gap in data. For carbon ions, the therapeutic energy range was mostly covered, but total and partial charged-changing cross-sections are, in most cases, the only available measurements. Some simple and double differential yields were also published, but only with few incident energies. Some data also exist for oxygen beams, but the energy range is not well covered, and

the cross sections are also limited to partial charge-changing in the best cases. Finally, helium beams, today seen as a viable therapy option in some cancer cases, were also investigated, but the corresponding measurements are sparse. While the energy range is well covered for those measurements, differential cross-sections are lacking.

Furthermore, one important topic in the context of hadrontherapy as mostly been left aside: the fragmentation of target nuclei, which is of particular importance when using proton beams. The fragments produced in this way have two distinct characteristics in comparison to the incident particles: a relatively low kinetic energy and a larger charge and mass number. Therefore, their LET is high and their RBE as well. Thus, they could have a non-negligible impact on treatment, and notably in the entrance channel of the beam, where healthy tissues are most often found [100].

As a consequence, the FOOT project was proposed to remedy to these missing measurements. The experiment was designed to extend and complete the projectile fragmentation cross sections for helium, carbon and oxygen beams, and also to provide measurements for target fragmentation. The international collaboration fixed two goals concerning cross-sections: isotopic double differential cross-sections for projectile fragmentation should be achieved with an accuracy better than 5% and differential cross-section with respect to kinetic energy for target fragmentation maximum uncertainty should not exceed 10%.

The measurement of the target fragmentation represents an experimental difficulty: due to their properties, the fragments originating from the target have a really short range in matter, and therefore the probability of them escaping from the medium is low. By reversing beam and target, in an inverse kinematic approach, this difficulty can be circumvented: the secondary fragments produced that way have a longer range, making detection easier. Because cross-sections are only dependent on the geometry of the two interacting nuclei, and since reversing projectile and target makes no difference in the center-of-mass energy, the measurements are equivalent.

However, in order to produce sufficient statistics, solid targets rather than gaseous ones need to be used. This implies the need for hydrogen enriched targets such as polyethylene C_2H_4, from which the cross-sections contribution from the C nuclei need to be subtracted. Therefore, a dedicated measurement on C targets need to be performed as well. Additionally, in order to reduce the impact of multiple coulomb scattering inside the target, as well as the probability of secondary fragmentation, thin targets must be used.

Given the large amount of data needed to be collected, being able to make use of the several facilities around the world that can deliver beams

of interest is necessary. In order to do so, the experimental setup needs to be easily movable, and must fit the space limitations of the various experimental and treatment rooms.

Furthermore, projectile fragments are forward peaked, and their angular distribution is wider with decreasing charge, as stated in section 1.2. Achieving the desired angular acceptance for all secondary fragments with a limited size apparatus and keeping the setup movable is therefore impossible.

The design of the experimental apparatus is a consequence of both of those points: the FOOT experiment is indeed the combination of two complementary setups. The first one is based on an emulsion spectrometer, optimized for the detection of fragments with low charge, emitted at large angles. The second one is an electronic setup based around a magnetic spectrometer, coupled with tracking detectors. It covers an angular acceptance up to 10° with respect to the beam axis, which encompass the expected angular distributions of most heavier fragment ($Z > 2$).

Given that the emulsion setup is not a part of the work reported in this document, only the electronic setup will be described in the following section. A detail description of the experimental setups can be found in [101].

2.2 The electronic setup

The electronic detector setup is composed of three separate regions: the upstream region, the magnetic tracking region and the downstream region. A global visualization of the experimental setup, extracted from the FOOT event display can be found in the figure 2.1.

Upstream region

The purpose of this region is to monitor the beam, determining its direction and count the number of ions crossing the target. It is also used to select the correct reconstructed vertex, or interaction point in the target. In order to reduce as much as possible the fragmentation probabilities outside the target, the thickness of the material crossed by the beam need to be minimized. The upstream region is therefore composed of two detectors: the Start Counter (SC), which is a thin plastic scintillator and the Beam Monitor (BM), which is a drift chamber. The upstream region also includes the target.

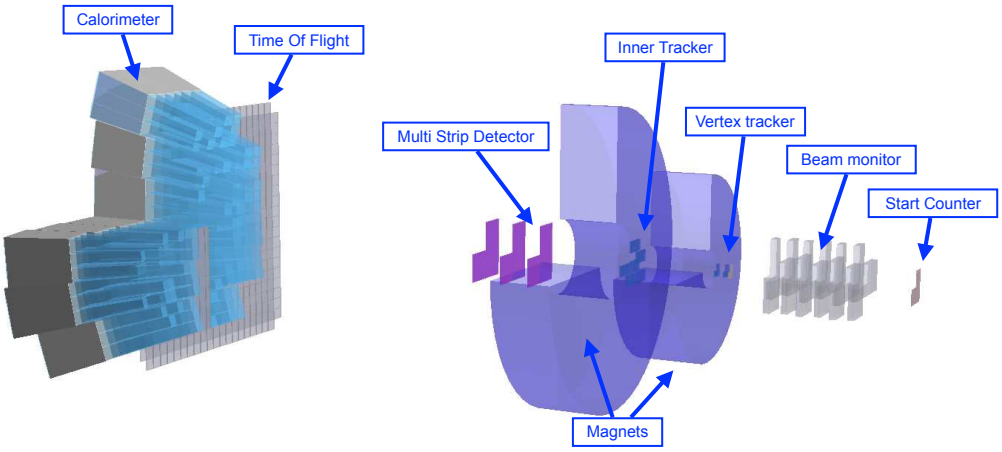

Figure 2.1: Schematic view of the electronic setup, obtained from the event display developed for the FOOT experiment. All the components used in this setup are show, from right to left: the start counter, the beam monitoring, the target, the vertex detector, the first permanent magnet, the inner tracker, the second permanent magnet, the multi-strip detector, the time-of-flight wall and the calorimeter.

Start Counter

The SC is made of a 250 μm thick foil of the EJ-228 plastic scintillator. The active surface of the detector has been designed to cover typical beam transverse size. It is contained within a 3D printed box fixed on an aluminum frame, and two entrance windows made of 10 μm thick aluminized mylar.

The light produced in the scintillator is collected on the sides by a total of 48 3×3 mm^2 SiPMs (AdvanSiD ASD-NUV3S). The output of the several SiPMs are read by group of 6, in a total of eight electronic channels. Both readout and powering of the SiPMs are handled by the WaveDAQ system [102].

This detector fulfills a few different roles for the experiment: it measures the incoming ion flux, with an efficiency $> 99\%$, provides the reference time for all the other detectors and finally is used to deduced the time-of-flight measurement in combination with the time-of-flight wall (TW).

Beam Monitor

The BM is a drift chamber made of twelve alternated layers. Each layer is composed of three drift cells. The cell area is delimited by eight 90 μm diameter aluminum field wires, and find at its center a 25 μm diameter gold-plated tungsten sense wire. Layers alternate between a horizontal and a vertical orientation of the drift cells. Two consecutive layers with the same orientation are staggered by half a cell to resolve left-right ambiguities in track reconstruction. The active area of the detector, for which the twelve layers can be exploited is close to 4×4 cm^2. Along the beam direction, the active length is 13 cm.

The BM operates with a gas mixture of Argon and CO_2 (80% and 20% respectively), at a relative pressure close to 0.9 bar. Two mylar windows, 100 μm thick, are used to contain the gas mixture inside the detector.

Each wire at the center of cells is polarized with a high voltage, which value is included in the 1850 to 2200 V, depending on the incoming beam. Those wires are also read out by a time-to-digital converter (TDC) board CAEN V1190B, which utilizes the SC trigger single as a reference time.

A BM efficiency close to 90% was evaluated for different combinations of ions beam and energies [103].

The BM detector plays a crucial role in order to disentangle pile-up ambiguity for the following tracking region. Indeed, it provides the direction of the beam as well as the interaction point inside the target, when used in combination with the vertex detector (VTX). It can also be used to discard events in which the beam has fragmented in the SC.

Magnetic tracking region

The tracking system of the FOOT experiment is based around three measuring stations, positioned before, in between, and after two permanent magnets. In order to conform to the global acceptance defined earlier, and to provide position measurements as accurate as possible while reducing the amount of material fragments have to cross, the first two detectors, VTX and the inner tracker (IT) are composed of silicon pixel sensors, while the last is a telescope of silicon micro-strips (MSD). A visualization of the tracking region can be found in figure 2.2, where the expected magnetic fields map are also visible.

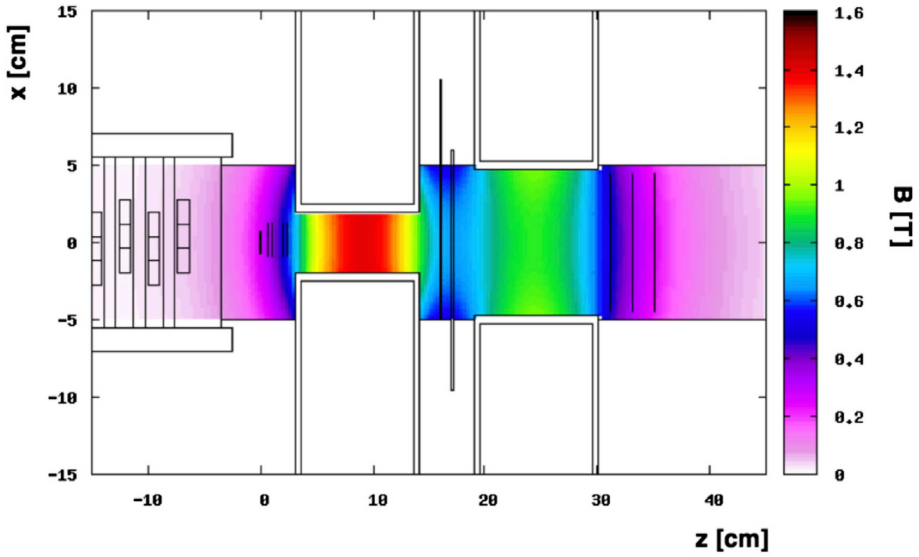

Figure 2.2: Detailed view of the magnetic region of the electronic setup for the FOOT experiment. The x coordinate is represented as a function of the depth in the setup, z. The BM, target, VTX, IT and MSD are visible, as well as the two permanent magnets. The magnetic field intensity applied by the magnets in this region is also visible. Taken from [101].

Magnetic system

The design of the magnetic system for the FOOT experiment was largely impacted by the need to produce a portable setup. Therefore, the magnetic system is kept in the air rather than being under a vacuum, and make use of two permanent magnets in Hallbach configuration, which allows for a detection station to be place in between.

In this configuration, an approximately dipolar magnetic field can be achieved inside the hole at the center of the cylindrical permanent magnet. Both magnets are envisioned for the experiment: they can respectively provide a maximum field of 1.4 T and 0.9 T at their center. Each magnet is composed of twelve units of Samarium-Cobalt, a material which retains its magnetic properties even in high radiation environments.

Because of the need to inter-align the several detectors composing the experimental setup, the mechanical structure holding the two magnets not only must withstand their weights and the magnetic forces they apply, but also, be able to displace the two magnets with respect to the beam line.

Vertex Detector

The VTX is itself composed of 4 different pixel sensor layers of 2x2 cm^2 active area. Their positioning, in close proximity to the target, allows for a geometrical acceptance up to 40°.

The several sensors are MIMOSA-28 (M28), from the family of CMOS Monolithic Active Pixel Sensors (MAPS), developed by the CNRS PICSEL group, Strasbourg []. They consist in a matrix of 928(rows) x 960(columns) pixels, with a pitch of 20.7 μm. The thickness of the epitaxial layer is 15 μm, for a total sensor thickness of 50 μm.

The sensor output is readout using a rolling shutter technique, with a frame readout time close to 190 μs. All the pixel of one row are read out in parallel, at the end of the column where a discriminator is placed. The data is then collected on a System on Chip (SoCKit) board, using a FGPA interfaced to both the sensors and the general DAQ control, through a DE10 Terasic board.

The VTX is used to achieve two goals: tracking of the fragments immediately after the target, as well as providing the point of origin of the fragments detected.

Inner Tracker

The IT detector is composed of two planes of M28 pixel sensors, to provide tracking of the fragments in between the two permanent magnets. They are arranged in a so-called plume fashion, inspired by the PLUME project [104]. The corresponding ladder are double-sided: two modules of 4 M28 sensors are glued on each side of a 2 mm thick supporting structure made of silicon carbide (SiC) foam. Each plane is composed of two ladders, for a total of 32 sensors. In order to reduce the dead area of the sensors, superimposing two ladders is a necessity. Overall, the active area for this detector is estimated to be around 8×8 cm^2.

Micro Strip Detector

After the two magnets, a final tracking detector is present, in the form of a 6 layers of single sided silicon detectors. Two consecutive layers have perpendicular orientation, allowing for the measurement of both coordinates in the plane orthogonal to the beam axis. Each sensor is thinned down to 150 μm. They are positioned 2 cm apart one another. Overall, they cover an active area of 9.6x9.3 cm^2, therefore fulfilling the 10° global acceptance requirement.

A strip pitch size of 50 μm was selected in order to reduce the fragment pile-up in the same strip. Coupled together with a readout with a pitch of 150 μm it allows for a spatial resolution of the order of a few tens of μm. Only one out of three strips are read by the acquisition. The analog signal provided by the strips are digitized by a 1MHz 12-bits ADC, and sent further down to the general DAQ.

The MSD allows for tracking of the fragments, and subsequent matching with the downstream detectors, and at term, could also be used to provide redundant measurement of $\mathrm{d}E/\mathrm{d}x$ for the fragments.

Downstream region

The distal part of the experimental setup is dedicated to the fragment charge and energy identification. It is composed of two detectors: the TW, which is formed by two walls of plastic scintillators and the calorimeter (CAL), made of crystal scintillators. Depending on the beam characteristics, this region is placed at least one meter away from the target.

Time-of-flight Detector

Two layers of 20 EJ-200 plastic scintillator bars compose the TW. They are wrapped with reflective aluminum and black tape. The layers are orthogonal two each other. Each bar is 2 cm wide, 44 cm long and 0.3 cm thick. The active area covered by the wall of scintillators is 40×40 cm^2.

Each edge of a bar is readout by 4 SiPMs, with a 3×3 cm^2 active area, for a total of 320 SiPMs. Each group of 4 SiPM provide a summed signal corresponding to a readout channel. The whole waveform are digitized and recorded at 3 to 4 Gsamples/s by the WaveDAQ system.

The TW plays a critical role in the experimental setup: it provides the measurement of the deposited energy ΔE, a final tracking point, as well as the final time for the event. The latter, when combined with the reference time from the SC, can be used to deduce the time-of-flight for a given fragment. This measurement, coupled with the energy loss, is needed in order to perform charge identification which will be explained in the following section 2.3.

Calorimeter

The CAL detector is the final detector composing the FOOT experimental setup. The calorimeter is an assemblage of 320 $Bi_4Ge_3O_{12}$ (BGO) crystals.

This material, with high density ($\rho = 7.1 g/cm^3$) allows for a high stopping power and a light yield of 10 photons/keV.

The crystals are arranged in a disk with a radius of approximately 20 cm. This is possible because of the truncated pyramid shape of each crystal, with a front face covering an area of 2×2 cm^2 and a length of 240 mm. Crystals are grouped by 9, in a 3x3 module, in order to ease the positioning and the weight distribution of the ensemble.

The readout of the crystals light is guaranteed by 25 SiPMs matrix, with an active surface of 2x2 cm^2. Those matrices are further interfaced with the WaveDAQ system, with a sample rate of 1Gsample/s. Temperature monitoring of the crystal, needed to compensate the response of the detector, is provided by the same board.

The calorimeter is dedicated to the measurement of the kinetic energy of the fragments. In the energy range used in particle therapy, the main mechanism of energy loss will be though electromagnetic interaction with the orbital electron of the crystal, and therefore a proper containment of the fragments can be achieved inside of the detector. However, several limits exists to the measurements performed by the calorimeter, notably, both neutron production and secondary intra-crystal fragmentation in the detector results in an underestimation of the fragment energy.

DAQ

Finally, the data acquisition system (DAQ) for the FOOT experiment should be addressed. The DAQ system indeed need to be able to process incoming data from various type of detectors and their associated acquisition. The system is based around three separate PCs: the first one is dedicated to the control of the run, the second one to its online monitoring, and the final one focuses on storage of the acquired data. Figure 2.3 illustrates the overall DAQ setup.

The supervision of the several detectors will be performed by a CAEN VME crate via 1Gbps Ethernet connections, which also host the TDC crate from the BM. Indeed, it is necessary to synchronize the response of the several detectors, through central signals, in order to construct a proper event.

The data collection is assured by a central switch which connects to the various sub-DAQ of the experiments, required in order to control the size and validity of the acquired data. Thus, the tracking system uses DE10 Terasic board to provide intermediary event formatting, zero suppression and data shipping. The SC, the TW and the CAL, are interfaced with

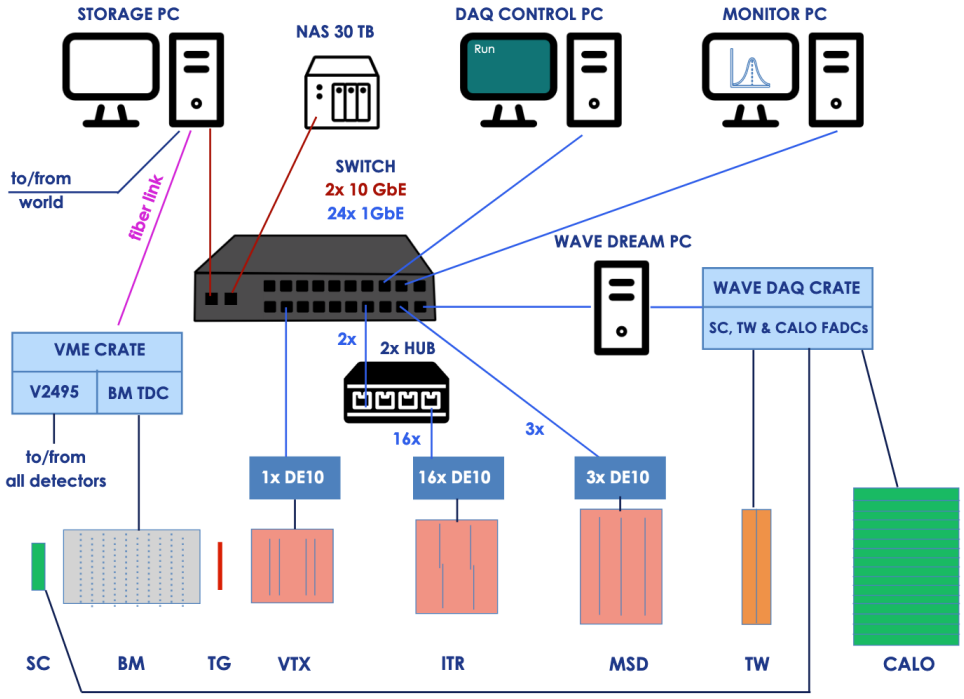

Figure 2.3: Schematic description of the DAQ for the FOOT experiment. The central switch links most components of the system: the three PCs, used for storage, control and online monitoring respectively; the NAS storage server; as well as the sub-DAQ for each detectors (the Terasic DE10 boards for the tracking system, the WaveDAQ for the time-of-flight system and the calorimeter). Only the VME crate, holding the BM TDC and the control board is not linked to the switched but directly to the storage PC. Taken from [105].

the WaveDAQ system, which provide data calibration, compression and shipping: the signals are digitized with a 16-bits resolution.

The overall readout rate for the system is limited by the slowest detectors, i.e. the pixel trackers. This results in an operating rate close to 1 kHz. The total expected event size is of the order of a few tens of kB, which storage is assured by SSD disks during data taking. The data is later on transferred to a network attached storage (NAS) system (30 TB) when the system is idle.

With this overview of the experimental setup, we can go on to understand how the collaboration aims to performs cross-section measurements by exploiting the information provided by each detector.

2.3 Cross-sections and fragments identification

In order to determine isotopic cross-sections with an uncertainty below 5%, a robust fragment identification need to be put in place. To do so, the collaboration aims to use redundant measurements in order to uniquely identify each fragment in term of charge number Z and mass number A.

Local reconstruction of the event and the need for a global reconstruction

To retrieve values for isotopic and elemental cross-sections, mass or charge distributions for the corresponding isotope or element needs to be obtained.

The charge identification, as mentioned above and developed later on in section 2.3, is performed mainly by the local reconstruction of the TW detector. For the mass identification, different identification techniques, based on basic relativistic equations, will be used [106], with U the modified Atomic Mass ($U = 931.5$ MeV):

$$A_1 = \frac{p}{U\gamma\beta c} \qquad A_2 = \frac{E_k}{(\gamma-1)Uc^2} \qquad A_3 = \frac{p^2c^2 - E_k^2}{2Uc^2E_k} \tag{2.1}$$

Therefore, the fundamental quantities providing those redundant mass identifications are the time-of-flight measurement τ_{tof}, the path of the particle through the apparatus l, its energy loss ΔE in the TW, the kinetic energy of the particle E_k and finally its momentum p.

Indeed, both β, the velocity of the fragment and γ the lorentz factor are dependent on τ_{tof} and l:

$$\beta = \frac{l}{\tau_{tof}c} \qquad \gamma = \frac{1}{\sqrt{1-\beta^2}} \tag{2.2}$$

Those three identifications share between them redundant information: they are correlated two by two by a common quantity. A_1 and A_2 share a dependency on τ_{tof} and l, A_2 and A_3 both make use of the kinetic energy of the fragment E_k, and finally A_1 and A_3 are dependent on the particle momentum p.

Some of the fundamental quantities necessary for mass identification here are directly available after local reconstruction of the data extracted from each detector. However, both the path of the particle l and its momentum p requires a more advanced processing, associating the information

produced by several detectors. This is known under the term 'global reconstruction' of the event.

Global reconstruction is indeed the core of the work presented in this document: the tracking of ejectiles algorithm (TOE), was developed towards this end for the FOOT collaboration. Its inner workings as well as its performances will be exposed in the following chapter.

Thus, in order to understand how the basic quantities taking a part in mass identification are determined, details about the local reconstruction of the different detectors should be addressed. However, the reconstruction of the kinetic energy of the fragment will not be examined: only the quantities necessary in order to establish the global reconstruction of the event will be developed.

The procedures and algorithms described in the following are grouped together in a common software suite called SHOE (Software for Hadrontherapy Optimization Experiment).

Time-of-flight reconstruction

The reconstruction of the τ_{tof} value is based around the information extracted from the SC and TW: the first provides the start time of the event, while the TW provides its end time. Those time evaluations have been performed by applying a digital constant fraction discriminator (CFD) to the two corresponding waveforms [107].

The idea behind this technique is to produce a constant time evaluation, no matter the shape of the waveform that is considered. Indeed, threshold methods are inherently dependent on the amplitude of the signals. The CFD algorithm proceeds as follows: first, the waveform is delayed, amplified and inverted; a bipolar signal is then obtained by summing the original and modified waveforms; the zero-crossing point of the resulting signal, through a linear interpolation, is then taken as the time evaluation for the waveform.

The 8 SC signals, formed by each 8 SiPMs groups, are summed for each event, and the resulting time evaluation is taken as the start time t_{start}. For the TW, each bar is read independently, and the corresponding time evaluation, t_{end} is obtained by using the contributions of both bar ends. However, since the SC is responsible for the reference time of the event, the phase difference between the clocks of the TW and SC need to be taken into account as well. Therefore, the time evaluation for the end of the event can be expressed under the following form:

$$t_{end} = \frac{t_1 - \Delta\phi_1 + t_2 - \Delta\phi_2}{2} \tag{2.3}$$

With 1 and 2 denoting each end of the bar, and $\Delta\phi$ the corresponding clock phase difference. In order to extract the value for the event time-of-flight for given TW bar, the two values produced this way are subtracted. However, an additional correction factor, Δt_j corresponding the possible experimental time offsets, notably coming from the propagation of signals through cables, need to be removed:

$$\tau_{tof,b} = t_{end} - t_{start} - \Delta t_j \tag{2.4}$$

Finally, it should be noted that the time-of-flight value required by the fragmentation identification only corresponds to a path length originating in the target and not the SC. Therefore, an additional correction factor $\Delta t_{primary}$, associated with the distance traveled by the incident primary ion between the Sc and the target, need to be taken out of this formula.

$$\tau^{(f)}_{tof,b} = \tau_{tof,b} - \Delta t_{primary} \tag{2.5}$$

Furthermore, the reconstruction of the position in the TW, by correlating the bars that are crossed by the fragment implies the availability of two values of t_end. The expression for the time-of-flight of a given fragment is therefore, with F, R denoting the front and rear bar respectively:

$$\tau^{(f)}_{tof} = \frac{\tau^{(f)}_{tof,F} + \tau^{(f)}_{tof,R}}{2} \tag{2.6}$$

Charge identification

The determination of the charge of the detected fragments is performed by making use of the Bethe-Bloch formula. This operation requires the knowledge, and therefore reconstruction, of both the τ_{tof} and ΔE_{TW} values [108].

The raw values for the energy loss in a bar is extracted from the TW waveforms: for each bar, the signals are collected independently, and the charge is then defined as the area of the signals. The total charge collected for any given bar is therefore, with 1 and 2 denoting each side of the bar:

$$Q_b = \sqrt{Q_1 Q_2} \tag{2.7}$$

This value corresponds to the raw energy loss of the incident particle inside a given bar. However, due to several factors such as material inhomogeneity or varying reflectiveness, it is not constant along the bar. Therefore, a mean charge value $< Q >$ for each bar was considered in energy calibration.

The energy calibration was performed according to a reference energy loss value ΔE_{MC}, obtained through Monte-Carlo simulations. Birks' law is used to model the behavior of the charge collected versus the energy loss. The two Birks' parameters, p_0 and p_1, are respectively the charge conversion factor and the saturation factor. Once this equation is established and the parameters are evaluated, it can be used to deduce energy loss values for a given bar b, for any charge measured:

$$\Delta E_b = \frac{< Q >}{p_0 - p_1 < Q >} \tag{2.8}$$

Once ΔE_b and τ_{tof} are acquired, the charge of the detected particle can be determined. To do so, the Bethe-Bloch formula is exploited: several parametrizations, denoted as $f_{BB}(\tau)$, relating the energy loss of a particle in a TW bar as a function of its time-of-flight are available, one for each charge number $Z < Z_{primary} + 1$.

For a given parametrization, the closest point to the measured one (defined by τ_{tof} and ΔE_b) is selected. For any time-of-flight value τ, the distance to the experimental point d is defined as:

$$d = \sqrt{(\tau_{tof} - \tau)^2 + (\Delta E_b - f_{BB}(\tau))^2} \tag{2.9}$$

Then, in order to determine the shortest distance between the curve parametrization and the experimental point, the range of possible time-of-flight values is iterated over, looking for an inflection point for $\partial d/\partial \tau$. In order to converge quickly to a solution, the bisection algorithm is used.

Once a minimum distance d is established for each given parametrization curve, the shortest one is selected among this set, therefore providing the corresponding charge identification.

The figure 2.4, left, shows a two-dimensional graphic of ΔE_b as a function of τ_{tof} reconstructed as specified above, for a set of simulated data using a 200 MeV/u oxygen ion beam impinging on a C_2H_4 target. The several parametrized curves used for charge identification are also shown.

Tracking system local reconstruction

The tracking system is at the center of the global reconstruction of the event: reconstruction of the particle trajectory would be virtually impossible without the information it provides, spread out along the experimental setup.

Individually, each detector composing the tracking system (VTX, IT, MSD, TW) provides reconstructed clusters, corresponding to the interaction point of particles that are detected.

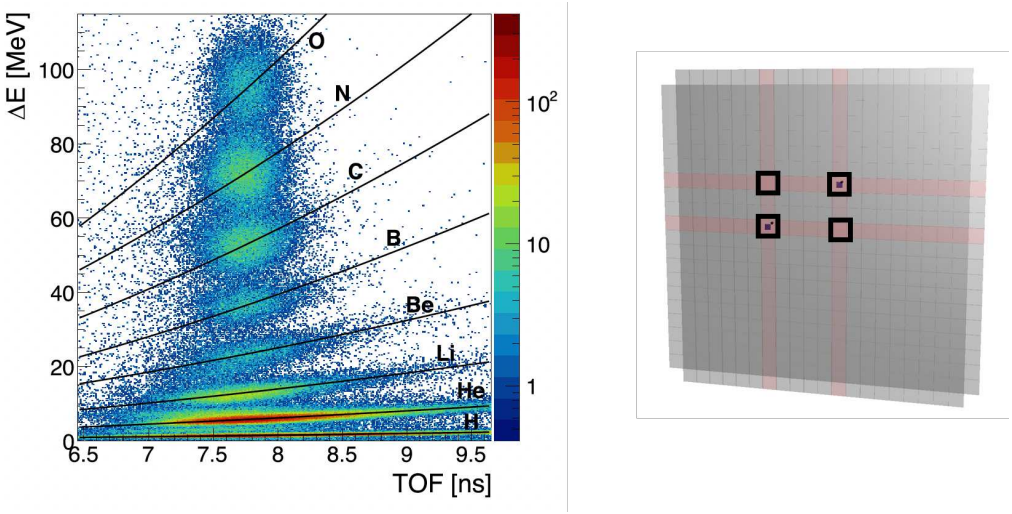

Figure 2.4: Left, ΔE as a function of τ_{tof} distribution for a simulated data set of a 200 MeV/u ^{16}O beam impinging on a 5mm C_2H_4 target. The different parametrizations of the Bethe-Bloch curve, one for each $Z \leq Z_{beam}$ are also shown. Taken from [101]. Right, event display for the TW corresponding to an event where the fragment spatial configuration can lead to a potential reconstruction of ghost clusters. The fired bars are shown in red, small rectangles show the reconstructed clusters, and the open black rectangles corresponds to the reconstruction possibilities for the clusters.

However, those clusters represent only part of the information on a particle trajectory. Furthermore, several particles can be produced in a fragmentation event; consequently, several clusters on the same detection layer can be reconstructed. Additionally, both experimental conditions such as electronic noise, as well as detector design choices can lead to the reconstruction of fake or ghost clusters. Therefore, the task of regrouping the clusters belonging to a same trajectory and extracting from it the parameters of interest can be difficult. A robust local reconstruction from the several detectors composing the tracking system is of utmost importance.

VTX, IT and MSD cluster reconstruction During data taking, any charged particle crossing a detection layer from the tracking system loses some of its energy inside the medium. This energy deposition results in one or several fired pixels or strips. However, the information on which pixels or strips was fired by the same incident particle cannot be retrieved directly, and need to be reconstructed afterwards.

As of now, this local reconstruction of the clusters from the pixels or

strips that are fired inside of their respective detectors share a same basic approach: for any given fired pixel or strip, an iterative search on its first neighbors is applied. All of the direct neighbors that are also fired are added to the cluster under construction. Subsequently, they are the origin point for a new first neighbor search, until either no more fired pixels or strips are found, or the cluster size reaches a user defined limit.

It should be noted that electronic noise can artificially augment the size of the clusters, if left unchecked. Therefore, for both pixel detectors, a hardware threshold is used to reduce the amount of pixels fired due to noise. For the MSD, an equivalent threshold is applied in software.

The reconstructed clusters have therefore their positions determined by the centroid of all of the fired pixels or strips. In consequence, the clusters for the pixel detectors are two dimensional, while the clusters for the MSD are one dimensional.

TW cluster reconstruction The TW is also subject to the local reconstruction of clusters within its two detection layers. However, because of its particular role in the experimental setup, both τ_{tof} and the charge identification need to be attached to the reconstructed cluster. Moreover, since it is the final position that can be extracted from the tracking system, the reconstructed cluster is two dimensional, even though each detection layer only provides one dimensional information. Its correct reconstruction is delicate: because multiples fragments can be produced during an event, the issue of ghost hits arise.

This can be understood by considering two fragments produced in the target and detected in the TW. If the only information used to reconstruct the cluster is the physical position of each bar, then a total of four different positions can be deduced as valid, without any way to differentiate between them, as can be seen in figure 2.4, right.

In order to avoid this situation, the signals collected at each end of any given bar can be leveraged. The time difference between the two signals τ_1 and τ_2, coupled to the velocity of light inside the bars c_{TW} allows us to derive the position x_b along any given bar:

$$x_b = \frac{(\tau_1 - \tau_2)c_{TW}}{2} \tag{2.10}$$

The determination of this value allows for a better selection of the corresponding bar in the other layer of the TW, and thus reduces the amount of ghost clusters that could be produced.

It should be noted that ghost hits is not the only factor that can lead to a poor reconstruction of a TW cluster: two fragments, emitted in close

proximity, could potentially reach the same bar on a given layer. This would results in a signal from the bar that is unusable. To circumvent this difficulty, the choice has been made to always start the local reconstruction with the layer tallying the highest number of fired bars. In this way, the two other reconstructed values in the TW, τ_{tof} and the charge identification can be correctly associated with the cluster.

VTX tracklets and vertex reconstruction: The VTX also has a special status when considering global reconstruction: its function is not limited to the reconstruction of the position of the impact points in its four layers.

Indeed, because it is directly placed behind the target, it is only slightly affected by the magnetic field of the experiment. Therefore, the trajectories of the fragments in the region covered by this detector are straight. This property can be leveraged to reconstruct tracklets from the reconstructed clusters, and those tracklets can be further used to reconstruct the point of interaction inside the target, the vertex [109].

The reconstructed vertex can then be used in the global reconstruction as an extra point to perform trajectory reconstruction; the tracklets can be used to deduce some of the parameters necessary to start the global reconstruction procedure.

Two different algorithms can be used to perform the local reconstruction of the tracklets. Both of them start the reconstruction with the final layer of the VTX.

The first one makes use of the track direction reconstructed locally in the BM: it is used to deduce an approximate vertex in the tranverse plane of the target. For each cluster in the final layer of the VTX, a straight track is drawn towards this vertex. The closest clusters to this track, in each following layer (in reverse order), are added as candidate for the tracklet reconstruction, and the parameters of the straight track are adjusted accordingly. Once every layer has been inspected, a final check is applied: the track thus formed must be composed of at least three clusters.

The second algorithm only uses information provided by the VTX: rather than looking at the BM track direction, it explores every two clusters combination between the last layer of the detector and the two in the middle. Each possible combination forms a candidate track, except for those which do not find their origin in the target. For each valid combination, the algorithm looks into the closest clusters to the track for the remaining layers. Again, once a cluster is added to a candidate track, the track parameters are re-adjusted accordingly. The same cluster count check is

applied to the final track candidates.

It should be noted that in order to reduce the amount of combination and reduce the computation time, for both those algorithm, once a cluster is selected by a track, it is marked as so and is no longer available for future track candidates.

The local reconstruction of the vertex is done through an algorithm scanning the target thickness and looking for a point which maximize the probability of several tracks crossing the target at this position.

For each track reconstructed, a Gaussian probability tube is constructed, with a width defined by the dispersion of the track at a point $\mathbf{v}$:

$$f_i(\mathbf{v}) = exp\left(-\frac{1}{2}(\mathbf{v}-\mathbf{r_i})^T V_i^{-1}(\mathbf{v}-\mathbf{r_i})\right) \tag{2.11}$$

where $\mathbf{r_i}$ is the point of closest approach to $\mathbf{v}$ for the track i, and V_i is its covariance matrix, which is diagonal and contains the position uncertainty of the track in x, y and z coordinates.

For each couple of tracks a and b, a temporary vertex, corresponding to the point inside the target that maximizes the product between $f_a(\mathbf{v_t})$ and $f_b(\mathbf{v_t})$ is constructed. A minimal probability value is required for those temporary vertices, all of those that do not match it are excluded.

Afterwards, the maximization of the function $Q(\mathbf{v})$ allows for the selection of the best temporary vertex using the complete set of reconstructed tracks:

$$Q(\mathbf{v}) = \sum_i^N f_i(\mathbf{v}) - \frac{\sum_i^N f_i^2(\mathbf{v})}{\sum_i^N f_i(\mathbf{v})} \tag{2.12}$$

From there, the complete set of tracks is again inspected: all the tracks which point of closest approach with respect to the selected vertex is below a certain value are marked as originating from this vertex. The final positions and errors for the selected vertex are finally recomputed using this ensemble of tracks.

An important point, in order to reconstruct both local tracklets in the VTX and full trajectories during global reconstruction, was left aside: those two procedures requires the application of an preceding alignment procedure. Indeed, the VTX on its own is a set of four CMOS sensors, which need to be precisely aligned in order to achieve a high spatial resolution of the reconstructed tracklets. A proper alignment between detector is also mandatory in order not to degrade the performances of the global reconstruction algorithm used.

Therefore, this alignment procedure allows for the determination and subsequent correction, in software, of the individual translational and rotational shifts of the sensors, unavoidable during the setup of the detectors. The procedure developed in [110] was integrated in the SHOE framework to remedy this issue.

Cross-section determination and performance study

Once the quantities necessary to produce the several mass identifications evoked in equation 2.1 are evaluated, the respective yields of each isotopes for a given charge can be extracted from the corresponding mass distribution.

This distribution is fitted with a superposition of Gaussian functions, one for each isotope considered, using a standard χ^2 minimization. Each individual yield Y_f is extracted from the deduced parameters, as well as its mass resolution.

From those yields, the differential cross-sections for a given fragment f, with respect to its emission angle θ or its production kinetic energy E_k are defined as:

$$\frac{\mathrm{d}\sigma_f}{\mathrm{d}\Omega} = \frac{(Y_f(\theta) - Bkg_f)^U}{N_{primary} N_t \Omega_\theta \varepsilon} \qquad \frac{\mathrm{d}\sigma_f}{\mathrm{d}E_k} = \frac{(Y_f(E_k) - Bkg_f)^U}{N_{primary} N_t \Omega_E \varepsilon} \tag{2.13}$$

For both of those equations, $N_{primary}$ is the number of incident particles, N_t the number of available interaction center in the target, Ω_θ and Ω_E the angular and energy phase spaces, and ϵ the reconstruction efficiency. The misidentification of fragments, in both charge and mass is represented by the value $Bkgd_f$. The U denotes the application of an unfolding procedure, to correct the distributions from experimental effects.

In an effort to derive limiting resolution values of the quantities leading to fragment identification, a Monte-Carlo performance study, using the FLUKA toolkit, was performed on a 200 MeV/u ^{16}O beam impinging on a 5 mm thick C_2H_4 target. In order to achieve cross-section determination within the uncertainties defined by the collaboration, this procedure is required. Thus, momentum, time-of-flight, kinetic energy resolutions impact on mass resolution were inspected.

In this study, the experimental resolutions of the quantities necessary for particle identification were included in the simulation by means of Gaussian smearing. Figure 2.5 reports the evolution of the mass resolution for all fragments with charge Z = 6, as a function of the kinetic energy resolution, the momentum resolution and the time-of-flight resolution respectively. For

each graphic, the other two resolutions were fixed to a default value. Those values were respectively: $\sigma(E_k)/E_k = 1.5\%$, $\sigma(p)/p = 3.7\%$ and $\sigma(\tau_{tof}) = 70$ ps.

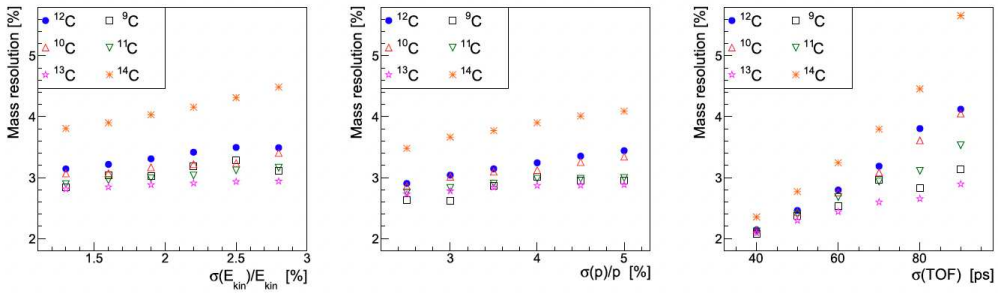

Figure 2.5: Mass resolution as a function of $\sigma(E_k)/E_k$, $\sigma(p)/p$ and $\sigma(\tau_{tof})$ on the left, middle and right, respectively. The data used for this study have been obtained through a Monte-Carlo simulation (FLUKA) of a 200 MeV/u ^{16}O beam impinging on a 5mm thick C_2H_4 target. The mass resolutions for all fragments with $Z = 6$ are shown. Taken from [101].

From figure 2.5, it immediately arises that the parameter impacting the mass resolution the most is the time-of-flight resolution. Indeed, in comparison, the mass resolution is a slowly varying function of both the kinetic energy and momentum resolution. However, in order to obtain results with an uncertainty included within the limits fixed by the collaboration, relatively harsh resolutions must be obtained for each parameter:

- $\sigma(p)/p \in [3, 5]$ %
- $\sigma(TOF) \in [70, 120]$ ps
- $\sigma(E_k)/E_k \in [1.5, 2.5]$ %

2.4 Conclusion

The FOOT experiment final goal is to provide the isotopic differential cross-sections data that are missing for most combinations of beam and target type in the therapeutic energy range, within a 5% uncertainty.

The experiment is composed of two movable setups, aiming to cover a large angular acceptance as well as the largest variety of fragments possible. The components of the electronic setup, for which the work presented in this document was developed, were described in further details.

In order to achieve cross-sections measurements within the limits fixed by the collaboration, a robust fragment identification needs to be used.

This isotopic identification relies on three different techniques, that make use of redundant information. In order to exploit those methods, a few fundamental quantities need to be evaluated from the measurements. Among them are the time-of-flight value, the kinetic energy of the fragment, its charge identification, its path along the experimental setups and finally its momentum.

While some of those values can be directly extracted from the local reconstruction of the detectors, the reconstruction of both the path of the fragment and its momentum requires a more advanced reconstruction, called global reconstruction, which correlates the information obtained from the entirety of the experimental setup. The TOE algorithm, exposed in full details in the following chapter, is a global reconstruction method aiming for track recognition and reconstruction, developed for the FOOT experiment.

L'objectif final de l'expérience FOOT est de fournir les données de sections efficaces différentielles isotopiques qui manquent pour la plupart des combinaisons de type de faisceau et de cible dans la gamme d'énergie thérapeutique, avec une incertitude de 5%.

L'expérience est composée de deux dispositifs complémentaires, visant à couvrir une large acceptation angulaire ainsi que la plus grande variété de fragments possible. Les composants du dispositif dit "électronique", pour laquelle le travail présenté dans ce document a été développé, ont été décrits plus en détail.

Afin d'obtenir des mesures de sections efficaces dans les limites fixées par la collaboration, une identification robuste des fragments doit être mise en place. Cette identification isotopique repose sur trois techniques différentes, qui utilisent des informations redondantes. Afin d'exploiter ces méthodes, un certain nombre de quantités fondamentales doivent être évaluées à partir des données produites par les différents détecteurs. Parmi elles, on trouve la valeur du temps de vol, l'énergie cinétique du fragment, l'identification de sa charge, sa trajectoire le long du dispositif expérimental et enfin sa quantité de mouvement.

Alors que certaines de ces valeurs peuvent être directement extraites de la reconstruction locale des détecteurs, la reconstruction de la trajectoire du fragment et de son impulsion nécessite une reconstruction plus avancée, appelée reconstruction globale, qui corrèle les informations obtenues à partir de l'ensemble du dispositif expérimental. L'algorithme TOE, exposé en

détails dans le chapitre suivant, est l'une de ces méthodes. Il vise à la reconnaissance et la reconstruction des trajectoires des fragments produits, et a été développé spécifiquement pour l'expérience FOOT.

CHAPTER 3

Global reconstruction with the TOE algorithm

Contents

As exposed in the earlier chapter, the reconstruction of both the path and the momentum of any detected fragment requires correlation between

the measurements coming from the several detectors composing the experimental setup. The tracking system is the centerpiece of this advanced form of reconstruction: without the position measurements it provides, spread out along the experimental setup, an accurate track reconstruction is impossible. Both the charge identification and time-of-flight reconstruction also plays a role in the global reconstruction, as will be seen later on.

Any recorded event can be characterized by the multiplicity of the produced fragments, and therefore, several reconstructed clusters can be found in each detection layer from the tracking system. The role of the global reconstruction is dual: first, the selection of the clusters belonging to a same track, representative of a fragment trajectory inside the experimental setup; second, the extraction of the parameters of interest from the corresponding reconstructed track. Those two distinct roles are respectively called track recognition and track fitting.

It should be pointed out that the measurements are impacted by experimental conditions: electronic noise, dead pixels or strips alters the quality of the local reconstruction, and can even lead to the reconstruction of non-physical clusters. The detector design can also lead to the reconstruction of ghost clusters, which are also non-physical. The global reconstruction method should be able to perform its task even in the presence of such measurements.

3.1 Track recognition

The goal of track recognition is to find the reconstructed clusters, formed by local reconstruction of the detectors, that belong to a same given trajectory. The method used to perform this selection must be developed and designed in accordance with the needs and characteristics of the experiment.

The TOE algorithm recognition procedure adopts a combinatorial track following approach: for any fragment under inspection, an history of the possible trajectories it could have taken is registered. This is possible only because the fragment multiplicity expected for any given event is relatively low (below ten fragments per event). Indeed, a higher number of piled-up events would lead, by definition, to a far higher fragment multiplicity for any given event, where the combinatorial approach would be both wasteful and impracticable in terms of computing time and efficiency.

The track recognition procedure of the TOE algorithm starts with the definition of so-called endpoints for the event under reconstruction. Those endpoints correspond to the possible final destinations for the tracks that will be reconstructed. They are also the basis for the static hypothesis

necessary to perform the reconstruction. For each of those endpoints, a different combinatorial history is established. Each history starts from a set of initial seeds, from which the trajectory is extrapolated to the following detection layers, which requires a parametrization of the track for any point in the experimental setup, called the track state. The initial seeds are considered in the overall procedure as dynamic hypotheses. As new suitable candidate measurements are found in each detection layer, they are added to the current history for the track under reconstruction, after a filtering and selection phase. They are then used to refine the trajectory. This process goes on until the endpoint corresponding to the static hypothesis is reached. This sequence of steps is characteristic of track following recognition procedures.

The track recognition procedure in the TOE algorithm was designed around three intertwined components that form the basis of the algorithm: the propagator, which is dedicated to the track state propagation from one detection layer to the next; the filter, which is used to compare the extrapolation of the track parameters to the available measurements and correct the track state accordingly; and finally the arborescence, or the memory of the procedure, which holds the possible trajectories for the track under reconstruction. Most of the design choices for the conception and implementation of the TOE algorithm originates from the form taken by the propagation model, which is itself a consequence of the design of the FOOT experiment and its inhomogeneous magnetic field.

Propagation model

In general, the propagation phase of any track following procedure can be performed by means of a simple trajectory modeling such as polynomials or helices. However, the use of those simple models requires a constant magnetic field, leading to a constant curvature for any trajectory. In the case of an heterogeneous magnetic field, the situation is more complex. Therefore, the choice was made to base the propagation model on the basic physical equation given by the Lorentz force.

Indeed, no other additional physical effects was included in the model used for propagation. The decision to excludes both multiple Coulomb scattering and energy loss was made in order to reduce the computing time necessary to perform a reconstruction. Even without the implementation of those terms, the algorithm can be shown to perform well (see 3.2).

Based on the selected system of coordinates, the equation of motion given by the Lorentz force can take several forms: it can be expressed as a function of time t, of the path length $s(t)$, or more simply as a function

of the depth in the setup z [111]. The latter was selected as the system of coordinates to be used by the TOE algorithm. This choice can be explained by the design of the experimental setup: each detector is positioned at a given depth, and the reconstructed clusters give position information on the plane orthogonal to the beam axis; therefore, performing the propagation through the z variable is the evident choice.

The derivation of the equation of motion used by the TOE algorithm starts from the classical form taken by the Lorentz force, with m the mass of the particle, γ its Lorentz factor, q its charge, $\mathbf{v}$ its velocity and $\mathbf{B}$ the magnetic field:

$$\gamma m \frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = q\mathbf{v} \times \mathbf{B} \tag{3.1}$$

By introducing the equality $\mathbf{v} = \mathrm{d}\mathbf{r}/\mathrm{d}t$, and using the chain rule, we get:

$$\gamma m \frac{\mathrm{d}z}{\mathrm{d}t}\frac{\mathrm{d}}{\mathrm{d}z}\left(\frac{\mathrm{d}z}{\mathrm{d}t}\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}z}\right) = q\frac{\mathrm{d}z}{\mathrm{d}t}\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}z} \times \mathbf{B} \tag{3.2}$$

If we consider $\mathrm{d}l = \sqrt{\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2}$ and let $R = \mathrm{d}l/\mathrm{d}z$, then with $\mathrm{d}l/\mathrm{d}t = v$ we get $\mathrm{d}z/\mathrm{d}t = v/R$ which leaves us with the following equation:

$$\gamma m \frac{\mathrm{d}}{\mathrm{d}z}\left(\frac{v}{R}\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}z}\right) = q\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}z} \times \mathbf{B} \tag{3.3}$$

If we now drop the vector form and compute each coordinate separately, we reach:

$$\gamma m \left(\frac{\mathrm{d}}{\mathrm{d}z}(\frac{v}{R})\begin{vmatrix}\frac{\mathrm{d}x}{\mathrm{d}z}\\ \frac{\mathrm{d}y}{\mathrm{d}z}\\ 1\end{vmatrix} + \frac{v}{R}\begin{vmatrix}\frac{\mathrm{d}^2x}{\mathrm{d}z^2}\\ \frac{\mathrm{d}^2y}{\mathrm{d}z^2}\\ 0\end{vmatrix}\right) = q\begin{vmatrix}\frac{\mathrm{d}y}{\mathrm{d}z}B_z - B_y\\ B_x - \frac{\mathrm{d}x}{\mathrm{d}z}B_z\\ \frac{\mathrm{d}x}{\mathrm{d}z}B_y - \frac{\mathrm{d}y}{\mathrm{d}z}B_x\end{vmatrix} \tag{3.4}$$

The z coordinate allows us to establish the equivalence:

$$\gamma m \frac{\mathrm{d}}{\mathrm{d}z}(\frac{v}{R}) = q\left(\frac{\mathrm{d}x}{\mathrm{d}z}B_y - \frac{\mathrm{d}y}{\mathrm{d}z}B_x\right) \tag{3.5}$$

By introducing it into the x and y coordinates, we finally obtain the expression for the equation of motion used by the TOE algorithm, with the norm of the momentum for the particle $p = \gamma m v$:

$$\begin{matrix}\frac{\mathrm{d}^2x}{\mathrm{d}z^2}\\ \frac{\mathrm{d}^2y}{\mathrm{d}z^2}\end{matrix} = \frac{qR}{p}\begin{pmatrix}\frac{\mathrm{d}x\mathrm{d}y}{\mathrm{d}z^2}B_x - \left[1 + \frac{\mathrm{d}x}{\mathrm{d}z}^2\right]B_y + \frac{\mathrm{d}y}{\mathrm{d}z}B_z\\ \left[1 + \frac{\mathrm{d}y}{\mathrm{d}z}^2\right]B_x - \frac{\mathrm{d}x\mathrm{d}y}{\mathrm{d}z^2}B_y - \frac{\mathrm{d}x}{\mathrm{d}z}B_z\end{pmatrix} \tag{3.6}$$

This final form is a system of two second order differential equations. Those equations rely on a few common variables: $\mathrm{d}x/\mathrm{d}z$, $\mathrm{d}y/\mathrm{d}z$, $\mathbf{B}(x, y)$ and the factor q/p. The dependency on the magnetic field can be rewritten as a dependency on the position x and y. The track state s, which is both the propagation entry point and result, directly ensue from those observations:

$$s = \begin{bmatrix} x \\ y \\ \dfrac{\mathrm{d}x}{\mathrm{d}z} \\ \dfrac{\mathrm{d}y}{\mathrm{d}z} \end{bmatrix} \tag{3.7}$$

In order to perform the propagation of any given particle inside the experimental setup through the propagation model described above, a value for the factor q/p as well as initial values for the track state need to be established.

At this preliminary point of the reconstruction, the only viable option to provide these values is to guess them, based on the information provided by the local reconstruction of the detectors. Therefore, hypotheses are formulated: first, a set of static hypotheses, which holds various q/p values; second, a set of dynamic hypothesis which corresponds to the possible initial track states.

Hypotheses

Static hypotheses

The reconstructed clusters from the final layer of the tracking system, i.e. the TW detector, are used to formulate the set of static hypotheses to be used by the recognition procedure. Any given static hypothesis regroups a combination of information invariant along the trajectory of the particle: the charge of the particle cannot be modified unless an extra nuclear reaction takes place along the path of the fragment, a situation which will prevent reconstruction; if we disregard the losses in energy inside the setup, then the norm of the momentum of the particle can be considered as constant as well; additionally, the endpoint of the particle is also fixed. Indeed, since it is what is used a basis for the formulation of the hypothesis, the reconstructed track must end its trajectory in the vicinity of this endpoint.

Therefore, the reconstructed quantities in the TW detector are harnessed: the position will provide the final endpoint for a given history and

Charge	Nucleon number
1	1, 2, 3
2	3, 4
3	6, 7
4	7, 9, 10
5	10, 11
6	11, 12, 13
7	14, 15
8	15, 16
Z > 8	2 Z

Table 3.1: Guess for the number of nucleons for a fragment with a given charge, used by the TOE algorithm in the formulation of a set of static hypotheses.

the reconstructed charge will be used as the basis for the determination of the q/p factor.

However, the determination of an initial guess for the norm of the momentum p presents a difficulty: none of the measurements performed in the TW detector are directly suitable to establish such a value. Thus, the charge identification is once again used: it is used to deduce a value for the number of nucleons of the corresponding fragment. As of now, the most probable isotopes for any given charge are considered, and are indexed in table 3.1.

The number of nucleons thus determined, the norm of the momentum can be initialized by taking into account the following conjecture: given the acceptance of the experimental setup, the fragments collected are mostly emitted in the forward direction, therefore, their respective energy by nucleon should be close to the one of the beam, as was explained in section 1.2.

Therefore, using the equation:

$$pc = \sqrt{E_k^2 - 2E_k mc^2} \tag{3.8}$$

which relates the norm of the momentum with the kinetic energy E_k of the particle and its mass m, a value for the norm of the momentum can be obtained, which depends on the number of nucleons determined in the previous step.

It should be pointed out that the momentum distribution for the lighter fragments are broader than the rest. In turn, this can impact the reconstruction efficiency of the procedure: if the norm of the momentum of the

particle is too far apart from its initial guess, the reconstruction might fail because the propagation model does not corresponds to reality. In order to avoid those scenarios, an additional static hypothesis with half the norm of the momentum with respect to the beam is added to the pool for fragments identified with a charge Z = 1.

Thus, each reconstructed cluster in the TW detector will give rise to possibly several static hypotheses. Each of those are added to a pool of static hypotheses for the event. Each hypothesis from the pool will lead to a different combinatorial history, and will go through the recognition procedure individually.

However, in order to reduce the computation time of the method and avoid unnecessary and redundant reconstruction, the hypotheses from the pool are actually regrouped through charge identification: rather than only considering one endpoint per static hypothesis, all the clusters which have been identified with the same charge are all considered as viable endpoints. Thus, once the final layer is reached, they are all checked against the final position for the track, which averts costly recomputations of the same trajectory, in case the selected endpoint was not the right one.

Dynamic hypotheses

The dynamic hypothesis, as stated before, corresponds to the initial values of the state. In order for the recognition procedure to converge towards an optimal solution, a reasonable first guess is required.

The local reconstruction done in the VTX is all that is needed to compute the values necessary for initialization. Indeed, the definition of the state of the particle requires values for x and y coordinates, as well as the initial track slopes in the corresponding directions, and their associated errors:

$$s = \begin{bmatrix} x \\ y \\ t_x = \dfrac{dx}{dz} \\ t_y = \dfrac{dy}{dz} \end{bmatrix}, \qquad S = \begin{bmatrix} \sigma_x^2 & 0 & 0 & 0 \\ 0 & \sigma_y^2 & 0 & 0 \\ 0 & 0 & \sigma_{t_x}^2 & 0 \\ 0 & 0 & 0 & \sigma_{t_y}^2 \end{bmatrix} \tag{3.9}$$

The definition of initial values for the x and y coordinates and the corresponding errors is dependent of the nature of the reconstructed vertex:

- If the vertex is composed of several tracklets, the position of the vertex can directly be used as the seed to start the propagation

- In case the vertex is made of only of tracklet, its position is not reliable, and therefore the cluster from the first detection layer of the VTX detector is used instead.

Initial values for the track slopes are deduced from the tracklets composing the corresponding reconstructed vertex. However, when several tracklets are available all of them could belong to the particle currently under reconstruction. Therefore, each distinct tracklets formed by the local reconstruction will lead to a new dynamic hypothesis. Each of those will be added to the current combinatorial history as a new start point for the propagation.

Once the several hypotheses are set up, based on the information provided by the VTX and TW detectors, the propagation model can be used to establish a prediction of the track state in the following detection layer. This is the role performed by the propagator, which is described in the following section.

Differential equation solver: the propagator

As seen in section 3.1, the propagation model used to reconstruct the trajectory of particle along the experimental setup is in reality a system of two ordinary differential equations of second order. Because there is no analytical solutions for these particular equations, a numerical solver is required in order for them to be integrated.

The solution proposed for the TOE algorithm was designed with modularity in mind: it is not particularly specific to the considered equation, and can indeed solve various differential equations of either first or second order. On top of that, several options are available regarding the precision of the integration; on a computational stand point, both fixed and adaptable stepping are ready for use. The main configuration used in the TOE makes use of both a high precision method and an adaptive step size. In consequence, those points will be developed into further details in this part.

One of the usual way to solve ordinary differential equations numerically is to apply a Runge-Kutta (RK) method. Here, the differential equations to solve take the general form:

$$\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} = f(t, y, \frac{\mathrm{d}y}{\mathrm{d}t}) \text{ with } y(t_0) = y_0,\ \frac{\mathrm{d}y}{\mathrm{d}t}(t_0) = \left.\frac{\mathrm{d}y}{\mathrm{d}t}\right|_0 \tag{3.10}$$

Therefore, they belong in a category of differential equations for which specialized RK methods were developped: the general Runge-Kutta-Nyström

(GRKN) methods. GRKN methods usually make use of an embedded formula pair that will produce two different integration for the same step; this can be later leveraged to produce and adaptive step size. The solution chosen here is known as GRKN-5(6) and was proposed by Fehlberg [112]. Usually, the various factors used in RK methods are gathered inside a Butcher array; the one for the GRKN-5(6) method can be found in appendix A.

Being at the center of the reconstruction procedure, an important part of the computational cost is spent performing propagation. Accordingly, a careful construction of the different pieces making this propagator is required. It is currently divided in three different sub-classes: the evaluator, which will compute different evaluations of the function to integrate; the solver, which will produce a solution to the differential equation for the current step, and finally the stepping policy which lays out how a step unfolds.

Evaluator

GRKN methods first stage is based on the production of a determined amount of evaluations of the second derivative, given by the differential equation to integrate. These evaluations, for the step n, can be computed as such:

$$f_0 = f\left(t_n, y_n, \left.\frac{\mathrm{d}y}{\mathrm{d}t}\right|_n\right) \tag{3.11}$$

$$f_i = f\left(t_n + \alpha_i h, y_n + \alpha_i h \left.\frac{\mathrm{d}y}{\mathrm{d}t}\right|_n + h^2 \sum_{j=0}^{i-1} \gamma_{ij} f_j, \left.\frac{\mathrm{d}y}{\mathrm{d}t}\right|_n + h \sum_{j=0}^{i-1} \beta_{ij} f_j\right) \quad \forall i \in [1, m] \tag{3.12}$$

Here, h corresponds to the step size, while the various factor α_k, β_{kl} and γ_{kl} are specific to the GRKN method in utilization. The method implemented in the TOE algorithm uses eight different evaluations of the second derivative, i.e.: $m = 8$.

Figure 3.1 illustrates the relation between evaluations and the solution of the differential equation. However, a simpler case is used: a first order differential equation, solved by a classical Runge-Kutta method. In this configuration, the four evaluations, denoted as f_k with $k \in [0, 3]$, correspond to estimations of the function slope at t_n, $t_n + h/2$ and $t_{n+1} = t_n + h$. The solution for the step, computed from those evaluations, is indicated by the point (t_{n+1}, y_{n+1}).

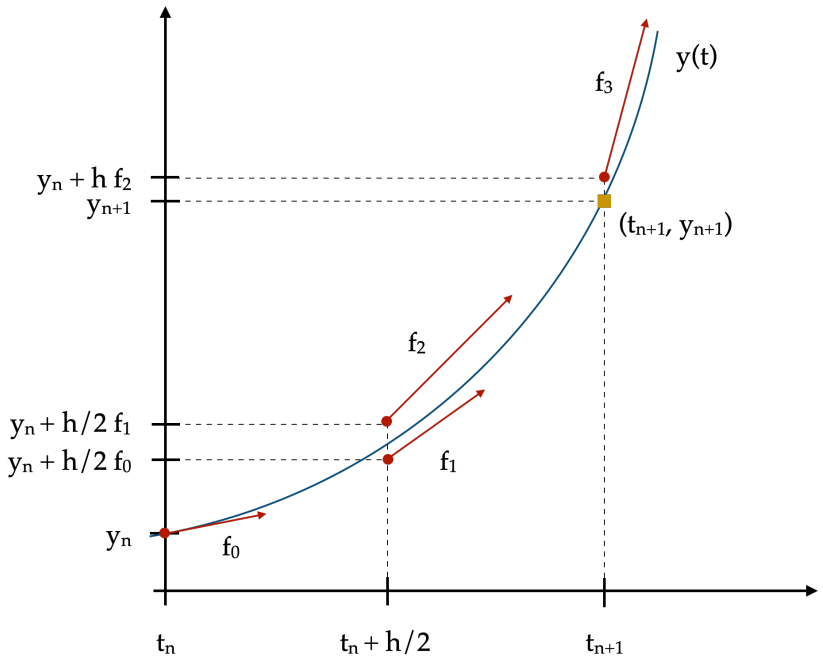

Figure 3.1: Illustration of the evaluation step of the classical Runge-Kutta method. A simple step between t_n and $t_{n+1} = t_n + h$ is considered. The analytical result of the integration of the differential equation, $y(t)$, is shown in blue. The four different evaluations of the differential equation are shown in red. The solution obtained by combining them corresponds to the yellow square indicated by the pair of values (t_{n+1}, y_{n+1}).

Solver

The solver is meant to compute the solutions to the differential equation. It must produce as well a solution for the first derivative of the function, if a second order differential equation is used. Those solutions are determined using the different evaluations computed at the last step, weighted by factors given by the method in use, as can be seen in the following formulas:

$$y_{n+1} = y_n + h \left.\frac{\mathrm{d}y}{\mathrm{d}t}\right|_n + h^2 \sum_{i=0}^{m} w_i f_i \tag{3.13}$$

$$\hat{y}_{n+1} = y_n + h \left.\frac{\mathrm{d}y}{\mathrm{d}t}\right|_n + h^2 \sum_{i=0}^{m} \hat{w}_i f_i \tag{3.14}$$

$$\left.\frac{\mathrm{d}y}{\mathrm{d}t}\right|_{n+1} = \left.\frac{\mathrm{d}y}{\mathrm{d}t}\right|_n + h \sum_{i=0}^{m} \dot{w}_i f_i \tag{3.15}$$

The computation of two different solutions for a step is a feature that allows for the setup of an adaptive step size as well as for a reduction of the global truncation error linked to the numerical integration: individually, the two solutions are respectively of fifth and sixth order, meaning that their truncation error for a step, τ_n, is on the order of $\mathcal{O}(h^6)$ and $\mathcal{O}(h^7)$. In other terms, for a method of order p, a constant C exist such that:

$$\tau_n < Ch^{p+1} \tag{3.16}$$

By combining those two solutions, a lower truncation error for the step can be achieved, as described in the following section.

Stepping policy

Defining a stepping policy is necessary in order to handle either a fixed or adaptive step size, based on the method in use. An adaptive step size brings in two major advantages over its fixed size counterpart: first, it allows a finer control over the overall numerical error linked to the integration; second it can result in a huge decrease of computation time.

In the case of a second order differential equation, the solver for the GRKN-$q(q+1)$ method computes a pair of solutions y_{n+1} with a truncation error of order $\mathcal{O}(h^{q+1})$, and $\hat{y}_{n+1}$ with a truncation error of order $\mathcal{O}(h^{q+2})$ at $t_{n+1} = t_n + h$. If the local error $e_n = y_n - \hat{y}_n$ is superior to the tolerance $\mathcal{T}$ specified by the user, then the step length h will be modified according to the following equation [113]:

$$h = 0.9 \left(\frac{\mathcal{T}}{e_n} \right)^{\frac{1}{q+1}} \tag{3.17}$$

This step will be repeated as long as the tolerance is not cleared. This in turn ensure a global numerical error of the order $N_{step}\mathcal{T}$.

Furthermore, the use of an adaptive step size can result in a drastic reduction of computation cost: if the solution to the differential equation is slowly varying with respect to t, then the step size can be increased without any consequences on the quality of the integration. As an example, a comparison between two numerical integrations of the differential equation $f"(x) = -2(f'(x) + f(x))$, by the propagator developed for the TOE algorithm is shown in figure 3.2. The exact solution of this equation is $f(x) = exp(-x)\cos(x)$. The methods used are GRKN-4 and GRKN-5(6) with a fixed an adaptive step size, respectively. The GRKN-5(6) method is able to perform the integration over the range $-2 < x < 2$ in a total of eight steps, while the GRKN-4 takes forty.

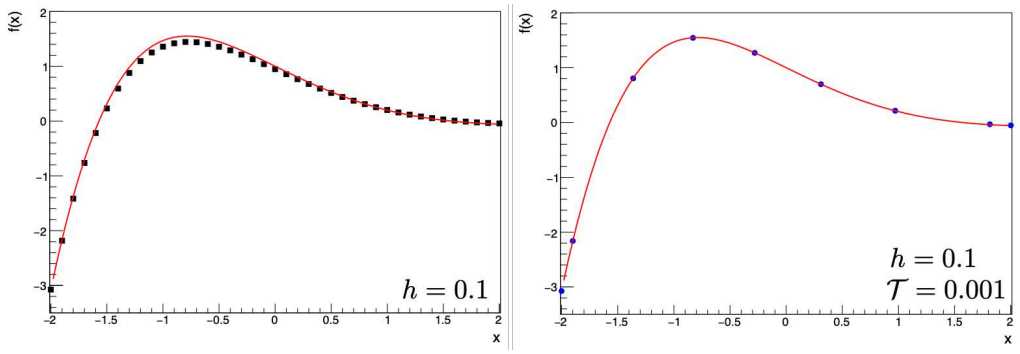

Figure 3.2: Two numerical integration of the same differential equation $f''(x) = -2(f'(x) + f(x))$ performed by the propagator developed for the TOE algorithm, in the range $-2 < x < 2$. For both of them, the solution corresponds to the red line. On the left, the classical GRKN-4 parametrization is used, with a step size $h = 0.1$, for a total of 40 steps. On the right, the GRKN-5(6) method performs the integration: from an initial step size $h = 0.1$ and a tolerance $\mathcal{T} = 0.001$, the integration is done in 8 steps.

The propagator role is to solve the system of differential equations given by the propagation model. This process is controlled by another part of the TOE algorithm: the filter.

Correction engine: the filter

One of the main characteristics of a track following procedure lies in the inclusion of the information provided by the several measurements (reconstructed clusters), available at each detection layer, into the propagation. Filtering techniques based on Kalman filters are a natural answer to solve this particular problem, and have made their mark as such over the years [114, 115].

Indeed, in such a procedure, as the clusters are evaluated sequentially, the predictive power of the propagation is improved by making use of past knowledge. The difficulty then lies in finding appropriate set of seeding clusters, which is a task undertaken when establishing the dynamic hypotheses.

Conceptually, a Kalman filter aim is to estimate the state of a system given noisy input information and a model for the observed process that can be stained by error. To do so, the filter evaluate towards which one its confidence lies the most: this is known as the gain of the filter. The gain computation is a direct consequence of the minimization of the mean square error between the actual state of the system and its estimation. This is why

the Kalman filter is considered as the optimal estimator in the minimum mean-square-error sense [116].

The principal limitation of this technique lies in the prerequisite for the model used: it must be linear. If this assumption falls short, then other filters could provide a better solution. This also means that in order to be used as the correction engine in the case of propagation in a magnetic field, the basic formalism need to be adapted.

Several solutions exists, the principal one being the so-called Extended Kalman Filter (EKF). While pretty commonly used, it requires the use of Jacobian matrices to handle the nonlinearity in the system. The favored solution used in this work makes uses of a statistical linearization technique called the unscented transformation (UKF) [117]. The changes that this particular formalism brings will be exposed in section 3.1.

Kalman filter basic formalism

The Kalman filter is constructed around two alternating phases: a prediction of the state is computed using the model, and is then corrected using the measurements. This basic procedure is summed up in figure 3.3: the variables involved into each phases are indicated, as well as the resulting state vector and covariance.

Let's assume the real state of a track can be described at a point k by vector a $\mathbf{s}_k$ and its associated covariance S_k. Given the configuration of our system, the discrete set of points k corresponds to the several detection layers.

The recursive application of the prediction and correction phases can begin starting from an initial state vector estimate $\hat{\mathbf{s}}_0$, and initial state covariance matrix $\hat{S}_0$, given by a particular dynamic hypothesis. The state at the point k is predicted using the following equations, with F_{k-1} the propagation model, $\mathbf{w}_{k-1}$ the modeling error and W_{k-1} its associated covariance:

$$\hat{\mathbf{s}}_{k|k-1} = F_{k-1}\hat{\mathbf{s}}_{k-1|k-1} + \mathbf{w}_{k-1} \tag{3.18}$$

$$\hat{S}_{k|k-1} = F_{k|k-1}\hat{S}_{k-1|k-1}F^T_{k|k-1} + W_{k-1} \tag{3.19}$$

The correction phase can be applied next: let $\mathbf{m}_k$ and M_k be the respective measurement vector and its covariance at point k. They are linked to the state to estimate through the equation:

$$\mathbf{m}_k = H_k\mathbf{s}_k + \mathbf{e}_k \tag{3.20}$$

With $\mathbf{e}_k$ the measurement error and H_k the observation matrix, which is dependent on the type of measurement performed on the current detection layer.

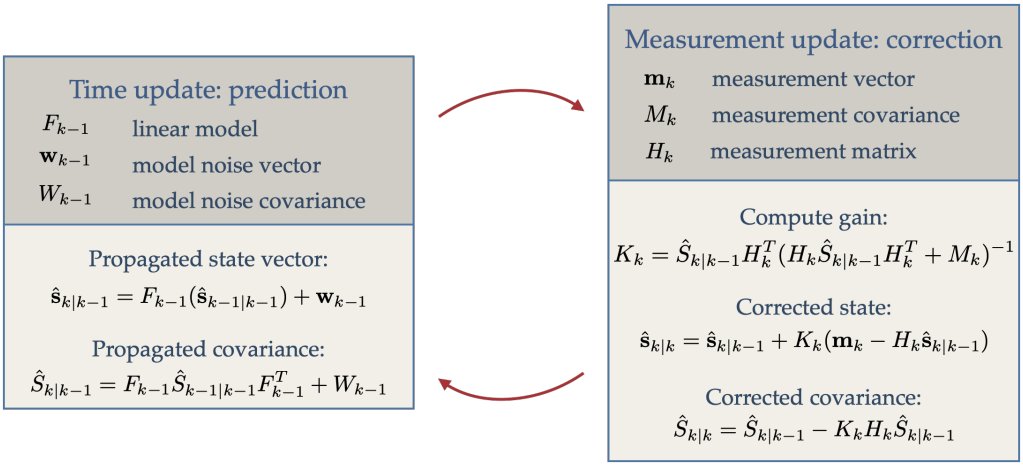

Figure 3.3: Overview of the Kalman filter basic formalism with two alternating phases: the time update delivers a prediction of the state to be tracked by the filter, while the measurement update provides a correction to the state estimation, using the measurement given has an input. The gain of the filter, computed from the estimation and the measurement, establishes the balance in trust of the filter between the model it uses and the measurements that are provided.

The gain of the filter is computed as such:

$$K_k = \hat{S}_{k|k-1}H_k^T(H_k\hat{S}H_k^T + M_k)^{-1} \tag{3.21}$$

It is then used to compute the corrected state according to the two following equations:

$$\hat{\mathbf{s}}_{k|k} = \hat{\mathbf{s}}_{k|k-1} + K_k(\mathbf{m}_k - H_k\hat{\mathbf{s}}_{k|k-1}) \tag{3.22}$$

$$\hat{S}_{k|k} = \hat{S}_{k|k-1} - K_kH_k\hat{S}_{k|k-1} \tag{3.23}$$

A basic example of the usefulness of the Kalman filter can be found in figure 3.4. Two different configurations of a ball thrown into a gravitational field are presented. For both of them, the state to determine is $\mathbf{s} = [y, z, \mathrm{d}y/\mathrm{d}t, \mathrm{d}z/\mathrm{d}t]$, in consequence, the position y of the ball as a function of its depth z is represented.

In the first case, the measurements and the model taken by the Kalman filter match. The prediction and correction are never far from each other, showing that the filter quickly converges to an optimal solution.

In the second case, an erroneous model is given as input to the Kalman filter. Therefore, the filter quickly learns to distrust its own model: after correction, the state is always by far closer to the measurements than the

prediction. Furthermore, the corrected state follows the real trajectory of the ball closely.

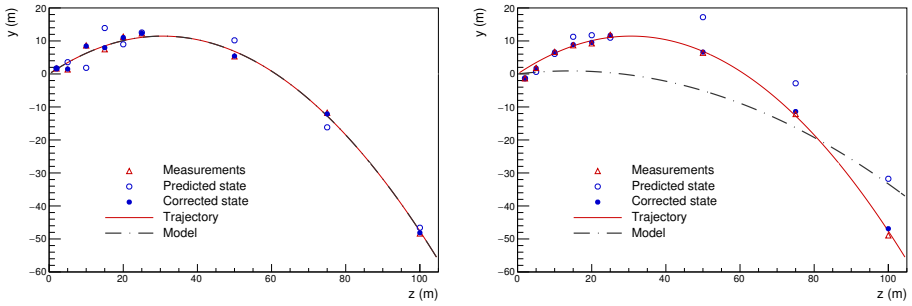

Figure 3.4: Usage of the classical Kalman filter for a ball thrown in a gravitational field, along the $y - z$ plane. In both cases, the measurements provided to the filter corresponds to the red hollow triangles, the predicted state to the blue hollow circles and the corrected state to the blue full circles. The model used by the filter is indicated by the dotted black line, and the real trajectory of the ball corresponds to the red line. On the left, the model used by the filter and the measurements both corresponds to the real trajectory of the ball. The model uses the conventional value for the gravitational constant: $g = -9.81$ m/s^2. The initial values also corresponds to the measurements: $v_{y0} = 15$ m/s and $v_{z0} = 20$ m/s. On the right, the model used by the filter do not correspond to the trajectory of the ball: $g = -4.91$ m/s^2, $v_{y0} = 3$ m/s, $v_{z0} = 23$ m/s.

Unscented Kalman Filter

The main aim behind the use of the unscented transform is to approximate how the probability distribution of the state is modified when the propagation model is applied to it. This basic principle contrasts with the idea behind the EKF which tries to linearize the propagation model in itself around the last prediction point.

In order to do, a set of $2N + 1$ sigma points, with N the dimension of the state vector, is selected in order to capture the mean and covariance of the underlying probability distribution. In a sense, this modified filter is similar to a particle filter but without the need to employ a large number of points drawn randomly to reach meaningful results.

The unscented transform is applied to the prediction phase of the classical Kalman filter. The first step of this modified time update is to determine the set of sigma points, which will represent the probability distribution of the state. A weight w_n, with $n \in [0, 2N]$ is assigned to each point. The

set is computed from the last corrected state, according to the following equations:

$$\boldsymbol{\mu}_0 = \hat{\mathbf{s}}_{k-1|k-1}, \text{ with } w_0 \in [0,1) \tag{3.24}$$

$$\boldsymbol{\mu}_i = \hat{\mathbf{s}}_{k-1|k-1} + \sqrt{\frac{N}{1-w_0}} A_i, \; w_i = \frac{1-w_0}{2N} \tag{3.25}$$

$$\boldsymbol{\mu}_{i+N} = \hat{\mathbf{s}}_{k-1|k-1} - \sqrt{\frac{N}{1-w_0}} A_i, \; w_{i+N} = \frac{1-w_0}{2N} \tag{3.26}$$

Here, $i \in [1, N]$ and A_i denotes the ith row of the matrix square root of estimated covariance of the state: $AA^T = \hat{S}_{k-1|k-1}$. The sigma point $\boldsymbol{\mu}_0$ corresponds to the mean value of the state. Its associated weight w_0 is set to 0.3 in the TOE algorithm.

Afterwards, the propagation model is applied to each sigma point through the propagator. Implementation wise, it implies the need for a way to force the step length of the numerical integration, even in the case of an adaptive step length: each sigma points should follow the same step size. Therefore, normal step operation is applied to $\boldsymbol{\mu}_0$; for each step decided on, the other sigma points are forced to follow the same step size.

Once the set of transformed points is obtained, i.e. $\boldsymbol{\nu}_n = f(\boldsymbol{\mu}_n)$ the resulting propagated state vector and covariance are deduced:

$$\hat{\mathbf{s}}_{k|k-1} = \sum_{n=0}^{2N} w_n \boldsymbol{\nu}_n \tag{3.27}$$

$$\hat{S}_{k-1|k-1} = \sum_{n=0}^{2N} w_n (\boldsymbol{\nu}_n - \hat{\mathbf{s}}_{k|k-1})(\boldsymbol{\nu}_n - \hat{\mathbf{s}}_{k|k-1})^T \tag{3.28}$$

At this point, the time update is completed, and the basic formalism can be used again.

A mock example illustrating such a procedure can be found in figure 3.5. An hypothetical one dimensional state is considered. Since the corresponding model is non-linear, the basic Kalman filter cannot be used. Therefore, three sigma points are picked from the probability distribution of the state, and propagated through the non-linear model. The resulting set of points is then used to infer the displaced probability distribution. Proceeding to the classical correction phase from there is straightforward.

So far, the formalism of both the classical KF and the UKF has been reviewed: based on a set of measurements, the two alternating phases will follow one after the other. The time update phase, which control the propagation, produces an estimation of the track state, which will be refined by measurements in the correction update.

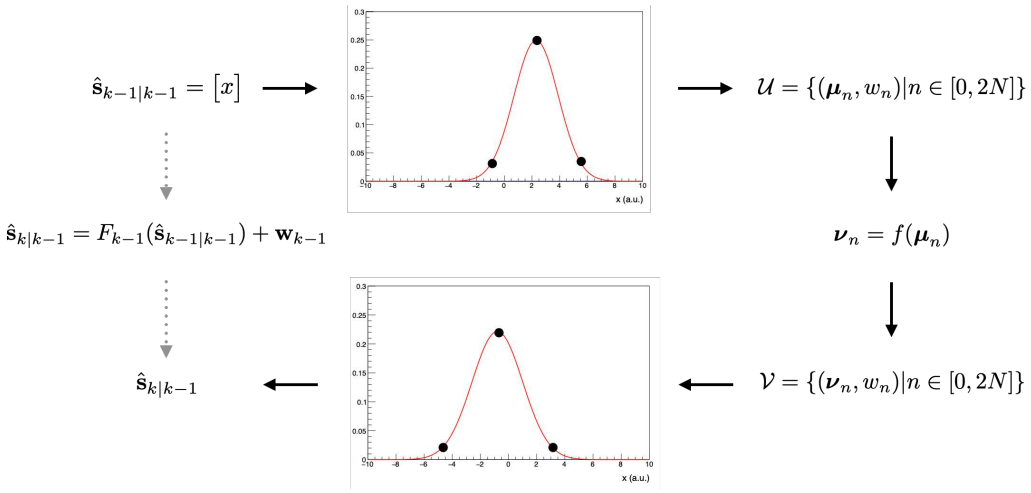

Figure 3.5: Illustration of the modified time update phase in the UKF. The example of a one dimensional state is taken. Since in this case the model is not linear, the classical form of the Kalman filter cannot be directly used. Therefore, three sigma points are derived from the probability distribution of the state. This probability distribution is illustrated by the Gaussian in the top panel. The purpose of the sigma points is to capture the moments of the distribution: the mean and the sigma value. The propagation model is applied to each of them, and from the transformed set, a new probability distribution for the state is deduced, which corresponds to the estimated state.

However, in the case of the TOE algorithm, an extra step need to take place before the propagation can occur again: not all of the measurements available at any given detection layer are valid candidates for the track reconstruction. An additional process must take place in order to select the viable clusters only.

Cluster selection: the confrontation procedure

For each given detection layer, a confrontation between the predicted state of the track and all the available measurements must be performed. This is necessary in order to authenticate which of those clusters could belong to the trajectory of the particle currently in reconstruction.

However, to achieve this confrontation, a selection criterion must be established. Here, considering the purpose of the procedure, the most obvious criterion is the distance between the predicted position and the measurement, in the plane orthogonal to the beam axis. Indeed, if the measurement

is far apart from the predicted state, either it does not belong to the trajectory, or the combination of hypotheses leading to this prediction is wrong.

The formalism developed in the last section allows for an easy measure of the distance between the predicted state and the clusters, regardless of their dimensions. The residuals of predictions $\mathbf{r}_{k|k-1}$, the associated covariance $R_{k|k-1}$ and the corresponding χ^2 value are given by the equations:

$$\mathbf{r}_{k|k-1} = \mathbf{m}_k - H_k \hat{\mathbf{s}}_{k|k-1} \tag{3.29}$$

$$R_{k|k-1} = M_k + H_k \hat{S}_{k|k-1} H_k^T \tag{3.30}$$

$$\chi^2 = {\mathbf{r}_{k|k-1}}^T {R_{k|k-1}}^{-1} \mathbf{r}_{k|k-1} \tag{3.31}$$

Selecting only the closest measurement to the predicted state is often a viable option. A different choice was made in the context of the TOE algorithm, because the propagation model suffers from two flaws: first, the initial value for the norm of the momentum can be quite far from the real one; second, it does not take into account any form of possible scattering by the incident particle. Therefore, limiting the choice to the closest cluster only could lead to an incorrect selection.

In consequence, the selection procedure is based on the creation of an imaginary point, called cutter point, for each measurement. This point is positioned to the maximum acceptable distance between the predicted state and the cluster under inspection. This cutter point is used to provide an adaptable threshold for selection: by comparing the χ^2 values for both the cutter point and the cluster, a decision between acceptance or rejection of the cluster can be made.

Cutter point

One cutter point is defined for each cluster available in a given detection layer: the error on the position for the measurements are integrated in the computed threshold. This allows for more stringent requirements on the match between predicted state and the measurement, if the position is determined with high accuracy. In turn, it reduces the amount of possible histories for a track, therefore reducing the computation cost for the reconstruction.

However, the measurement error is not the only factor forming the cutter candidate: an *ad-hoc* parameter, referred to as geometrical cut, also intervenes. Since part of the purpose of this value is to emulate the cumulative effects of scattering along the experimental setup on the transversal coordinates, several different cuts are associated to each detector. Furthermore, since the magnetic field impacts mainly the trajectory along the x

coordinate, the values of the geometrical cuts are not the same in x and y. It should be pointed out that this geometrical cut takes the form of an integer value.

The four detection layers of the VTX detector are grouped into two geometrical cuts; the four detection planes of the IT are grouped two by two into a total of four geometrical cuts; each layer of the MSD has its own geometrical cut; finally, the TW detector corresponds to two cuts (x and y).

A final factor takes part in the computation of the cutter point: multiple scattering has an higher impact on lighter ions, which should also be taken into account in the computation of this threshold. Therefore, an additional factor l is used as well. Its value is $l = 2.5$ for hypotheses corresponding to protons and deuterons, and $l = 1$ for the rest of the hypotheses. This factor is in fact an integral part of the static hypothesis detailed in section 3.1.

The cutter point is therefore the combination of the predicted state, the error made on the measurements $e_{x/y}$, and two *ad hoc* parameters necessary to take into account the cumulative impact of scattering of particles in the experimental setup. Since the geometrical cuts $g_{x/y}$ are asymmetrical, its position, $c_{x/y}$, lies on an ellipse. This position, necessary to deduce the actual selection threshold for the measurements, is therefore dependent on the angle between the predicted state and the cluster under inspection, in the plane orthogonal to the beam axis. The two axis length $a_{x/y}$ of the ellipse are defined as:

$$a_{x/y} = l g_{x/y} e_{x/y} \tag{3.32}$$

These values, together with the angle between the predicted state position and the measurement, denoted as φ, and the predicted state position $\hat{\mathbf{s}}_{x/y}$, are used to deduce the position of the cutter candidate:

$$c_x = \hat{\mathbf{s}}_x + \frac{a_x a_y \cos(\varphi)}{\sqrt{(a_x \cos(\varphi))^2 + (a_y \sin(\varphi))^2}} \tag{3.33}$$

$$c_y = \hat{\mathbf{s}}_y + \frac{a_x a_y \sin(\varphi)}{\sqrt{(a_x \cos(\varphi))^2 + (a_y \sin(\varphi))^2}} \tag{3.34}$$

At this point, the equations (3.29),(3.30), (3.31) can be used for both the cutter point and the cluster. However, since the χ^2 value is used as the discrimination factor, the covariance matrix of residuals $R_{k|k-1}$ should be the same in both computations. Accordingly, the covariance matrix of the measurements M_k is modified; it is setup according to the axis of the

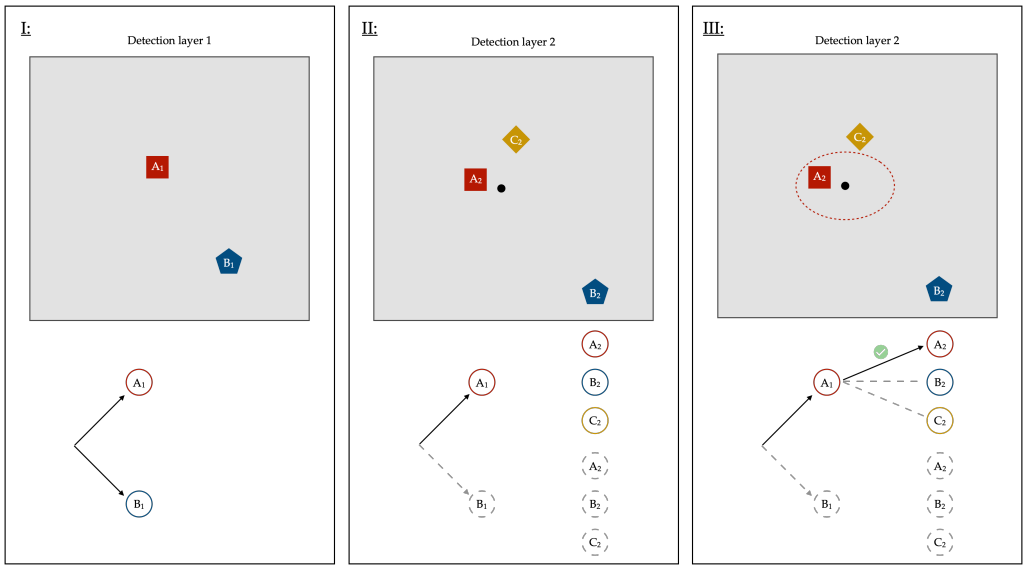

Figure 3.6: Schematic representation of the cluster selection procedure, panels I to III. The propagation starts from panel I, in an fictional first detection layer. The two measurements present A_1 and B_1 are both added as seeds to the arborescence. From there, the branch holding cluster A_1 is propagated to the second detection layer, and the estimated state is shown as the black dot. The second layer holds three different clusters: A_2, B_2 and C_2. The clusters sharing the same letter, color and form belong to the same track and should therefore be matched by the procedure. The cutter point for the cluster A_2 is symbolized by the corresponding colored eclipse in panel III. Since the cluster position is within its corresponding ellipse, it is therefore added to the history as a distinct branch.

ellipse:

$$M_k = \begin{pmatrix} a_x^2 & 0 \\ 0 & a_y^2 \end{pmatrix} \tag{3.35}$$

With this modification, the χ^2 value for the cutter point is in most cases close to unity.

Once both values are obtained, they are compared: if the χ^2 value from the measurement is lower than the one for the cutter, then the cluster is accepted as a possible point in the trajectory of the particle, and the history is updated accordingly. In the case it is superior, the cluster is rejected.

Figures 3.6, 3.7 and 3.8 illustrate the cluster selection procedure and its impact on the corresponding history, on two pixel detection layer. Two real tracks, denominated as A and B are to be reconstructed. The first detection layer shows two reconstructed clusters (A_1 and B_1), while the second

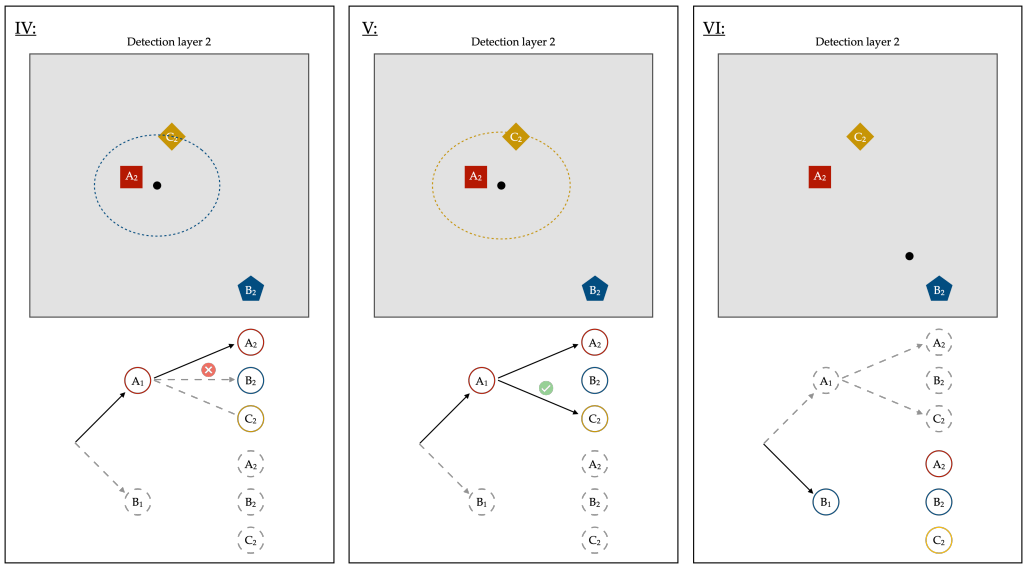

Figure 3.7: Schematic representation of the cluster selection procedure, panels IV to VI. The cutter point for the clusters B_2 and C_2 are symbolized by their corresponding colored eclipse, in panel IV and V respectively. The cluster C_2 is the only of the two which is within its corresponding ellipse, it is therefore also added to the history as a new branch. In panel VI, the procedure is then started back from the branch starting with the cluster B_1, with the black dot representing the propagated state for this branch.

presents three (A_2, B_2, C_2). The clusters on the first detection layer are used as seed for the recognition. The respective initial states corresponding to those clusters is then propagated, individually, to the second detection layer. Each available cluster is compared to the predicted state using the corresponding cutter point, symbolized by an ellipse. The arborescence ends up with three different history.

The geometrical cuts used to compute the position of the cutter point are by definition dependent on the geometry and beam used to produce the experimental data. The optimization of those values is the subject of the following section.

Optimizing geometrical cut values

The problem of optimizing the geometrical cut values in order to obtain the best performances possible for any configuration of the experimental setup is difficult by nature: because a particle has to cross the several layers composing the setup in order to be reconstructed, the cut values are interdependent.

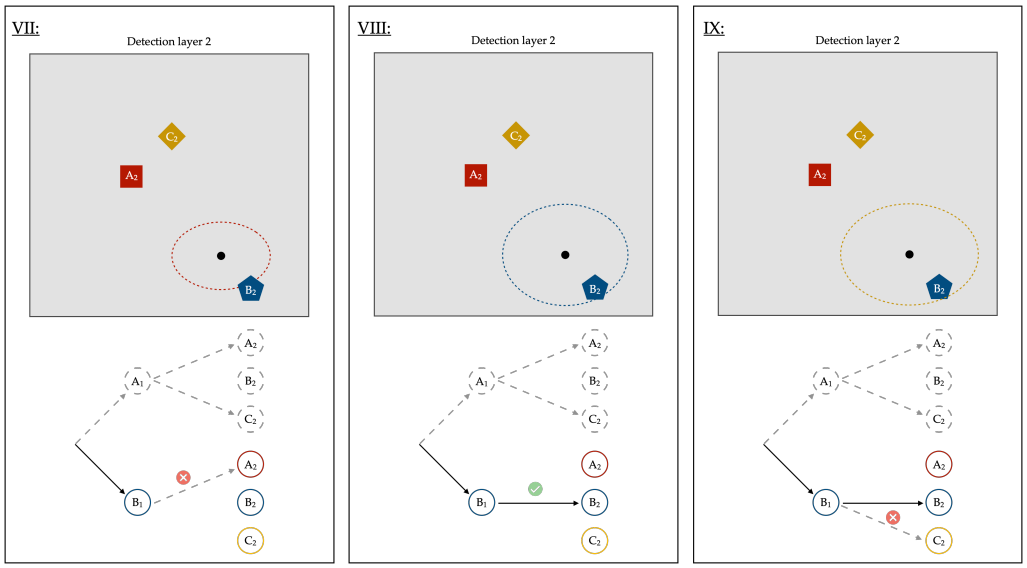

Figure 3.8: Schematic representation of the cluster selection procedure, panels VII to IX. In each panel, the cutter point for one of the cluster in the detection layer, is represented by the corresponding colored eclipse. Only the cluster B_2 is considered valid, and so is added to the arborescence.

Indeed, if the ones corresponding to the first layers are too small, then no matter how large the following ones are, the performances will be limited by the tracks already rejected in the first planes. Conversely, if the cut values for one layer are too high, a state where all measurements in the said layer would be accepted can be reached. The resulting size of the arborescence would increase drastically, therefore augmenting the computational cost of the overall method.

In order to devise the best approach to determine the optimal values for those geometrical cuts, two complementary methods have been implemented, in a dedicated software tool. The first one, inspired from the gradient descent algorithm, converges towards optimal values, but does so at a heavy time cost. The second one, the comb method, is fast but can only deliver a rough approximation of the optimal values.

Both of those methods aim to maximize a scoring function provided by the user. This function corresponds to the goals of the experiment: in our case, a combination of efficiency and purity (as defined later on in section 3.2) is used. Because of the statistical nature of those quantities, the optimization method needs to be run on a large number of events at a time to produce meaningful results.

	VTX		IT				MSD						TW	
	x	y	x	y	x	y	x	y	x	y	x	y	x	y
16O200C2H4	17	21	59	49	63	61	6	5	3	4	2	3	7	9
16O200C	18	13	74	42	69	67	6	5	3	4	2	3	9	11
16O400C	8	9	39	29	32	29	4	4	5	2	6	3	5	10
12C200C	15	21	73	51	77	41	9	6	6	4	4	2	12	15

Table 3.2: Set of geometrical cut values used for the simulated data set used in this chapter. The characteristics of each set can be found in section 3.2. These cut values have been obtained after respective optimization for efficiency and purity, as described in section 3.2.

Gradient-descent method The gradient-descent method is quite simple in its concept: from a set of initial cut values, a baseline score is determined by performing the reconstruction on the selected set of events. Afterwards, the smallest possible variation on a cut, i.e. an increment or decrement of 1, is introduced sequentially on all the cuts. For each of the possible configurations, the reconstruction is fulfilled. The corresponding score is stored for each one of them. The cut variation providing the best increment on the scoring function is applied to the geometrical cuts, therefore defining a new baseline score. This procedure is applied until no cut variation increases the score further.

Comb method The comb strategy bases itself on the gradient-descent strategy, but an additional set of steps is undertaken once the small variation scan is over. Rather than directly applying the modification from the winning variation, a rough scan of the possible values for the selected cut is performed.

The variation corresponding to the winning score for the rough scan is then further examined: a finer scan exploring the values of the cut neighboring the winning value is applied. From this finer scan, a final variation is proposed for the corresponding cut, which also defines a new baseline. A global scan over all the cut values can take place once more, skipping over the cut set in the last step.

This procedure is applied until the score cannot be further increased by varying any of the cut values, or if all the cuts have been evaluated.

Thus, for each geometry and beam configuration where data is acquired, the geometrical cut values need to be optimized using one or both of those methods. Table 3.2 lists the cut values used for the four different simulated

data sets used in the section exploring the performances of the algorithm (section 3.2).

The combinatorial approach use in the TOE algorithm recognition procedure is therefore enabled by the cluster selection strategy. However, it could not be accomplished without a way to store the information as the procedure unfolds. This is the role played by the arborescence.

History of possible trajectories: the arborescence

The arborescence role is to record the track history: as the track recognition procedure unfolds, a determined number of measurements is available at each given layer; even though among them only one corresponds to the true track currently in reconstruction, several can be considered as viable. For each of the measurements clearing the confrontation step of the procedure, a new branch is added to the arborescence.

Thus, the recognition procedure must take place for every branch of the arborescence, until either it reaches the final detector and the branch is recorded as a viable trajectory for the track, or the branch is marked along the way invalid. Given the configuration of the experimental setup, branches are not directly dropped if no cluster is added to them at a detection layer: skipping several layers is allowed. Invalidity for a given branch can arise in three different scenarios:

- Seeding procedure: the tracklets reconstructed in the vertex detector are used to provide the initial measurements for the recognition procedure. Given $N_{tracklet}$ the number of clusters in the tracklet, if the number of corresponding clusters accepted in the arborescence is below $N_{tracklet} - 1$, then the branch is dropped.

- Last cluster origin: before starting propagation in a new detector region, the origin of the last cluster for each branch is cross-checked. If it does not come from the detector directly before the current one, the branch is marked as invalid.

- Final endpoint cross-check: the clusters in the TW act as final cross-checks; only the ones associated with the current static hypothesis are verified. If none pass the selection procedure, the branch is rejected.

The implementation of the arborescence is based on a custom tree data-structure. It is a hierarchical structure by concept, with several root nodes corresponding to the dynamic hypotheses, and their subtrees of children.

Each node keep tracks of its parent node and possibly several children as well as the relevant information for the recognition procedure.

In order to expose the underlying concepts behind the recognition procedure, only one history (or branch) at a time has been considered so far. However, the recognition procedures does not focus on one branch of the arborescence until completion: propagation and filtering occurs in a wave-like pattern.

Wave-like approach

In order to be as efficient as possible, each detection layer of the experimental setup should be visited only once. Therefore, the recognition procedure adopts a wave-like approach. When a new detection layer needs to be reached, all the valid branches at the time are propagated towards the plane. After a branch is split into several trajectories, the pool of branches to propagate to the next detection layer is increased accordingly. However, this requires a special tool able to retrieve all of the leaves (i.e. the outermost nodes) of the arborescence before propagation can start.

Furthermore, this approach allows for a memory optimization: all of the branches will be confronted to the same set of clusters in any detection layer. By propagating them together, the set of available clusters can be kept in memory for the confrontation procedure and filtering of each branch, and discarded afterwards.

In turns, this implies that the propagation needs to be restarted from the previous corrected state of the track, which needs to be stored inside the arborescence.

Figure 3.9 represents part of the recognition procedure, from detection layer n to detection layer $n + 3$. It emphasizes, through colored arrows, that the propagation step in between plans is done with all branches of the arborescence concurrently, rather than focusing on one of them until it reached the final detection layer.

Once all detection layers have been visited, the arborescence outputs a set of possible trajectories for each endpoint considered. Additional arborescences, corresponding to matching static hypotheses could also lead to several possible trajectories for the same endpoint. All of these trajectories need to be extracted and compared: out of them, only one can match the real trajectory of the particle. This corresponds to another procedure, called trajectory shearing.

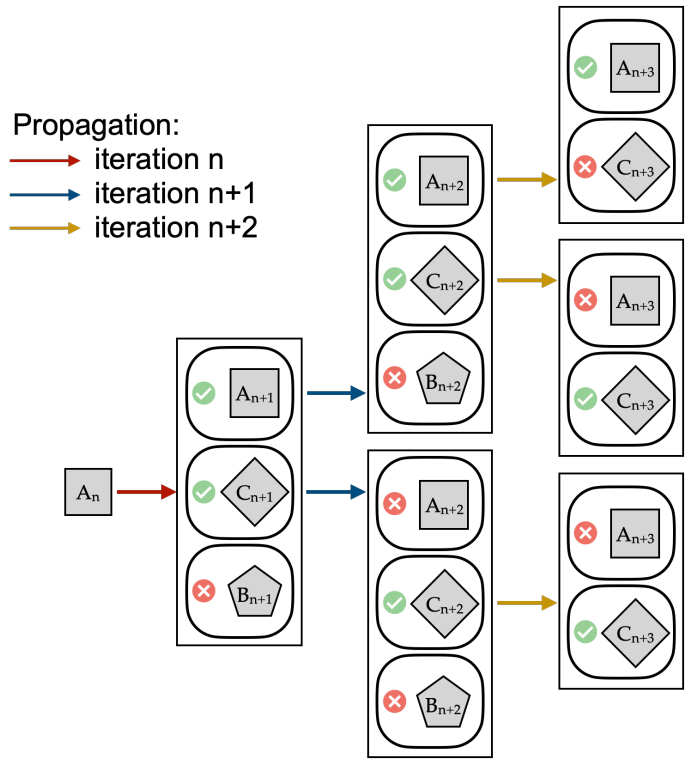

Figure 3.9: Illustration of the wave-like approach adopted for the procedure, for a branch starting with the cluster A_n until the layer $n+3$ is reached. In between detection layer, all the available branches forking from this one are propagated concurrently to the following detection layer, as denoted by the colored arrows. At each detection layer, the cluster selection, represented by the several rectangular box is applied, each available cluster potentially leading to a branch separation.

Trajectory shearing

The combinatorial approach used by the recognition procedure invariably leads to the reconstruction of several possible trajectories for a same endpoint, which identify the particle under reconstruction. Therefore, it is necessary to select only one of those possible trajectories in order to associate it to the particle and be able to extract parameters of interest from it.

To do so, a so-called shearing factor is computed for each possible trajectory leading to the same endpoint.

A fictional example of a reconstruction with several possible trajectories can be found in figure 3.10: two different static hypotheses corresponding to the same endpoint lead to the reconstruction of respectively three and two

possible trajectories. All of them are compared using the shearing factor, and the winning trajectory is outlined with full dark lines.

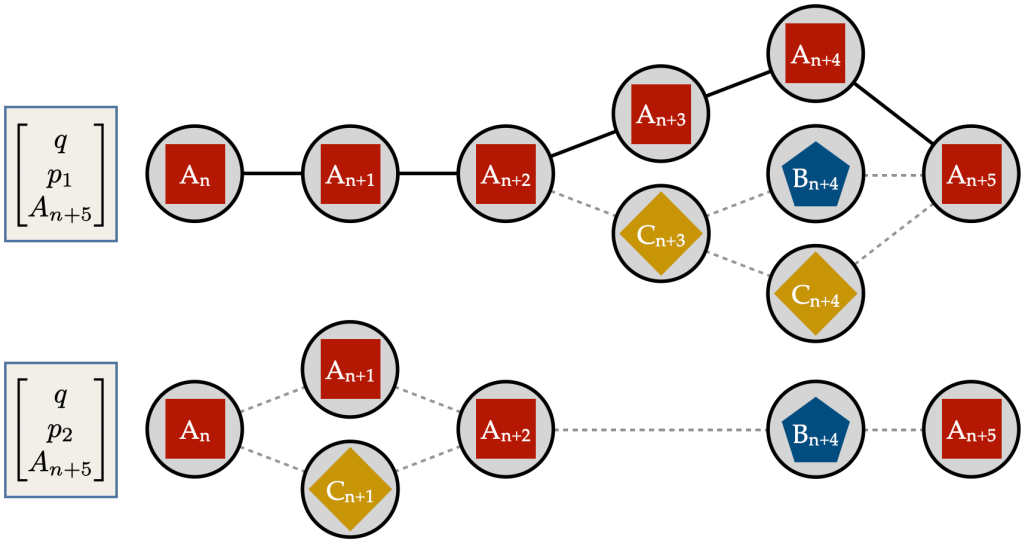

Figure 3.10: Schematic illustration of the need for a shearing procedure. In this example, two different arborescences share the same endpoint (A_{n+5}). The first one registered three different trajectories while the second registered two. Out of them, only one can correspond to the trajectory of the particle: the winning trajectory, according to the shearing procedure corresponds to the black lines between clusters. In this case, the reconstruction successfully associate the correct clusters to the trajectory. If any other trajectory was selected, at least one cluster would have been wrongly associated with the track. This is represents indeed a scenario with an optimal outcome for the recognition procedure.

The shearing factor can only be based on the overall distance between the track state and the corresponding measurements: the lowest this distance is, the more probable the hypothesis leading to this history is.

However, as seen in the cluster selection procedure (section 3.1), the cluster lookup is not done in the same surface area for each detection layer, due to the geometrical cuts. Therefore, the shearing factor should also be taking into account this effect in order to be reliable: not all clusters will contribute with the same weight to the computation.

The shearing factor is therefore computed based on χ^2 values, in pretty much the same way the cluster selection unfolds, but with a twist: rather than using the state of the track at the prediction phase to compute residuals, the one obtained after correction is used. Once again, the measurement covariance matrix is modified according to equation (3.35).

The $\chi^2_{corrected}$ value for each cluster is therefore stored in the arborescence, along with the corrected track state. For any given history j, the shearing factor $\mathcal{S}_j$ is given by the following formula, assuming that N_j clusters were selected:

$$\mathcal{S}_j = \sqrt{\frac{1}{N_j - 1} \sum_{i=0}^{N_j - 1} {\chi^2_{corrected}}^{(i)}} \tag{3.36}$$

For a given endpoint, a shearing factor is computed for each possible trajectory. The history with the lowest shearing factor is selected as the winning trajectory, and registered in order to be used in the remainder of the track reconstruction.

The inner workings of the track recognition procedure of the TOE algorithm have been exposed in full details. Looking at the performances of this procedure is the mandatory step that follows.

3.2 Evaluation of performances

The evaluation of the performances of the track recognition procedure requires the definition of a few different criteria: the efficiency of the recognition, the cluster purity of the track reconstructed, the fake yield associated with the recognition and finally the clone multiplicity are all necessary to quantify how well the algorithm performs [118].

By essence, the evaluation of performance requires some form of "truth" about the track to reconstruct to be known. Therefore, the performances are based on Monte-Carlo simulations that are able to provide the required information.

Furthermore, since the performances of the global reconstruction algorithm are the focus here, any erroneous charge reconstruction by the TW detector and local reconstruction was ignored: the charge used as input for the algorithm is here obtained from the Monte-Carlo truth.

Definitions

Reference set

In order to assess the efficiency of the procedure, a set of track of reference $N_{reference}$ must be established. This set corresponds to all the tracks that an algorithm performing ideally should find and reconstruct. In the case of the TOE algorithm, it is taken as all the tracks which are within the acceptance

of the experimental setup, and have formed a minimum of one cluster in each detector. Therefore, all of those tracks are uniquely identified by their final cluster, originating from the TW detector.

Recognition efficiency

The definition of the track recognition efficiency is dependent on a criterion which specifies if a reconstructed trajectory corresponds to the track in the reference set or not. A track is considered as found by the algorithm if it includes one of the endpoints of the reference set, given the major role played by it in the recognition procedure. The recognition efficiency can then be expressed as:

$$\varepsilon_{recognition} = \frac{N_{recognition}}{N_{reference}} \tag{3.37}$$

Fake yield

The recognition procedure can also reconstruct trajectories that do not belong to the reference set: those trajectories correspond therefore to fake reconstructions. This is an important measure in the evaluation of the performances of the algorithm: if this number is to high, the reconstruction algorithm cannot be applied on experimental data. The fake yield is defined as:

$$\varepsilon_{fake} = \frac{N_{fake}}{N_{recognition} + N_{fake}} \tag{3.38}$$

Cluster purity

Finally, given that the purpose of the recognition part of the algorithm is to identify which reconstructed clusters belong to a same fragment trajectory, the cluster purity of the reconstructed track should be quantified. For each track reconstructed by the algorithm, the origin of each cluster can be checked using the Monte-Carlo truth: the cluster at the endpoint serves as an indication for all of the others. With $N^{(m)}_{correct}$ the number of clusters correctly associated to a MC track m (corresponding to a single particle), and $N^{(m)}_{total}$ the corresponding total number of clusters in the reconstructed track, we get, for a set of M reconstructed tracks, a cluster purity that can be expressed as:

$$\rho = \frac{\sum_{m=0}^{M} N^{(m)}_{correct}}{\sum_{m=0}^{M} N^{(m)}_{total}} \tag{3.39}$$

Clone multiplicity

Another evaluation point for the performances of the algorithm, this time in term of computational cost, is the quantification of the number of multiple cloned trajectories produced for a same endpoint, removed by the means of the shearing procedure. Indeed, the more clones are produced, the more time will be spent performing the reconstruction. Therefore, this number should be kept as low as reasonably possible while retaining a good recognition efficiency. The clone multiplicity can be determined individually for each track m reconstructed, which gives an average of:

$$< N_{clone} >= \frac{\sum_{m=0}^{M} N_{clone}^{(m)}}{N_{recognition} + N_{fake}} \tag{3.40}$$

Simulated data sets

The evaluation of the performances of the algorithm is only possible by comparison with values of reference. This information is inherently available in the data produced by Monte-Carlo simulations.

Four different data sets are used in this section: for the rest of the chapter, they will be referred to as 16O200C2H4, 16O200C, 16O400C and 12C200C. Looking at data produced in different configurations is key in order to understand how well the algorithm works, and to discover any flaws.

All of those files were produced by the FLUKA Monte-Carlo code, with performances studies as their purpose. Therefore, they make use of the most accurate geometrical description of the experimental setup available at the time.

- 16O200C2H4: uses a ^{16}O beam at 200 MeV/u, interacting in a 5 mm thick C_2H_4 target, with density 0.94 g/cm^3. A total of 10^7 primary ions have been simulated, but only fragmented events have been stored in the data set, out of which 10^5 have been analyzed.

- 16O200C: uses a ^{16}O beam at 200 MeV/u, interacting in a 5 mm thick C target, with density 1.83 g/cm^3. A total of 10^7 primary ions have been simulated, but only fragmented events have been stored in the data set, out of which 10^5 have been analyzed.

- 16O400C: uses a ^{16}O beam at 400 MeV/u, interacting in a 5 mm thick C target, with density 1.83 g/cm^3. A total of 10^5 primary ions have been simulated, and all events have been stored in the data set. All of them have been analyzed.

	16O200C2H4	16O200C	16O400C	12C200C
Efficiency ε (%)	98.7 ± 0.4	98.7 ± 0.4	99.8 ± 0.5	99.0 ± 0.7
Fake yield y_{fake} (%)	5.87 ± 0.07	8.06 ± 0.08	2.39 ± 0.05	9.5 ± 0.2
Multiplicity $\langle N_{clone} \rangle$	22.29 ± 0.06	32.48 ± 0.09	3.33 ± 0.01	36.1 ± 0.2
Purity ρ (%)	95.1 ± 0.1	94.5 ± 0.1	99.6 ± 0.1	94.2 ± 0.2

Table 3.3: Overall efficiency, fake yield, clone multiplicity and purity for the application of the recognition procedure on four simulated data sets. The values of efficiency, fake yield and purity are expressed as percentages.

- 12C200C: uses a ^{12}C beam at 200 MeV/u, interacting in a 5 mm thick C target, with density 1.83 g/cm^3. A total of 10^6 primary ions have been simulated, but only fragmented events have been stored in the data set, out of which $3{\times}10^4$ have been analyzed.

Results

Table 3.3 reports the overall evaluation of the recognition procedure performances: the quantities defined in the section 3.2 are assessed for each data set.

It is apparent that no real distinction in term of efficiency separates the data sets: no matter the configuration, its overall value is close to 99%. The data set 16O400C shows large discrepancies in comparison to the others, when looking at the three remaining parameters. This can be explained by the nature of the data contained in those sets: 16O400C is the only one which mainly contains non-fragmented events.

Therefore, the majority of the particles that are reconstructed for the 16O400C data set are oxygen ions, with a charge Z=8. For the other sets, light particles, i.e. with Z=1 or Z=2 are principally reconstructed. Because they are more prone to scattering, the size of the lookup ellipse in the algorithm need to be augmented. This has the downside of accepting more clusters in the history, therefore leading to a possibly higher clone multiplicity and a decreased purity. Furthermore, if the amount of available clusters is increased, fake reconstructions are more likely to be conducted to completion.

This can be easily demonstrated by modifying a key parameter applied to the lighter particles only: the *ad-hoc* parameter l, evoked in section 3.1. For example, considering the data set 16O200C2H4 and reducing the value of l from 2.5 to 2 leaves the overall efficiency and purity unaffected, but drops down the values for the fake yield and clone multiplicity to $y_{fake} =$

		16O200C2H4	16O200C	16O400C	12C200C
Z=1	ε (%)	98.1 ± 0.7	98.7 ± 0.6	95 ± 3	99 ± 1
	ρ (%)	90.6 ± 0.2	89.8 ± 0.2	86.9 ± 0.8	90.7 ± 0.3
Z=2	ε (%)	98.2 ± 0.6	97.9 ± 0.6	98 ± 3	99 ± 1
	ρ (%)	94.0 ± 0.2	94.3 ± 0.2	93.4 ± 0.9	93.7 ± 0.3
Z=3	ε (%)	99 ± 2	100 ± 2	100 ± 11	100 ± 3
	ρ (%)	98.5 ± 0.6	98.7 ± 0.5	98 ± 3	99.2 ± 0.8
Z=4	ε (%)	100 ± 3	100 ± 3	100 ± 18	100 ± 4
	ρ (%)	99.0 ± 0.7	99.2 ± 0.7	99 ± 5	100 ± 1
Z=5	ε (%)	100 ± 2	100 ± 2	100 ± 14	100 ± 3
	ρ (%)	99.3 ± 0.6	99.4 ± 0.6	98 ± 3	100.0 ± 0.7
Z=6	ε (%)	100 ± 1	100 ± 1	100 ± 9	100 ± 3
	ρ (%)	99.6 ± 0.4	99.6 ± 0.4	100 ± 2	100.0 ± 0.8
Z=7	ε (%)	100 ± 1	100 ± 1	100 ± 7	
	ρ (%)	99.9 ± 0.3	99.9 ± 0.4	100 ± 2	
Z=8	ε (%)	100 ± 1	100 ± 1	100.0 ± 0.5	
	ρ (%)	100.0 ± 0.3	100.0± 0.5	100.0 ± 0.1	

Table 3.4: Efficiency and purity computed separately for each fragment charge, for the four simulated data sets.

$4.77 \pm 0.06\%$, $\langle N_{clone} \rangle = 9.56 \pm 0.02$.

Thus, the reconstruction of particles with charge Z=1 is the main culprit behind a high clone multiplicity and fake yield and therefore can impart a high computational cost if the parameters of the algorithm are not tuned appropriately. It should be noted that reducing the value of the factor l has a negative impact on the efficiency of the reconstruction of the particles with charge Z=1. Nonetheless, if need be, the value of the parameter l can be tuned down to speed-up the reconstruction procedure.

In order to dive deeper into the performances of the recognition procedure, efficiency and purity should be investigated individually for each fragment charge. Table 3.4 summarizes these results.

For all charges $Z > 2$ both purity and efficiency are around or above 99%. Efficiency for Z=1 and Z=2 is not far from this value: taking into account the associated error, it fluctuates around 98% for all data sets. However, the association of clusters to the track is wrong in a non-negligible proportion of reconstructions: the purity for charge Z=1 and Z=2 is respectively around 90% and 94% across data sets.

Figure 3.11 and 3.12 display the purity distributions of fragments with charge Z=1 and charge Z=2 as a function of both the emission angle θ

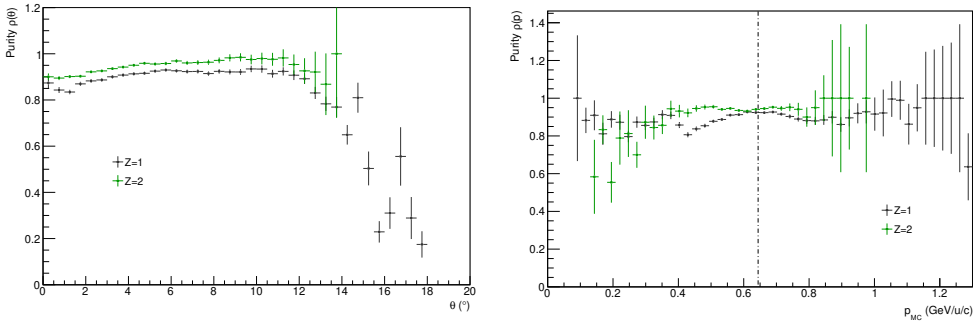

Figure 3.11: Purity distribution as a function of the emission angle θ on the left and as a function of the true momentum per nucleon of the particle p_{MC} on the right, for the 16O200C2H4 data set. The distributions corresponding to the fragments with Z=1 are shown in black, and the fragments with Z=2 in green. The momentum per nucleon deduced from the beam energy and used to establish the static hypothesis is indicated by the vertical black dotted line, in the corresponding distribution.

and the momentum p_{MC} per nucleon, for the data sets 16O200C2H4 and 16O200C respectively. Because of the need for significant statistics in order to produce meaningful results, only those two sets have been considered here.

It is apparent that light fragments with an emission angle superior to the acceptance devised for the FOOT experiment ($\approx 10\%$) can be reconstructed by the TOE algorithm. However, the purity sharply declines with the angle, past $10 - 12°$. Furthermore, the momentum distribution of the light fragments is broad: therefore, when the real momentum of the particle differs to much from the hypothesis used in the algorithm, the purity of the reconstruction suffers.

3.3 Track fitting

Once a track has been reconstructed, i.e. the clusters belonging to it have been identified and grouped together, the extraction of the parameter of interests for the experiment can be performed; in our case, both the length of the path and the momentum of the particle.

Path length evaluation

The evaluation of the length of the path taken by the particle in the experimental setup is a pre-requisite in order to extract the velocity of the particle

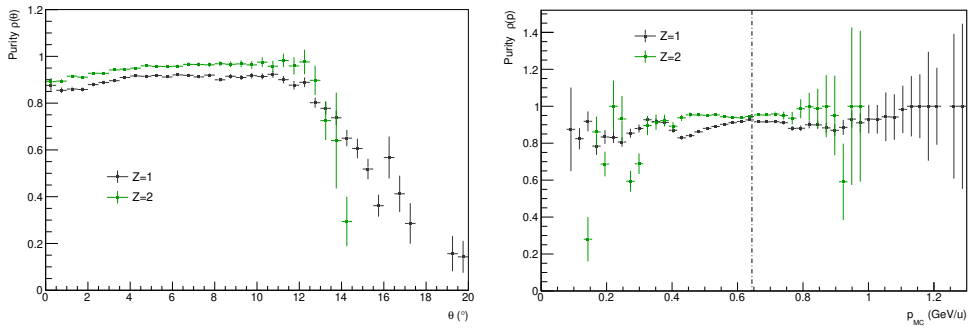

Figure 3.12: Purity distribution as a function of the emission angle θ on the left and as a function of the true momentum per nucleon of the particle p_{MC} on the right, for the 16O200C data set. The distributions corresponding to the fragments with Z=1 are shown in black, and the fragments with Z=2 in green. The momentum per nucleon deduced from the beam energy and used to establish the static hypothesis is indicated by the vertical black dotted line, in the corresponding distribution.

through its β value. Its estimation must involve the selected clusters for the trajectory, which represents the only spatial information available for the track.

Since the β value is also dependent on the τ_{tof} value, the two detectors involved in its measurement define natural limits for the length of the path. However, it should be pointed out that in the ST detector, the primary ion is detected, while the object of interest, represented by the reconstructed track, is one of the fragments it produced afterwards, in the target. Therefore, the τ_{tof} value, corrected from the primary ion contribution, should be used.

Track model

In order to extract the length of the path, a track model or parametrization must be employed, and this model should match the expected characteristics of the track given the configuration of the experiment. The nature of the magnetic field used in the FOOT experiment implies a slight deviation of most tracks along the x coordinate, while leaving the y coordinate mostly unaffected.

In consequence, simple polynomials are fitted, separately, to the bending (x-z) and non-bending plane (y-z): the former is matched with a polynomial of order three, while the latter uses a polynomial of order one. An ordinary least squares method is used to perform the minimization and deduce the

	$R^2 > 95\%$		$R^2 > 99\%$	
	$x-z$	$y-z$	$x-z$	$y-z$
16O200C2H4	100 %	96.4 %	100 %	92.2 %
16O200C	100 %	96.5 %	100 %	92.4 %
16O400C	100 %	76.4 %	100 %	48.7 %
12C200C	100 %	96.8 %	100 %	93.5 %

Table 3.5: Percentage of fits, performed to obtain the track parameters from the selected clusters, for which the determination coefficient is above 95% and 99% respectively. The percentages are presented independently for the $x-z$ and $x-y$ planes, for the four data sets.

corresponding parameters.

Table 3.5 shows the percentage of fits which results in a determination coefficient, from the linear regression used to fit the points, above 95% and 99%, independently for the two plane, for the four different data sets. In all cases, the polynomial of order three used in the $x-z$ plane is a good match for the shape of the trajectory. In the three data sets composed only of fragmented events, the hypothesis of a straight line in the $y-z$ plane holds well.

Parameter extraction

Once the parameters for the two polynomials have been obtained, they can be leveraged to extract the length of the path between the position of the reconstructed vertex, and the endpoint in the TW:

$$l = \int_{z_{vertex}}^{z_{TW}} \mathrm{d}z \sqrt{\frac{\mathrm{d}x}{\mathrm{d}z}^2 + \frac{\mathrm{d}y}{\mathrm{d}z}^2 + 1} \tag{3.41}$$

The derivatives with respect to z are immediately deducible from the parameters of the corresponding polynomials, and therefore, the numerical integration has a low impact on the overall computational cost of the procedure.

Once the length of the path of the particle is obtained, its β value can be deduced using equation (2.2).

Resolution on β

In order to evaluate how well the algorithm performs its task, establishing the resolution on the β value it provides is a crucial point. To do so, the reconstructed value is compared to the Monte-Carlo information: the relative

	$\mathcal{R}_\beta$ (%)			
	16O200C2H4	16O200C	16O400C	12C200C
Z=1	19.92 ± 0.05	21.18 ± 0.05	29.3 ± 0.3	22.4 ± 0.1
Z=2	6.15 ± 0.01	6.24 ± 0.01	20.6 ± 0.3	6.71 ± 0.03
Z=3	3.30 ± 0.03	5.85 ± 0.04	12.4 ± 0.5	5.04 ± 0.05
Z=4	5.18 ± 0.05	4.56 ± 0.04	2.3 ± 0.2	3.00 ± 0.04
Z=5	3.55 ± 0.03	4.37 ± 0.03	2.2 ± 0.1	6.88 ± 0.06
Z=6	4.60 ± 0.02	4.78 ± 0.03	9.5 ± 0.3	4.45 ± 0.05
Z=7	4.60 ± 0.02	3.37 ± 0.02	2.32 ± 0.06	
Z=8	5.11 ± 0.03	5.67 ± 0.03	7.19 ± 0.01	

Table 3.6: Resolution on the β value, independently for each fragment charge, for the four data sets used for the evaluation of performance of the recognition procedure.

difference between the two is computed for each reconstructed track. The standard deviation from zero can then be computed for the corresponding distribution. For a total of M reconstructed tracks:

$$\mathcal{R}_\beta = \sqrt{\frac{1}{M-1}\sum_{m=0}^{M-1}\left(\frac{\beta_{MC}^{(m)} - \beta_R^{(m)}}{\beta_{MC}^{(m)}}\right)^2} \tag{3.42}$$

Table 3.6 references the resolutions on the parameter β for the four simulations evoked in section 3.2. The resolutions have been evaluated independently for each charge.

Again, the reconstruction struggle with fragments with charge $Z = 1$: while the overall resolution for all other charges and data sets is below 10%, it is constantly above 20% in their case, over all simulated data sets.

The results obtained for the set 16O400C differ from the rest of the simulation files, which does not come as a surprise: since the fit in the $y - z$ plane is not reliable in a quarter of the reconstructions ($R^2 < 95\%$), the corresponding length value is erroneous and so the resolution on β is degraded.

Momentum computation

In order to establish single and double differential cross-sections, for all of the fragments isotopes produced during the experiment, within the collaboration goals, an estimated resolution on momentum of the order of 5% for the heavier ions must be obtained. This requirement is what motives most

of the decisions taken in order to reconstruct the momentum in the TOE algorithm.

Momentum estimation

Once the β value for a track is obtained, a first estimation of the momentum of the track can be performed using the formula:

$$p = \gamma\beta cUA \tag{3.43}$$

With U the unified atomic mass, A the number of nucleons, γ the Lorentz factor. However, this value is dependent on both the β value obtained before, and the number of nucleon determined in the static hypothesis.

The first dependency represents an issue in the sense that it augments the correlation between variables that should be kept independent in order for the three identification methods developed in section 2.3 to be effective. The second dependency can lead to an artificial degradation of the momentum resolution: if the number of nucleons is wrong by one, the corresponding disparity in momentum is consequent.

Momentum scanning

In order to reduce the impact of misidentification on the resolution of momentum, as well as reducing the dependency on the τ_{tof} value, an additional step was introduced to achieve momentum determination (as well as a recomputation of the number of nucleons A_R).

The basic idea behind this subsequent step is to refine the value of the momentum provided by the reconstruction by scanning the momentum range around the value estimated on the last step.

Indeed, at this point in the algorithm, a track model and its parameters was defined. This model is derived using the clusters selected in the recognition procedure. Therefore, a reconstruction of the trajectory of the particle in the experimental setup can be performed again, and the track model can be used as a comparison point. The exact same propagation model can be used, and only the norm of the momentum inserted at the start of the propagation will be modified.

The iterative part of the procedure can be described as follows:

① Retrieve the norm of the momentum per nucleon $\hat{p}$ from the estimated momentum

② Initialize (or increment by a step h, in the subsequent iterations) the value of the correction factor f, ranging between [A_R -2, $A_R + 2$]

③ Compute the value of the updated momentum $p = f\hat{p}$

④ Define N evaluation points, selected at random between the initial z position defined by the first layer of the VTX detector and the final one defined by the TW detector

⑤ Perform the propagation through the magnetic field using this updated momentum value p until one of the evaluation point is reached

⑥ At the evaluation point n, compute the distance s_n between the propagated state and the track model in the plane orthogonal to the direction of propagation

⑦ If some evaluation point has not been visited, go back to step ⑤

⑧ Compute and store the overall score of the propagation against the track model:

$$\mathcal{S} = \sqrt{\frac{1}{N-1}\sum_{n=0}^{N-1} {s_n}^2} \tag{3.44}$$

⑨ While $f \leq A_R + 2$, go back to ②

The output of this sequence of steps is therefore a set of pairs formed by the norm of momentum value used and the associated score for the reconstruction. In theory, the score function $\mathcal{S}(p)$ should present a minimum in the region that was iterated over.

This minimum corresponds to the final momentum value provided by the reconstruction: it is extracted by adjusting a polynomial of order two to the few points surrounding the minimal score computed during the iterative part. From there, the momentum is deduced by finding the zero crossing of the derivative of the polynomial. Table 3.7 shows the percentage of fits where the value of the determination coefficients is above a 95% and 99% threshold, for the four simulated data sets used to evaluate the recognition performances. Overall, the polynomial of order two hypothesis holds well.

Finally, once the final momentum value is obtained, the other parameters for the track are adjusted accordingly: both the number of nucleon and the mass of the particle are re-evaluated. The migration matrices can be used to evaluate the viability of these reconstructed values: in these matrices, the reconstructed number of nucleons A_R is shown against the true number of nucleons of the particle A_{MC}. Each cell indicates the percentage of entries it holds with respect to the column (or true nucleon number) it

	$R^2 > 95\%$	$R^2 > 99\%$
16O200C2H4	93.9 %	79.2 %
16O200C	94.3 %	78.5 %
16O400C	74.4 %	73.3 %
12C200C	97.1 %	80.4 %

Table 3.7: Percentage of fits performed after momentum scanning for which the determination coefficient is above 95% and 99% respectively, for the four data sets.

belongs to. The respective contribution of each nucleon number to the total of reconstructed tracks is also presented, in the row below the matrix.

Figure 3.13 and figure 3.14 show the final migration matrices for two simulated data sets: 16O200C2H4 and 16O200C, respectively. For both of them, the diagonals, which corresponds to a correct identification, are well defined: the yields are around 90% for $A = 1, 2, 3, 4$, which also represents a total of respectively 65.1% and 70.7% of all reconstructed tracks. For the heavier fragments, the diagonal fluctuates around 65-75%. The number of nucleons $A = 5$, $A = 8$ and $A > 16$ are poorly reconstructed, because no static hypothesis can lead to the creation of such a value. However, those values represents a low percentage of the total of tracks ($< 3\%$ in both cases).

Resolution on momentum

The track fitting procedure in the TOE algorithm was designed around the needs of the FOOT collaboration. Therefore, the resolution on momentum was the key factor leading to the development of this portion of the algorithm.

Just like the resolution on the β value, it requires Monte-Carlo simulation in order to be extracted: the idea is to established how much the reconstructed values differs from the "real" ones; Monte-Carlo simulations are one way to access this information.

For each correctly reconstructed track, the relative difference between the reconstructed momentum p_R and the Monte-Carlo one p_{MC} can be established. Looking at the deviation from zero of this value gives us the resolution. The resolution on momentum can be evaluated as a function of the momentum itself, but also as a function of the emission angles, θ and φ. The corresponding resolution therefore corresponds to the distribution of N tracks included in a given range.

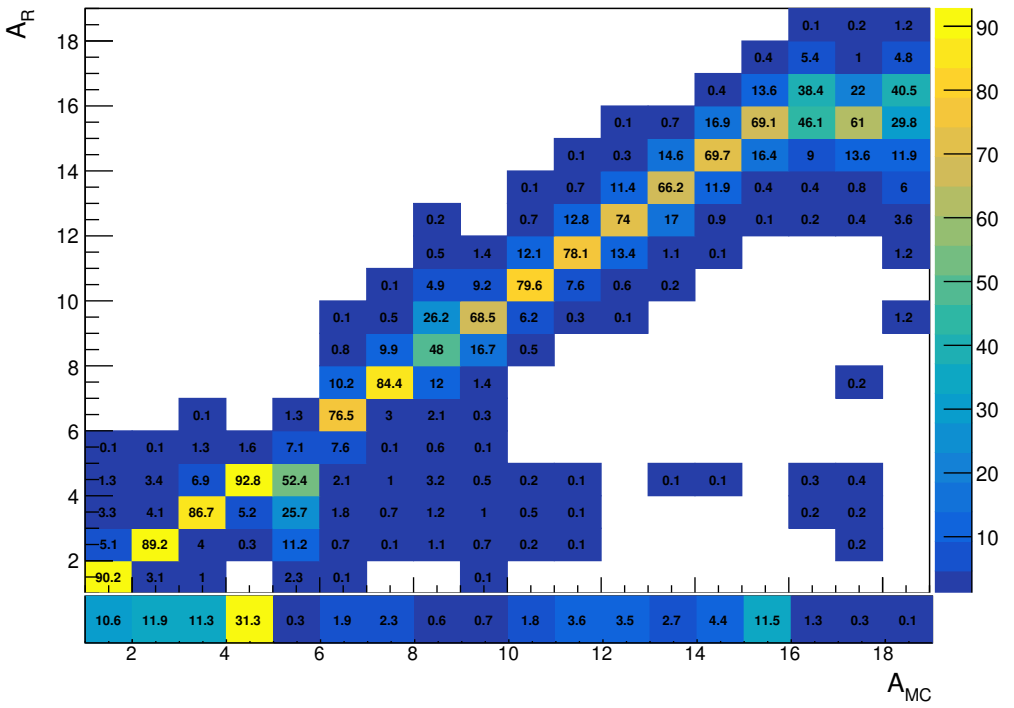

Figure 3.13: Migration matrix after momentum scanning for the data set 16O200C2H4. The reconstructed number of nucleon A_R is shown against the true number of nucleons A_{MC}. Each cell shows the ratio between the number of entries it holds over the total of the column it belongs to, as a percentage. The proportion of reconstructed tracks corresponding to the column is indicated in an additional row below the migration matrix.

$$\mathcal{R}_p(\mathcal{X})|_{\mathcal{X}_{min}\leq\mathcal{X}<\S_{max}} = \sqrt{\frac{1}{N-1}\sum_{n=0}^{N-1}\left(\frac{p_{MC}^{(n)}-p_R^{(n)}}{p_{MC}^{(n)}}\right)^2}, \text{ with } \mathcal{X}=p,\theta \text{ or } \varphi \tag{3.45}$$

Figure 3.15 and 3.16 show the resolutions on momentum for the two simulated data sets with large fragment statistics: 16O200C2H4 and 16O200C. On top, the momentum resolution is presented as a function of the momentum, while on bottom it is shows as a function of the emission angle θ. The resolution of each fragment charge is shown individually.

For the heavier charges, i.e.: $Z > 2$, the limit of 5% wanted by the collaboration is achieved overall, in both data sets.

Tracks with $Z = 2$ have an average momentum resolution value not far

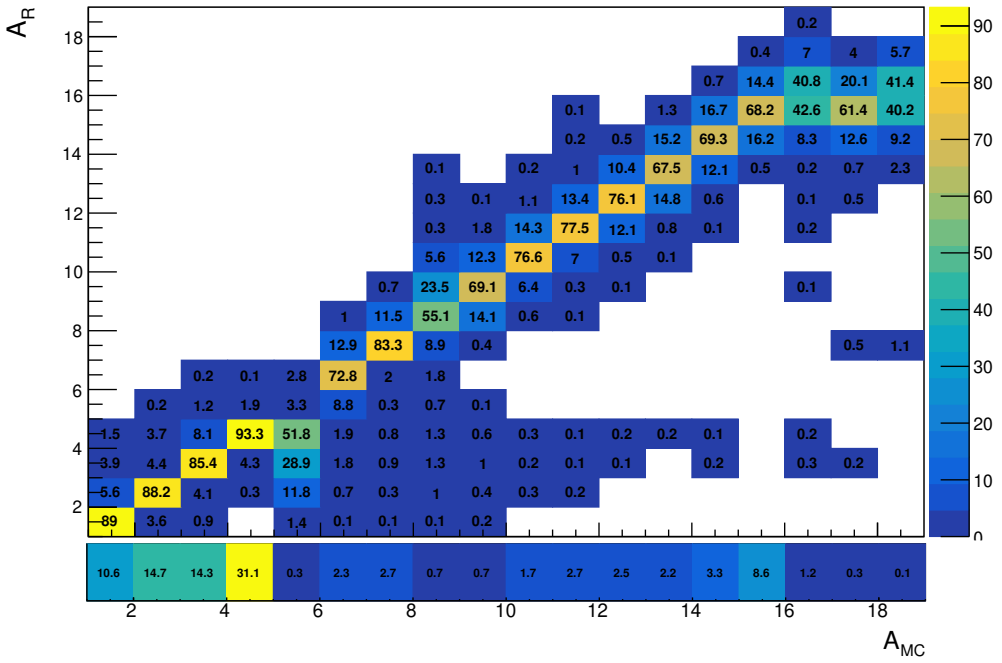

Figure 3.14: Migration matrix after momentum scanning for the data set 16O200C. The reconstructed number of nucleon A_R is shown against the true number of nucleons A_{MC}. Each cell shows the ratio between the number of entries it holds over the total of the column it belongs to, as a percentage. The proportion of reconstructed tracks corresponding to the column is indicated in an additional row below the migration matrix.

from the expected 5%. However, for tracks with low momentum and those with a high θ ($> 10°$), the resolution goes past 10%.

For $Z = 1$, the dependency on the momentum is not as visible as for $Z = 2$, expect for really low momentum values ($p < 0.6 GeV/c/u$), whereas the dependency on θ is clear: the resolution quickly degrades with increasing angle.

3.4 Conclusion

In this chapter, the inner workings of the global reconstruction algorithm developed for the FOOT collaboration were presented. The algorithm is divided into two parts: a first procedure allows for the recognition of the tracks, while the second establishes the parameters of interest, i.e. the momentum and length of the reconstructed track.

The recognition procedure is based around three components: a propagator, which role is to propagate the state of the track under reconstruction along the experimental setup; a filter, which allows the comparison between the estimation of the track state given by the propagator, and the reconstructed clusters in each detection layer; finally, the arborescence keeps track of the history of the reconstruction, and enables the combinatorial approach used in the algorithm. In order to start the reconstruction, several hypothesis need to be formed: an initial value for the charge and the momentum of the particle is required, together with values for the starting position and slopes of the track in the $x - z$ and $y - z$ planes.

The performances of the recognition procedure were evaluated with respect to two main criteria: the efficiency of the reconstruction, exposing the proportion of track recognized by the algorithm that belong to the set of all tracks that could be reconstructed; and the cluster purity, quantifying the performances of the cluster selection. For the four data sets studied, the overall recognition efficiency of the algorithm was found to be around 99%, for a cluster purity above 94%. The main culprits at the origin of the loss in cluster purity are the light fragments, with $Z = 1$ or $Z = 2$. Indeed, their broad angle and energy distributions, and the fact that they are more prone to multiple Coulomb scattering makes their reconstruction difficult.

The second part of the algorithm, the track fitting procedure, makes use of a basic polynomial track model to extract the parameters of interest after reconstruction. This model allows for an easy retrieval of the track length. The momentum is then deduced from this value, and is further refined using a dedicated procedure.

The determination of the track parameters was evaluated through the determination of the momentum resolution achievable by the algorithm. In order to achieve its goals, the FOOT experiment need to achieve a momentum resolution of the reconstructed track lower than 5%. For the heavier incident particles ($Z > 2$), this objective is achieved by the TOE algorithm. For lighter fragments, the performances are more nuanced: the resolution on the momentum ranges between 5% to 40%.

Dans ce chapitre, les rouages de l'algorithme de reconstruction globale développé pour la collaboration FOOT ont été présentés. L'algorithme est divisé en deux parties : une première procédure permet la reconnaissance des trajectoires, tandis que la seconde établit les paramètres d'intérêt, à

savoir la quantité de mouvement et la longueur de la trajectoire reconstruite.

La procédure de reconnaissance est basée sur trois composants : un propagateur, dont le rôle est de propager l'état de la trajectoire en reconstruction le long du dispositif expérimental ; un filtre, qui permet la comparaison entre l'estimation de l'état de la trajectoire donnée par le propagateur, et les amas reconstruits dans chaque plan de détection ; enfin, l'arborescence garde en mémoire l'historique de la reconstruction, et permet l'approche combinatoire utilisée dans l'algorithme. Pour commencer la reconstruction, plusieurs hypothèses doivent être formulées : une valeur initiale pour la charge et la quantité de mouvement de la particule est requise, ainsi que des valeurs pour la position de départ et les pentes de la trajectoire dans les plans x–z et y–z.

Les performances de la procédure de reconnaissance ont été évaluées en fonction de deux critères principaux : l'efficacité de la reconstruction, exposant la proportion de trajectoires reconnues par l'algorithme qui appartiennent à l'ensemble des trajectoires pouvant être reconstruites ; et la pureté, quantifiant les performances de la sélection des amas. Pour les quatre ensembles de données étudiés, l'efficacité de reconnaissance globale de l'algorithme s'est avérée être d'environ 99%, pour une pureté supérieure à 94%. Les principaux coupables à l'origine de la perte de pureté sont les fragments légers, avec Z = 1 ou Z = 2. En effet, leurs larges distributions en angle et en énergie, et le fait qu'ils sont plus enclins aux diffusions coulombiennes multiples rendent leur reconstruction difficile.

La deuxième partie de l'algorithme, la procédure d'ajustement de la trajectoire, utilise une modélisation de la trajectoire par un polynome pour extraire les paramètres d'intérêt après la reconstruction. Ce modèle simple permet de déterminer la longueur de la trace. L'impulsion de la particule est ensuite déduite de cette valeur, et est affinée à l'aide d'une procédure dédiée.

La détermination de la résolution sur la quantité de mouvement reconstruire a permis l'évaluation de la seconde portion de l'algorithme. Afin d'atteindre ses objectifs, l'expérience FOOT doit obtenir une résolution sur l'impulsion de la particule inférieure à 5%. Pour les particules incidentes les plus lourdes ($Z > 2$), cet objectif est atteint par l'algorithme TOE. Pour les fragments plus légers, les performances sont plus nuancées : la résolution sur la quantité de mouvement de ces fragments varie entre 5% et 40%.

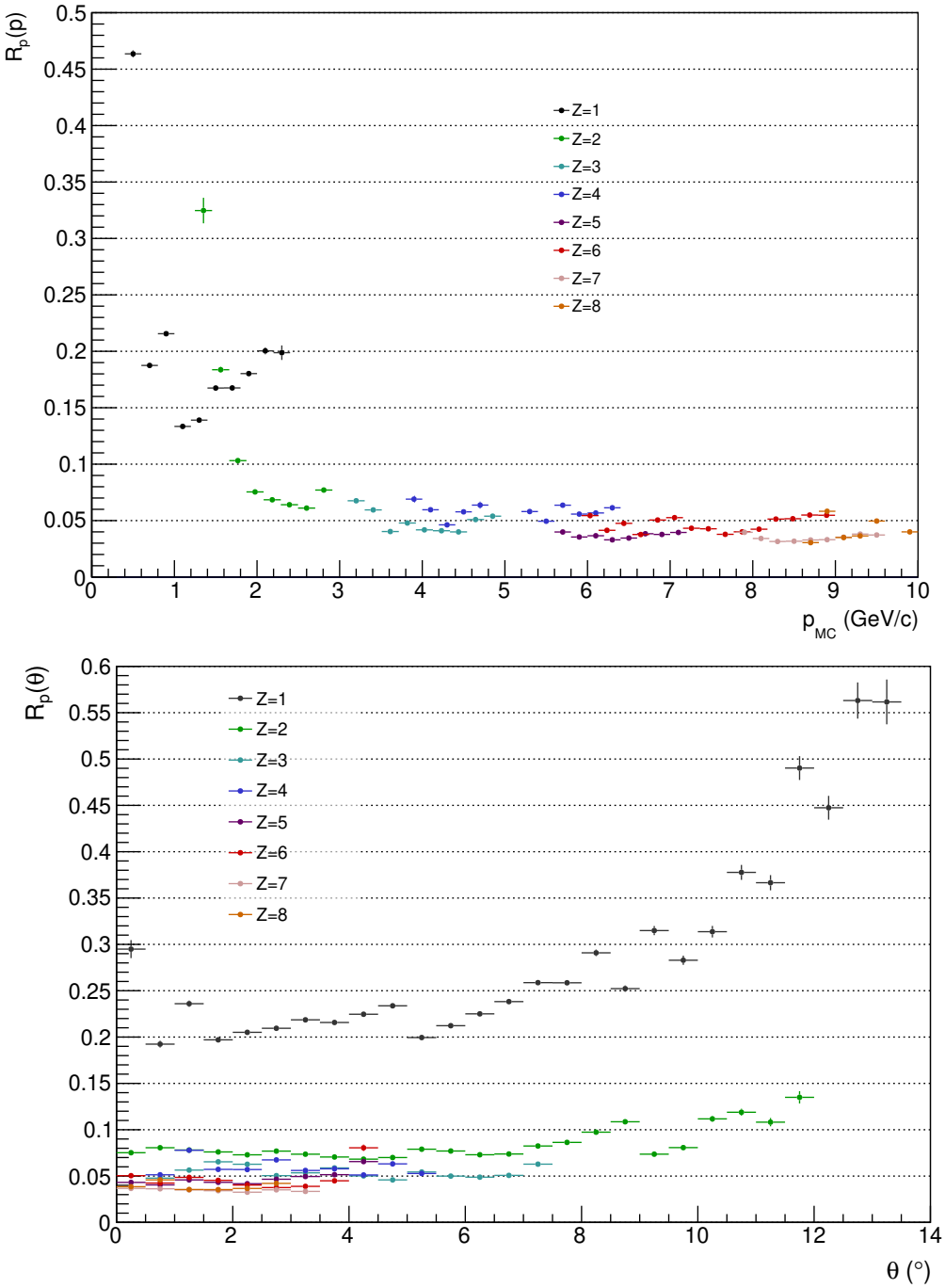

Figure 3.15: Resolution on momentum as a function of momentum on top, and the emission angle θ at the bottom, for the 16O200C2H4 data set. Each fragment charge distribution was computed individually.

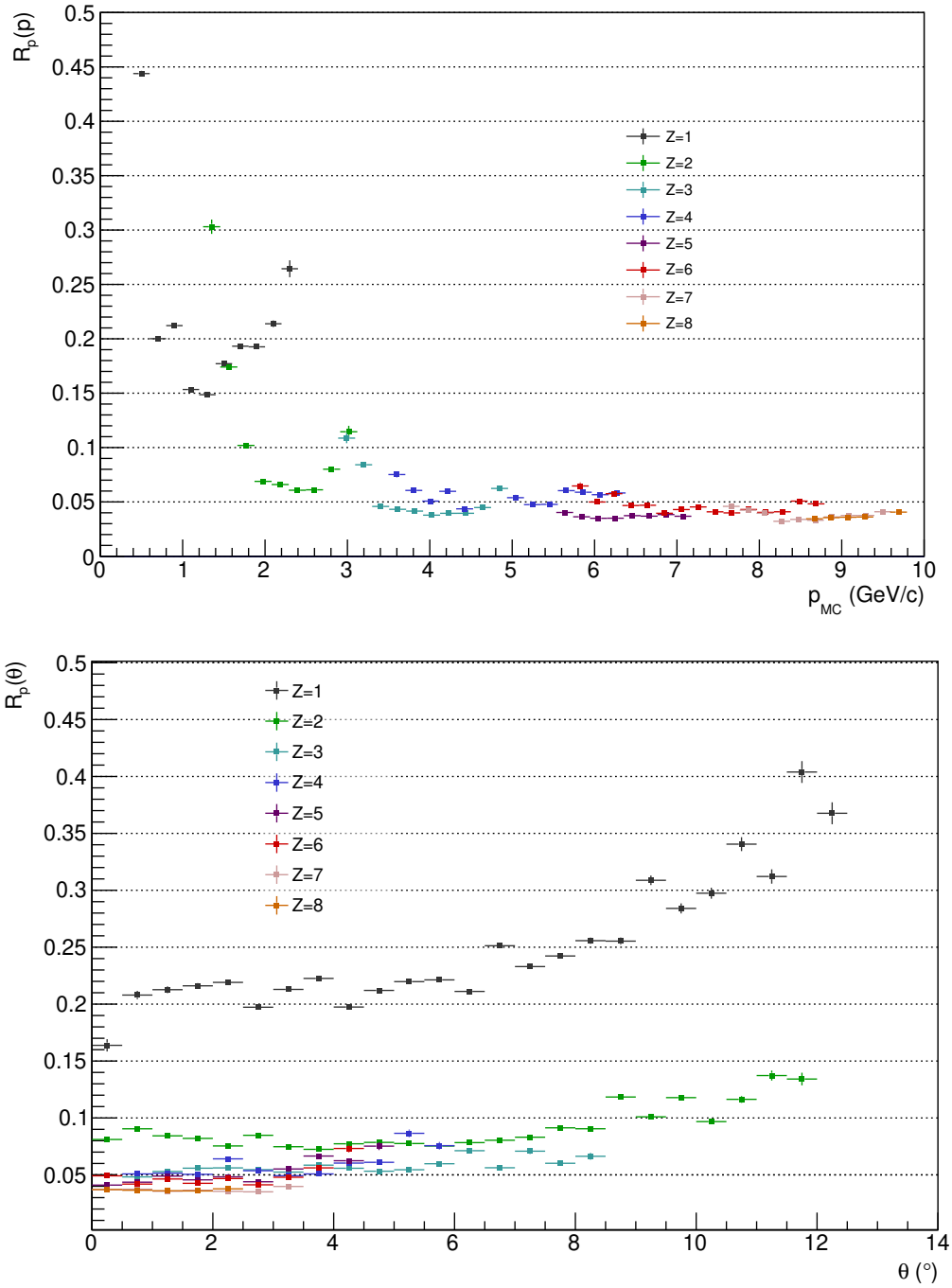

Figure 3.16: Resolution on momentum as a function of momentum on top, and the emission angle θ at the bottom, for the 16O200C data set. Each fragment charge distribution was computed individually.

CHAPTER 4

Unfolding algorithms applied to prompt gamma spectroscopy

Contents

As explained earlier in chapter 1, range uncertainties can compromise some of the physical advantages of hadrontherapy. Currently, and particularly in the case of protontherapy, large range uncertainties lead to the application of substantial safety margins. An extensive study [119] reports safety margins for various center in the U.S., where a relative and an absolute term are usually used in conjunction: the relative term varies from 2.5% to 4.6% of the initial prescribed range, while the absolute term varies from 1 to 3 mm. Those safety margins can influence the treatment planning phase: for example, the choice of beam incidence can be limited because of the potential irradiation of OARs, as is illustrated in figure 4.1.

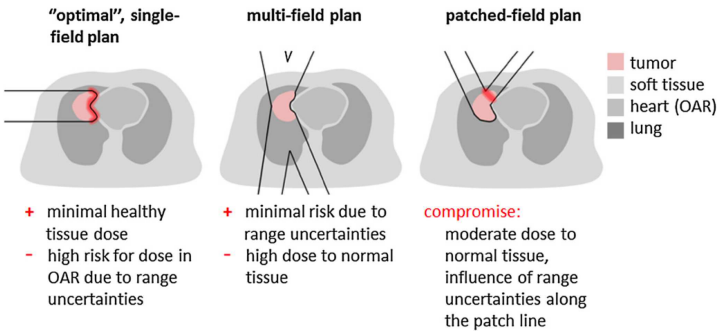

Figure 4.1: Illustration of the influence of range uncertainties in treatment planning for protontherapy. An optimal single field plan is disregarded by the planning system because the range uncertainties could lead to an important irradiation of the heart. A multi-field plan can be devised to minimize the risk to OAR with respect to range uncertainties, but leads to the high dose delivered to healthy tissues. Therefore, a compromise, under the form of a patched field plan (on the right) has to be executed. Taken from [93].

Hence, the reduction of range uncertainties is a crucial factor in order to further improve treatments using this modality. The measurement of

the range of the proton in the patient, so-called *in vivo* range monitoring, is therefore pursued as mean to reduce the size of the safety margins. As mentioned in section 1.4, several modalities exist to this day. They are based on the nuclear reactions underwent by the beam in the tissues. Among them, two rely on the detection of photons emitted after such interactions.

The first one, PET monitoring, relies on the detection of the coincident photon pair emitted after the annihilation of a positron. The positron is itself the result of a beta decay from a radioactive isotope created by the interaction of the beam in the medium it crosses. The second method makes use of the prompt gamma emitted by de-excitation of residuals nuclei left after nuclear reactions. For both of those methods, the longitudinal profile of the emission has been showed to be correlated with the depth dose profile of therapeutic ion beams.

However, the use of prompt gammas has several potential advantages over PET: first, the production of prompt gammas is largely superior to the one of the radioactive isotopes, in protontherapy; second, it has been shown that the emission profile of those gammas is more closely correlated with the dose profile of the beam [94]; third, PET suffers from various effects degrading its potential resolution and usage, such as the lifetime of the radioactive isotopes, or biological washout; fourth, prompt gammas are emitted almost instantaneously following the nuclear interaction, allowing for real time range verification.

In this context, an experiment joining a research group from the University of Trento (INFN) and from the University of Strasbourg (CNRS) was performed to study the feasibility of a novel strategy for real time monitoring in proton therapy using the detection of prompt gammas. Before explaining in details the aim and the nature of the experiment, a brief overview of the prompt gamma scene and its limitations is necessary.

4.1 Prompt gamma monitoring

The idea to use prompt gamma emissions as a mean to monitor treatment was originally formulated in 2003 by Stichelbaut and Jongen at the 39th PTCOG meeting. Three years later, the first study of the concept, and initial proof of its feasibility was published [120]. Since then, various research groups around the world worked on the subject, and several modalities as well as prototypes have been devised. Range information can be extracted from different sources: spatial patterns, or prompt gamma imaging (PGI) has been extensively studied; but prompt gamma spectroscopy (PGS) [121, 122, 123] as well as prompt gamma timing (PGT) [124, 125] have also

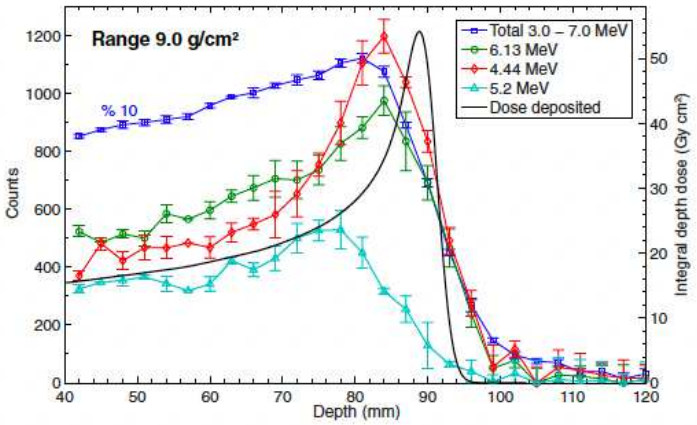

Figure 4.2: Prompt gamma spatial emission spectrum, for all photons with an energy between 3 and 7 MeV, in blue; for photon with an energy of 6.13 MeV in red; for photon with a energy of 4.44 in green and with a energy of 5.2 MeV in light blue. The Bragg curve for the corresponding proton beam is shown in black. Taken from [121].

been proposed. Figure 4.2 shows an example of correlation between the depth dose profile of a proton beam and the corresponding prompt gamma longitudinal profile. The emission profiles from the main gamma lines found in the prompt gamma spectrum are also shown.

Pre-requisites for prompt gamma monitoring

As seen in section 1.4, several approach to beam delivery co-exists in charged particle therapy. Pencil beam scanning (PBS) is the main irradiation strategy where range verification through the emitted prompt gammas is actively considered: in this mode, the targeted volume defined by the treatment plan is scanned in three dimension by a beam focused to a diameter of around 1 cm (FWMH). The beam is usually deflected laterally by dipole magnets, and the variation in depth is ensured by modification of the beam energy. The treatment plan is therefore organized in energy layers, each including a finite number of beam spots of the same beam energy, but with different lateral positioning. These spots are typically delivered within a few milliseconds and separated by beam breaks of roughly the same time. Modification of the beam energy generally takes a few seconds.

The use of beam spots is often considered a pre-requisite for the use of prompt gamma monitoring: each spots has a predefined range, making range verification easier. However it imposes a huge load on the detec-

tion and acquisition system needed to perform the monitoring: the time window for collection of data is quite short, and the corresponding rate of proton delivery is important. Typically, 10^8 protons can be considered as a representative number of protons for the distal spots. Assuming a prompt gamma production yield close to 0.1 per incident proton [126, 127], this translates into 10^7 prompt gamma emitted in 4π solid angle, per spot.

This raw value will consequently be reduced by the detection apparatus solid angle and detection efficiency; it necessitates the use of fast detectors and acquisition systems able to handle the corresponding throughput, characterized by intense and narrow data acquisition periods. Therefore, detection efficiency as well as a proper consideration of the achievable detector load must play a central role in the conception of a detection apparatus [128]. The interplay between achievable field of view and spatial resolution should also be studied.

One other aspect that need to be taken into account in order to achieve prompt gamma monitoring is the nature of the emitted spectrum, and notably its range: it extends from 0 to 10 MeV, with a few prominent gamma lines from the ^{16}O(p,p')^{16}O* (6.13 MeV) and ^{16}O(p, p'α)^{12}C* (4.44 MeV) reactions. This wide range essentially renders the detection systems usually used in nuclear medicine (in SPECT for example) unusable because of the size and weight of the collimators that would be required. Furthermore, in clinical conditions, the treatment room and the beam nozzle in itself contribute to background radiations, most notably under the form of neutron emission, therefore degrading the acquisition conditions for the range verification method.

PGI and collimation designs

Most studies exploit the spatial distribution of the prompt gamma counts to provide range verification. As explained above, this requires specific detection apparatus: to this day, two different categories have been developed. First, the use of mechanical collimation, such as knife-edge system (KES) or multi-slit collimation, has been explored in numerous studies. Second, collimation through electronic means, via the use of so called Compton cameras, was also extensively studied.

Before detailing further what are the benefits and downside of each modalities, it should be pointed out that most method of range verification rely on the characterization of the shape of the prompt gamma profile by comparison with a profile with a high statistical content [129, 130]. In clinical settings, the reference profile needs to be generated by Monte-Carlo simulation (or derivatives, such as the hybrid analytical-Monte-Carlo model

described in [131]), or provided by the TPS. This implies a dependency of the field on the hadronic models used in said simulations, and indeed, most of the study published to this day rely on Monte-Carlo simulation, in one form or another. This point is crucial for the field when several articles have already highlighted the inaccuracy of those specific models in the proper reproduction of both prompt gamma yield and energy spectra [96, 98, 63, 99].

Mechanical collimation The simplest form of mechanical collimation leading to the retrieval of the prompt gamma emission profile is the multi parallel slit camera [132, 133, 134, 129]. When using such a device, the detection apparatus is setup so that the individual slits are orthogonal to the beam axis, most often in a region covering the distal part of the beam spot. This principle requires the need for multiple detectors to cover the space defined by each slit. The acquired profile is then a direct one-dimensional projection of the emission profile for prompt gamma, along the beam path.

Another concurrent modality, more advanced in its development, is the so-called knife-edge slit collimation [126, 130, 135, 136]. It is a variant of the pinhole collimation, made possible because in the case of range verification, one-dimensional emission profile are enough to provide the required information. Again, multiple detectors are required in order to provide spatial information of the emission profile, which is obtained through a one-dimensional reversed projection of the detector array. One of the key advantage of this method is the possibility of performing background estimation with a closed collimator, estimation that can be later leveraged to augment the signal-to-background ratio.

Electronic collimation This form of collimation is achieved mainly in the so-called Compton cameras [137, 134, 138]. In its simplest form, two detection stages are required: a scattering detector, where the incident photon will first be scattered, and a second detection stage where it will be absorbed. When applied to prompt gamma emission, the energy of the incident photon is a priori not known; therefore, because of the underlying physical concept at the origin of such a camera, a total absorption of the scattered photon in the absorber is required. Alternatively, three stages Compton camera can be used to resolve the incident photon kinematics, but at the cost of detector efficiency.

In such a device, a cone-of-origin, using the information on the scattered and absorbed photon, is established. Either backprojection or statistical image reconstruction methods need to be used in order to produce an

emission profile from the several cones thus reconstructed. In this sense, Compton cameras trade reconstruction time cost for increased spatial resolution. However, they are particularly sensitive to noise, which can come from several sources: among them, coincident detection of separate gammas or other particles, or even pair production events, which are not negligible for photon of several MeV of energy, drastically reduce the available statistics.

A final point concerning the elaboration of such devices need to be addressed: understandably, most of the study aiming for the optimization of the detection efficiency of the corresponding prototypes are based on Monte-Carlo simulations. Be it KES [126], multi parallel slit cameras [132], or Compton cameras [138], they all make use of the internal hadronic models of Monte-Carlo simulations to estimate optimal parameters of their respective detector geometry to detect prompt-gamma emission. Once again, benchmarking of those models with appropriate experimental data is of critical importance.

PGS and PGT

As stated above, other means of range verification have been proposed through the years, such as PGS and PGT. These methods, while not exempted of the flaws of PGI, shows a novel approach to the problem of range verification:

- PGS: prompt gamma spectroscopy is based on the identification of characteristic prompt gamma lines and their relative contribution to the detected spectra. The magnitude of the discrete lines can be related to nuclear reaction cross-sections, and by doing so leveraged to deduce the corresponding range. Furthermore, it could potentially be used to evaluate the respective concentration of various element at specific depth in a patient, providing information on the evolution of hypoxic tumors for example. However, this method still has to face the issue of the available statistics to perform its role efficiently.

- PGT: prompt gamma timing uses the time-of-flight distributions of the detected gamma to achieve range verification. Indeed, those distribution are directly correlated with the transit time of the incident proton inside the patient: a change in the mean value of the distribution corresponds therefore to a range modification, while an increased thickness (or density) of the target medium will lead to a broadening of the corresponding time distributions. Once again, this method

relies on statistical counts that might not be achievable in clinical conditions.

The statistical issue

With this overview of the prompt gamma scene, and as clearly stated in a review by *Krimmer et al.* [127], the use of prompt-gamma as a mean for range verification in protontherapy (and more generally charged particle therapy) is severely compromised by the low amount of statistics reasonably achievable in clinical scenarios, regardless of the method selected for monitoring.

As it stands now, the methods that have been developed are heavily impacted by background radiations, a situation which can be remedied by techniques such as time-of-flight windowing, requiring the use of either the information from the beam radiofrequency (when a cyclotron is used to deliver the beam) or dedicated beam monitoring devices [139, 140]. In the case of KES, measurements taken with a closed collimator can provide a similar effect. Applying a selection on the energy range of the detected photons has also been an effective solution proposed to circumvent this issue.

However, all of those methods are performed at the expense of the final detection efficiency, and therefore reduces the available statistics.

Figure 4.3 shows the accuracy of range verification (expressed as the standard deviation of the range estimation) as a function of the number of protons delivered, for three different proton beam energies, using a knife-edge camera. As the number of protons in the spot lowers, the accuracy is reduced. In scenarios with a low amount of available protons, the standard deviation of the range estimation can overshadow the detected range shift. The difficulty of a proper detection of the range shift and correct estimation of said shift, in low statistics conditions, has also been raised in [141, 129].

In order to remedy to the low available statistics in clinical conditions, a few solutions have been proposed through the years: notably, the accumulation of the counts from several beam spots through a Gaussian blurring kernel [136]. However, this can result in the non-detectability of local range shifts due to tissue heterogeneities.

A recent study underlined the fact that the expected use of prompt gamma imaging could be used as an input parameter in TPS, so has to produce a definite number of beam spots with sufficient proton counts to be able to provide range verification [142].

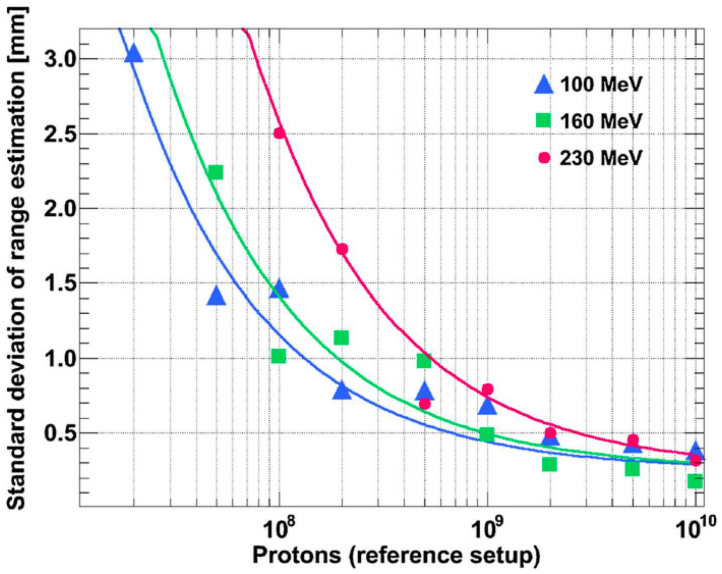

Figure 4.3: Standard deviation of the range estimation as a function of the number of protons in the beam spot. Three different beam energy are considered: in blue triangles, 100 MeV; in green squares 160 MeV; finally, in red dots, 130 MeV. Taken from [126].

Conclusion

Despite the flaws mentioned in the above section, PGI has been successfully used in clinical scenarios in the last few years [135, 136]. Conceptually, the field is highly dependent on Monte-Carlo simulations, which in turn implies the need for a proper evaluation of the hadronic model responsible for the production of secondary particles, including the emission of prompt gamma, by experimental data, over the entirety of the therapeutic energy range [143]. This is of further importance because most detection apparatus prototypes rely, in their optimization phase, on those same models. Additionally, the statistical issue inherent to the method need to be addressed in the following years.

4.2 Experiment

The objective behind the experimental data collected at the CYRCé cyclotron, in Strasbourg, was mainly to established a proof-of-concept for real-time range verification in proton therapy, based on a innovative method aiming for prompt gamma emission enhancement in tumoral volumes. This enhancement is the consequence of the loading of the tumor by drug-delivered

Salt (M)	Density (g/cm^3)	Mass fraction			
$CuSO_4$ (1)	1.249	Cu: 5.08	S: 2.57	O: 82.59	H 9.76
$Y(NO_3)_3$ (1)	1.383	Y: 6.43	N: 3.04	O: 81.58	H: 8.95
NaH_2PO_4 (2)	1.276	P: 4.86	Na: 3.60	O: 82.14	H: 9.40
NaH_2PO_4 (0.2)	1.028	P: 0.60	Na: 0.45	O: 87.99	H: 10.96

Table 4.1: Summary of the various salts and resulting solutions used as targets in the experiment. Their respective concentration, density, and the mass fraction of each element in the solution are reported.

elements, carefully selected because of their specific prompt gamma emission: for this particular experimental run the chosen element were ^{63}Cu, ^{31}P and ^{89}Y. It should be pointed out that the injection of said elements, through techniques such as vectorization, is a proper concern that will need to be addressed in the development of this technique.

On the Strasbourg side of the collaboration, the experiment had additional objectives: it was the first time our detector (a $CeBr_3$ scintillating crystal) was used to detect high energy gammas, therefore the entirety of the acquisition chain had to be validated; the group also aimed towards the benchmarking of the accuracy of the Monte-Carlo models available in the GEANT4 toolkit, through the application of deconvolution algorithms to the obtained spectra.

Experimental setup

The experimental data presented in this chapter have been collected at the CYRCé cyclotron, in Strasbourg, which is able to provide a proton beam of an energy of 24.85 ± 0.08 MeV [144]. The setup is presented in figure 4.4. An aluminum collimator with a diameter of 10mm was used to further focus the beam once it had exited the beam line exit window (50 μm aluminum). The beam was stopped in a T25 cell flask, with a wall thickness in the entrance thinned down to 88 μm (from 1.3 mm).

The flasks contained several possible solutions made of water and salts, such as $CuSO_4$, $Y(NO_3)_3$ and NaH_2PO_4. Water was also irradiated and used as a control point. Table 4.1 summarize the solutions used during the experiment, together with their concentration (in molar M or mol/L), and the respective mass fraction of each element in the final solution.

Two distinct detection apparatus were used to acquire the prompt gamma spectrum: the first one handled by the group from Trento, was a 3"x3" $LaBr_3$:Ce crystal, fully encapsulated, coupled to a photomultiplier tube

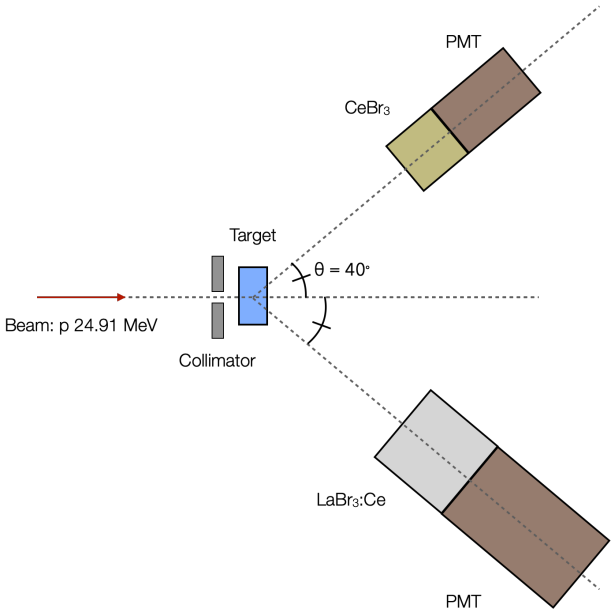

Figure 4.4: Schematic top view of the experimental setup used during the acquisition of data. The aluminum collimator, the T25 flask used as target, and the two detectors are present.

(PMT) (Hamamatsu R6223), polarized to a high voltage of 900V; the second one, managed by the Strasbourg group, was a 2"x2" $CeBr_3$ crystal, fully encapsulated and coupled to a PMT (Hamamatsu R6231-100) as well, but polarized at a high voltage of 800V. The encapsulation of the $CeBr_3$ crystal results in an entrance window made of 400 μm aluminum and around 1 mm of teflon, with a density of 0.5 g/cm^3. Both detectors were placed at 17.5 cm from the center of the target, and formed an angle of 40° with respect to the beam axis.

As stated before, the Strasbourg group was in charge of the $CeBr_3$ detector. The rest of this chapter will therefore be focused uniquely on this part of the setup.

Data acquisition chain

The data acquisition system for the $CeBr_3$ crystal was handled by an 8-channel WaveCatcher board and associated acquisition software. This board is a switched capacitor digitizer based on the SAMLONG chip [145], produced by a collaboration between the IRFU (CEA) and the LAL (CNRS). The signals coming from the PMT are converted to digital signals with a 12-bit resolution at a maximum sample rate of 3.2GS/s. The available

Solution	a	b	c
Water	$(2.18 \pm 0.07) \times 10^{-8}$	$(5.55 \pm 0.07) \times 10^{-4}$	0.05 ± 0.01
$CuSO_4$	$(2.15 \pm 0.05) \times 10^{-8}$	$(5.68 \pm 0.06) \times 10^{-4}$	0.05 ± 0.01
$Y(NO_3)_3$	$(2.16 \pm 0.05) \times 10^{-8}$	$(5.69 \pm 0.06) \times 10^{-4}$	0.05 ± 0.01
NaH_2PO_4 (2M)	$(2.10 \pm 0.05) \times 10^{-8}$	$(5.69 \pm 0.06) \times 10^{-4}$	0.04 ± 0.01
NaH_2PO_4 (0.2M)	$(2.11 \pm 0.05) \times 10^{-8}$	$(5.65 \pm 0.05) \times 10^{-4}$	0.047 ± 0.007

Table 4.2: Summary of the parameters of the second order polynomial used as calibration curve for the spectra acquired during the experiment. Each solution irradiated, and resulting spectrum, has its own individual calibration.

dynamic range is 2.5V.

Once the acquired waveforms were written to file, in binary format, subsequent reading, extraction and analysis of the data was performed by the **surfer_girl** software suite, developed specifically for this purpose, in the context of the work presented here. It is able to convert the data from binary to ROOT compatible format [146]. It can be used to further refine the information and extract relevant measurements, such as amplitude, charge, baseline, rise and fall time of the acquired signal. A constant fraction discriminator can be applied as well. Its capabilities have been extended to encompass analysis tools, such as calibration, smearing, or the application of a deconvolution algorithm (see section 4.4).

Detector calibration and resolution

In order to provide qualitative comparison between acquired spectra with different solutions, in a first time, and comparison with corresponding simulations in a second time, the application of a calibration curve to the experimental data is required. Each spectrum has been calibrated individually, make using of the most prominent gamma lines it is composed of: 0.511 MeV, 4.44 MeV and 3.93 MeV (single escape) and finally 6.93 MeV and 5.62 MeV (single escape).

For all the acquired runs, as slight non-linearity was found in the calibration curves, due to the saturation of the detector and photomultiplier. Therefore a polynomial of order 2 was used to fit the selected calibration points. The curve takes the form: $E(A) = aA^2 + bA + c$ (MeV), with A denoting the extracted amplitude from the signal, in arbitrary units. Table 4.2 summarizes the parameters used for each irradiated solution, and figure 4.5 shows the calibration curve for the water run.

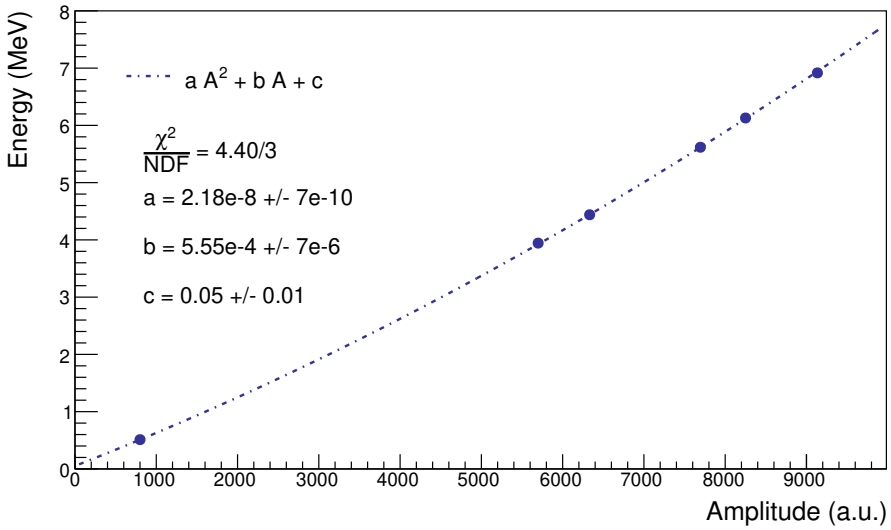

Figure 4.5: Calibration curve of the water spectrum, in energy (MeV) as a function of the amplitude of the signal (a.u.). The several points used for calibration corresponds to the blue dots. The resulting χ^2/NDF of the fit, together with the parameters and their associated errors are also shown.

Furthermore, the comparison between experimental and simulated data requires the extraction of the energy resolution of the detector as the function of the incident photon energy. Indeed, raw simulation output, such as the deposited energy in the detector volume, by each crossing particle, does not include these effects. It is also required in order to construct the system matrix response, to be used in the deconvolution algorithm. Since prompt gammas are the focus of this chapter, the resolution presented here corresponds accordingly to the photon response of the detector.

The energy resolution of the $CeBr_3$ crystal was extracted from previous measurements on a number of radioactive sources. Indeed, most peaks in the acquired spectrum lie on the Compton continuum of others, therefore, the width of the peaks are not a reliable source in order to determine the energy resolution of the detection system. Four different gamma lines were used: 0.511 MeV (^{22}Na), 0.662 MeV (^{137}Cs), 1.274 MeV (^{22}Na) and 1.333 MeV (^{60}Co). The full width at half maximum (FWMH) of each peak, as well as the energy, was extracted using a Gaussian fit. The resolution curve takes the form:

$$\frac{FWMH(E)}{E} = a + \frac{b}{\sqrt{E}} \tag{4.1}$$

Figure 4.6 shows the resolution curve of the $CeBr_3$ crystal and associated

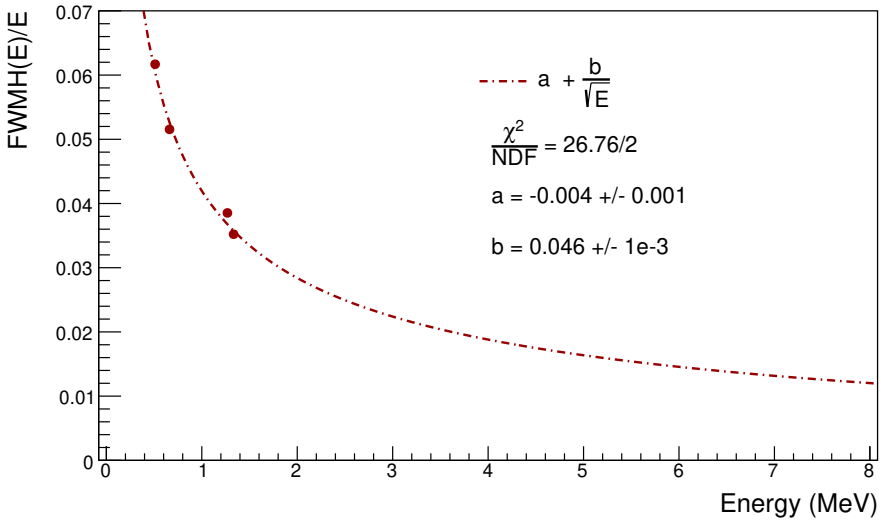

Figure 4.6: Energy resolution of the $CeBr_3$ detector (FWMH(E)/E) as a function of the energy (MeV). The four points used for the fit of the curve correspond to the red dots. The χ^2/NDF result of the fit is shown, together with the resulting parameters and their associated errors.

PMT, polarized at 800V, for incident photons.

With this, the overview of the experiment is complete. The description of the simulation models used in this chapter follows.

4.3 Simulation

The Monte-Carlo simulations presented in this chapter have been performed using the GEANT4 toolkit, version 10.07 [56]. This toolkit, developed by the CERN, is written in C++ and heavily based the object-oriented programming paradigm. Contrarily to most Monte-Carlo simulation toolkit, it is open source. Originally developed for high-energy physics, it has been extended to a lower energy range over the last decades, making its use in the context of medical physics possible.

One of the aim of this chapter is to evaluate the performances of the toolkit in its various possible configurations in the reproduction of prompt-gamma emission. A concise explanation of the steps involved in the propagation of particles in a medium through Monte-Carlo simulations can be found in section 1.2. This section will be dedicated to the details of some of the hadronic models found in the GEANT4 toolkit.

One may consider that only electromagnetic models, handling notably the behavior of photons, are of importance here. However, because of the physical origin of prompt gamma, the hadronic models, handling the production of secondary particles, are the true target of this evaluation: prompt gamma are a consequence of inelastic interaction between the incident charged particle and the target nuclei. Therefore, while different hadronic models have been considered, the remaining models are the same for each physics list: notably, the electromagnetic model used can be found in the class
G4EmStandardPhysics_option4.

At the energy range of interest for our experiment, several models can be used to simulate the interaction between the incident particle and the target nuclei and the steps necessary for the secondary particles thus formed to reach equilibrium. The available models can follow one of two approach: the combination of an intranuclear cascade model, a pre-equilibrium model and a de-excitation model can be use to emulate all the steps aforementioned; it can also be achieved by the use of pre-compiled data coming from an externally generated database, referencing all of the possible reaction channels and their associated output parameters.

Among the model studied in this chapter, two belong to the first category: the Intra-Nuclear Cascade of Liège (INCL) [147, 148, 149] and the BInary Cascade model(BIC) [150], both coupled with the so-called precoumpound model, handling pre-equilibrium emissions, and the same composite de-excitation model. Another model, belonging to the second category, has also been evaluated: the ParticleHP model, based on the TENDL database [151].

In the GEANT4 toolkit, description of physical interactions is done through the implementation of a dedicated physics list. The three physics list studied here, corresponding to the three hadronic models of interest are respectively **QGSP_INCLXX_HP**, **QGSP_BIC_HP** and **QGSP_BIC_AllHP** .

Intranuclear cascade models

Most intranuclear cascade models share the same basic underlying concept: above a pre-defined energy threshold (defined arbitrarly between 10 to 100 MeV), the dynamics of the overall reaction can be described by a sequence of independent interaction between individual nucleons of the target and the projectile, taking place in a mean field potential largely dictated by the target nuclei.

The two models considered here make a distinction between participant and spectator nucleons: a nucleon is considered as participant either if it is part of the incoming particle, or if it has interacted with another participant or was generated during the cascade. The nucleons left in the remnant of the target nucleus are considered as spectators. Collisions between participants are forbidden, both in the INCL and the BIC model.

Both of the models makes use of the same reaction cross-sections to compute collision probability in the cascade, in the energy range of interest: the Barashenkov parametrization, available in the GEANT4 toolkit under the **G4BGGNucleonInelasticXS** class. In both cases, collisions can only be considered if they respect Pauli's exclusion principle; however, the application of this condition differs in the two models.

Nucleons exiting the area defined by the target nucleus are considered as ejectiles for both algorithms: after correction of their energy due to the crossing of the Coulomb barrier, they are no longer tracked. Once the cascade is over, the remaining nucleons inside the target nucleus are fused together in a fragment that will undergo evaporation through the pre-coumpound and de-excitation models.

The rest of this section explore the differences between the two intranuclear cascade models used in this work.

Binary cascade model

In the BIC model, the nucleus is constructed from a set of randomly distributed nucleons, according to one of two nuclear density, based on the total number of nucleons:

- For $A > 16$, a Woods-Saxon form is employed, with $a = 0.545$ fm, $R = r_0 A^{1/3}$, and $r_0 = 1.16(1 - 1.16A^{-2/3})$:

$$\rho(r) = \frac{\rho_0}{1 + exp[(r - R)/a]},\ \rho_0 = \frac{3}{4\pi R^3}\left(1 + \frac{a^2\pi^2}{R^2}\right)^{-1} \tag{4.2}$$

- For $A < 16$, a harmonic oscillator shell model is used, with $R^2 = 0.8133A^{2/3}$ fm^2:

$$\rho(r) = (\pi R^2)^{-3/2} exp(-\frac{r^2}{R^2}) \tag{4.3}$$

The momentum of the nucleon is also selected at random between 0 and the Fermi momentum given by the previously computed nuclear density:

$$p_F(r) = \hbar c(r\pi^2\rho(r))^{1/3} \tag{4.4}$$

In the case of an incident proton, the collective effect of the nucleons form the target nuclei is represented by an optical potential determined by the Fermi momentum, with m the proton mass:

$$V(r) = \frac{{p_F}^2(r)}{2m} \tag{4.5}$$

During the cascade, only the participants are propagated inside the target nucleus, through a Runge-Kutta integration method. In case of a collision, a strict Pauli blocking is applied: the collision is authorized only if the resulting nucleons have a momentum superior to the local Fermi momentum $p_F(r)$.

The cascade process ends when the average energy of all the participant within the nuclear boundary is below a given threshold. It should also be pointed out that the cascade can only occur if the projectile kinetic energy is above a fixed threshold of 45 MeV/u, well above the basic configuration of our experiment. Therefore, the data from the raw BIC model and subsequent de-excitation models does not include any form of intranuclear cascade: a composite nucleus made of the combination of the selected target nucleus and the incident proton is formed and passed down directly the chain to the pre-coumpound model.

Nevertheless, in order to explore the effect of the BIC cascade, a modified version of the algorithm with a kinetic energy threshold set to 15 MeV was used in another set of simulations. In the rest of this chapter, the data coming from the revised algorithm will be referenced as BIClt while the data coming from the default model will be denominated as BIC.

Intranuclear cascade of Liège

The INCL model defines a diffuse nuclear surface, corresponding to a Woods-Saxon nuclear density distribution, up to a maximum distance $R_{max} = R_0 + 8a + r_{int}$, with $a = 0.510 + 1.63 \times 10^{-4} A$ fm, $R_0 = (2.745 \times 10^{-4} A + 1.063) A^{1/3}$ fm and r_{int} a term linked to the total cross section for the interaction under consideration; ρ_0 is taken so that the distribution is normalized to A:

$$\rho(r) = \frac{\rho_0}{1 + exp(\frac{r - R_0}{a})} \text{ for } r < R_{max} \tag{4.6}$$

In this model, the position and momentum determination of the nucleons is not considered independent. Once again the momentum is taken at random between 0 and the value of the Fermi momentum; from this value,

the radius at which the nucleon is placed is computed through the equation:

$$\left(\frac{p}{p_F}\right)^3 = -\frac{4\pi}{3A}\int_0^{R(p)} \frac{\mathrm{d}\rho(r)}{\mathrm{d}r} r^3 \mathrm{d}r \tag{4.7}$$

All the nucleons inside the target nucleus experience a constant potential well, defined by a radius dependent on their momentum. At odds with the BIC model, participant as well as spectators are propagated inside of the target nucleus. As mentioned earlier, collision can be Pauli blocked in this model as well; however, the depletion of the Fermi sphere is taken into account by looking at the local phase-space occupation probabilities.

One of the key aspects of the INCL model is its use of a coalescence model after a nucleon escapes the target nucleus: if other nucleons are sufficiently close in phase space, the initial nucleon can carry along the others, resulting in the emission of a light cluster.

Finally, the INCL model makes use of a self-consistent determination of the stopping time for the cascade, which is parametrized according to the formula:

$$t_{stop} = 29.8A^{0.16}(\mathrm{fm}/c) \tag{4.8}$$

Evaporation models

The purpose of the various evaporation models presented in this section is to bring the products of the cascade towards and equilibrium state, which can be subdivided into two parts: a pre-equilibrium model handles the transition from the kinetic model to the low energy range (a few MeV); after this step, the remnant nucleus from the cascade can be left with a non-null excitation energy shared by its nucleons. The subsequent processes allowing it to reach its ground state are managed by various de-excitation models

Pre-coumpound model

The default pre-coumpound model defined in the GEANT4 toolkit is available in the class **G4PreCoumpound**. It is an exciton model: within those, the excited nuclear state is defined by the excitation energy and a number of excited particles and corresponding holes, also known as excitons [152]. The remnant nucleus configuration from the cascade model is used as the basis for the pre-compound model.

Further transitions and therefore modification of the number of excitons n, with $\Delta n = +2, 0, -2$ is defined according to the partial transition

probability:

$$\lambda_{\Delta n}(n, E) = \frac{2\pi}{\hbar}|M_{\Delta n}|^2 \omega_{\Delta n}(n, E) \tag{4.9}$$

With E the excitation energy of the system, $M_{\Delta n}$ the matrix defining the probability of energy transition due to the disturbed nuclear state, and $\omega_{\Delta n}$ the density of the final state.

Transitions to states with different number of nucleons compete with further emission of nucleons from the remnant nucleus, as well as more complex fragments. The emission probability of the first kind follows the Weisskopf-Ewing model [153], while the second type can only be emitted if several excitons condense during the application of the model, according to the probability:

$$\gamma_N = N^3 \left(\frac{N}{A}\right)^{N-1} \tag{4.10}$$

With N the number of excitons in the cluster and A the number of nucleons inside the remnant nucleus. Only light cluster (up to α particles) can be emitted this way.

De-excitation models

As mentioned earlier, at the level of the de-excitation model, the residual nucleus is characterized by an excitation energy, shared between its nucleons. If this energy is higher than the energy required for the separation of some of the nucleons from the residual, nucleons and light cluster can be ejected. Otherwise, evaporation through gamma emission can take place, which is of particular importance in the context of a study on gamma prompt. The de-excitation models are grouped together in the GEANT4 toolkit in a class called **G4ExcitationHandler**.

Fermi break-up The first model of importance leading to the further emission of light cluster is the Fermi break-up model (FBU). The model provides probabilities for several possible final configurations, based on the density of the phase-space states. In the toolkit, this model can only be applied to nucleus with $Z < 9$ and $A < 17$. It is, in most cases, the model responsible for the production of an α cluster from an excited ^{16}O nucleus, leaving a ^{12}C excited nucleus and the subsequent emission of the 4.44 MeV gamma line.

Particle evaporation: Weisskopf-Ewing and generalized evaporation models If the FBU model cannot be applied to the residual nucleus,

a combination of two models providing probability for the emission of a wide range of fragments, in both their ground and excited state, is used. Both the Weisskopf-Ewing model and the Generalized Evaporation Model (GEM) [154] generates the probability to reach a given configuration through the density of states. The first handles the production of light particles, while the second manages fragment heavier than α particles, up to ^{28}Mg.

Photon evaporation When all other models have failed to produce secondary particle, the residual nucleus is send to the photon evaporation model. The photons emitted during the de-excitation chain can have one of two nature: they are either coming from a continuum gamma transition based on a dipole approximation of the residual nucleus; or they are coming from discrete gamma transitions evaluated from the associated database, provided as an additional data package (**G4PhotonEvaporation5.7**).

Database model

The TALYS Evaluated Nuclear data Library (TENDL), version 1.4, is the basis of the PHP model used in this chapter. This library consists of a set of nuclear reaction data for various light incident particles (up to alpha particles) for $\approx$ 2800 isotopes with an half-life longer than a second, ranging from ^{1}H to ^{291}Mc, for an energy range up to 200 MeV. The data are evaluated using the TALYS nuclear reaction code, and are formatted under the ENDF format.

To produce the library, the input parameters for the TALYS code are adjusted so that the resulting cross-sections agree with experimental data. In some cases, experimental data is directly found in the library. The ouput of the TALYS code include total, elastic and inelastic cross-sections, elastic and inelastic scattering angular distributions, exclusive channel cross sections, energy and double-differential spectra, as well a photon production for discrete states and continuum, among others.

The *modus operandi* of the TALYS code follows in essence the one of the two other models detailed before: it couples a model handling direct reactions with several others in charge of subsequent particle evaporation. However, optical model calculations are the basis for the direct reaction models: they represent the interaction between the incident particle and the target nucleus by means of a complex mean-field potential, which describes all the competing reaction channels. Optical model calculations are performed by the algorithm ECIS-06. The level density is modeled according to a combination between the Constant Temperature Model (CTM) and the Fermi gas model.

TALYS also makes use of a pre-equilibrium model to describe the evolution of the reaction after the first stage of the reaction but before the statistical equilibrium of the coumpound nucleus is reached. The pre-equilibrium model implemented in TALYS is the two-component exciton model. The emissions corresponding to the de-excitation stage are handled by two different mechanisms: the multiple Hauser-Feshbach and multiple pre-equilibrium decays. The first is responsible for the introduction of the emission of prompt gammas.

As mentioned earlier, one of the aim of this chapter is to compare the acquired experimental data and the output of the simulations of the experimental setup, with the various hadronic models presented in this section. In order for the comparison to be qualitative, i.e. to identify which gamma lines are well reproduced by the models and which are not, it is necessary to further process the spectra using deconvolution algorithms. The next section presents an overview of the algorithms used.

4.4 Unfolding algorithms

One of the key issues in the analysis of spectrometric data lies in the extraction of the correct information out of the experimental data: indeed, the obtained spectra are a convolution of different effects that alter the initial spectrum. In our case, the aim is the identification of specific gamma lines, which corresponds to particular nuclear reactions between the incident proton and the medium.

Physical effects such as Compton scattering, pair creation or Doppler broadening underwent by the incident gamma particle will impact the spectrum negatively: specific gamma lines can effectively be masked by those effects and therefore disappear from the resulting spectrum. Furthermore, the introduction of a detector implies an additional perturbation to the initial spectrum, because of the measurement apparatus inherent resolution, the resulting detected spectrum is "blurred"; signals coming from various nuclear reaction can overlap and therefore be indistinguishable by eye.

In order for any analysis to produce conclusive results, these effects should be removed from the detected spectrum, and only then can comparisons between experimental and simulated data be performed. Several so-called unfolding or deconvolution algorithms have been proposed through the years to solve this specific problem [155, 156, 157].

A convolution integral equation takes the following form, with $y(t)$ the measurement, $x(t)$ the initial value and $h(t)$ the impulse response of the

detector:

$$y(t) = \int_{-\infty}^{\infty} x(\tau)h(t-\tau)\mathrm{d}\tau \tag{4.11}$$

Which can be rewritten under the matrix form:

$$\boldsymbol{y} = H\boldsymbol{x} \tag{4.12}$$

Unfolding belong to the category of ill-posed problems: several different functions can solve a convolution equation. In consequence, the direct deconvolution methods are particularly sensitive to noisy data: small errors can lead to important oscillations in the result of the algorithm. In consequence, regularization techniques must be employed, but in doing so, only an approximate solution can be obtained [158].

Two different algorithms have been re-implemented for the purpose of the work presented in this document: the Gold [159] and the Richardson-Lucy [160, 161] deconvolution algorithms. An additional boosting procedure, that can be applied to either one of them was also put in place, as described in [162]. Both of those algorithms belong to the iterative classification (as opposed to the direct methods). They are also both positive definite: if the input provided is positive, the solution will be as well. Those two points makes them particularly suited for the task at hand.

Gold unfolding algorithm: The value of the estimated spectra $\hat{\boldsymbol{x}}$ at step $i+1$ is obtained through the following equation, where $A = H^T H$ and $\boldsymbol{y}' = H^T \boldsymbol{y}$:

$$\hat{x}_{i+1}(k) = \frac{y'(k)}{\sum_{l=0}^{N-1} A_{kl}\hat{x}_i(l)}\hat{x}_i(k) \tag{4.13}$$

Richardson-Lucy unfolding algorithm: This algorithm follows the Bayesian statistical approach, and the estimated spectra $\hat{\boldsymbol{x}}$ at step $i+1$ can be computed through the equation:

$$\hat{x}_{i+1}(k) = \hat{x}_i(k)\sum_{l=0}^{N-1} H_{kl}\frac{y(l)}{\sum_{m=0}^{N-1} H_{ml}\hat{x}_i(m)} \tag{4.14}$$

Procedure

The boosting procedure, introduced earlier, is required in order to concentrate the area of the peak found by deconvolution into their correct energy. It also allows for the reduction of fake peaks that can arise as artifacts of the procedure. Indeed, disentangling the contribution of different gamma

lines in spectra composed of overlapping peaks is a challenging task, a situation that becomes even more complex when single and double escape peaks are present. However, left unchecked, the statistical properties of the initial spectrum will not be preserved; an additional step, to normalize the spectrum, as been implemented with respect to the original method.

The overall algorithm, using either the Gold or the Richardson-Lucy method and including the boosting procedure, can be described as follows:

① Setup initial values for the estimated spectra: $\hat{\boldsymbol{x}}_{0,0} = [1, 1, \ldots, 1]^T$

② Either: start loop on R repetitions, i.e. $r = 1$, or increment r value

③ Either: start loop on I iterations, i.e. $i = 1$, or increment i value

④ Compute new value for the estimated spectra $\hat{\boldsymbol{x}}_{i,r}$ through the application of the selected algorithm (using either equation (4.13) or (4.14))

⑤ If $i \neq I$, go back to ③

⑥ Set $i = 1$

⑦ Apply boosting procedure, with p the boosting coefficient:

$$\hat{x}_{1,r+1}(k) = [\hat{x}_{I,r}(k)]^p\,, \text{ with } k \in [0, N-1] \tag{4.15}$$

⑧ If $r \neq R$, go back to ②

⑨ Apply renormalization procedure, with N_y be the number of counts in the measured spectra, and $N_{\hat{x}_{I,R}}$ the one in the deconvoluted one:

$$\hat{x}(k) = \hat{x}_{I,R}(k)\frac{N_y}{N_{\hat{x}_{I,R}}}, \text{ with } k \in [0, N-1] \tag{4.16}$$

Response matrix

As seen in this chapter, the response matrix H of the detector is the central point of any deconvolution algorithm. In our case, it is obtained by leveraging the possibilities offered by Monte-Carlo simulations: in a dedicated simulation run, where the geometry detailed in section 4.2 was used, an isotropic gamma source over 2π in the direction of the beam, with a uniform energy distribution between 0 and 10 MeV, was implemented.

The initial position of the incident photons was selected at random in a cylinder of 7 mm of diameter, reaching from the entry of the target to 1.1

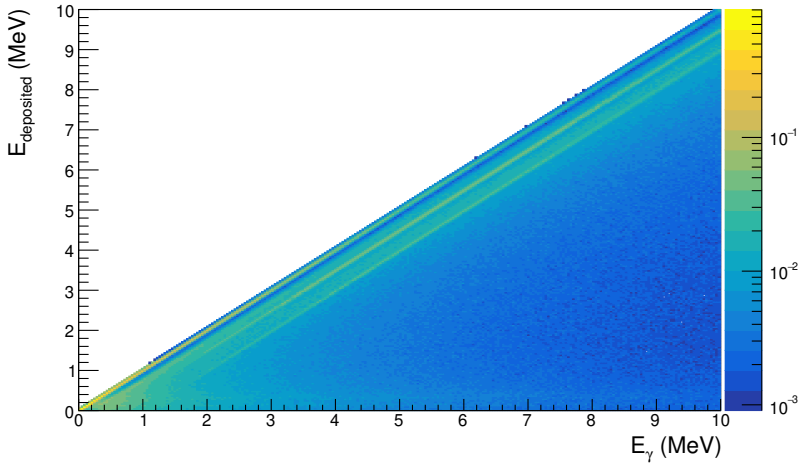

Figure 4.7: Response matrix used in the deconvolution algorithms. The deposited energy (MeV) is shown with respect with the initial energy of the photon (MeV). The matrix is normalized column by column.

cm along the beam axis, in order to encompass the possible emission range of the prompt gammas up to the Bragg peak of the proton beam. The resulting deposited energy in the $CeBr_3$ crystal, taking into account the detector resolution evoked in section 4.2, and the associated initial energy of the incident photon, were registered for all events reaching the detector. A total of 10^9 initial photons were generated.

The response matrix was obtained by scoring the number of photons with initial energy E_γ depositing energy $E_{deposited}$. Once the full matrix relating the deposited energy to the incident one was obtained, each column was furthermore normalized so that the sum of all entries in the column was equal to unity. In this way, the cell $H_{i,j}$ of the response matrix depicts the probability for a photon with incident energy $E_\gamma^{(j)}$ to be detected with an energy $E_{deposited}^{(i)}$. The response matrix used here in the unfolding algorithms can be found in figure 4.7.

4.5 Results

The data acquired during the experiment where subject to a hardware threshold on the minimal amplitude required for a signal to be registered. However, it does not corresponds to a fixed energy threshold for each spectrum, since the calibration varies in between runs. Therefore, the study of the acquired spectra requires the application of a software threshold, set at

0.7 MeV. In consequence, the same configuration, and therefore threshold, has been applied in all the results presented in this section.

Due to the acquisition system dead time, estimating the yields of the prompt gamma was not achievable with the acquired data. Therefore, in the following, all the spectra shown, deconvoluted or not, coming from experiment or simulation, are normalized to their integral, in order to be able to compare their shapes.

Unfolding algorithm selection and parametrization

In order to determine which algorithm to use in order to retrieve deconvoluted spectrum from the data acquired during the experiment, one can perform a basic test aiming to check how well a known distribution can be retrieved by applying a deconvolution algorithm on the corresponding measured distribution.

Figure 4.8 shows both the measured distribution and the corresponding initial distribution of photon energy used in the determination of the response matrix. Additionally, the output of the classic Gold and Richardson-Lucy algorithms, applied on the measured distribution, are also present. Here, 50 iterations were enough to reach a good agreement between the expected distribution and the one from the algorithms. It should be noted, however, that the Gold algorithm struggles on the higher energy range, where a slight discrepancy can be found.

Both of those algorithms, in their classic form and in their boosted one have been applied to the water spectrum obtained during the experiment. A total of 50 iterations have been applied for all of the algorithms: in the case of the boosted ones, 5 iterations, 10 repetitions and a boosted coefficient of 1.1 was used. Figure 4.9 shows the calibrated water spectra on top, and the results of the unfolding algorithms on the bottom: on the left, the Gold (and boosted Gold) was applied, while on the right, the Richardson-Lucy (and boosted Richardson-Lucy) was used. The three main gamma lines resulting from an ^{16}O(p,p'γ)^{16}O reactions, of respective energies 7.12, 6.92, 6.13 MeV, and the one generated by the ^{16}O(p, p' α γ)^{16}O reaction, with an energy of 4.44 MeV have been marked by a down arrow.

From those two figures, a few conclusions regarding the deconvolution algorithms arise: first, the Gold family of algorithms fails to accurately represents the two higher energy peaks coming from nuclear reaction on ^{16}O, while they are present in the spectrum obtained by unfolding with the Richardson-Lucy algorithms; second, the boosting operation allows for a faster convergence towards specific gamma lines. Therefore, in the follow-

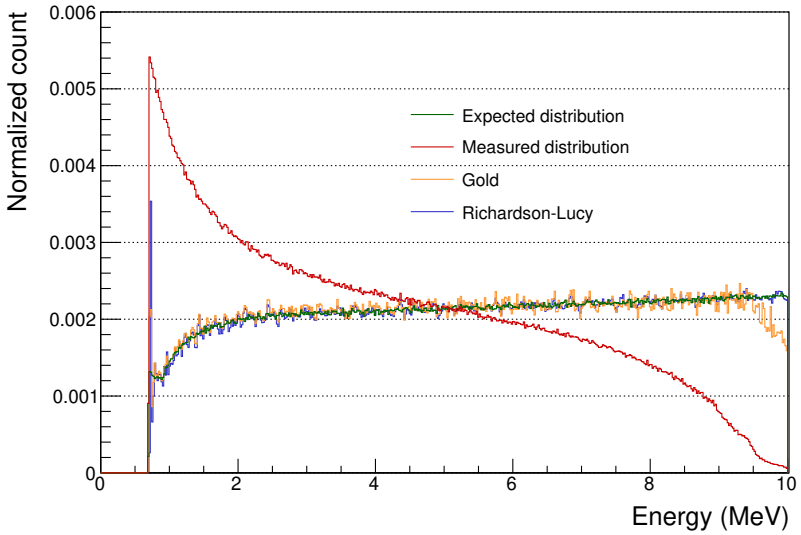

Figure 4.8: Normalized count as a function of the energy (MeV), for the expected distribution (in green) and measured distribution (in red) used for the generation of the response matrix. The results of the classic Gold and Richardson-Lucy algorithms is also shown, in yellow and blue, respectively.

ing, only the boosted Richardson-Lucy algorithm will be used to investigate the acquired data.

It should also be pointed out that remnant artifact peaks are present in the deconvoluted spectra, with this number of iterations. Notably, the pronounced peak around 6.5 MeV corresponds to no known gamma-lines given the available target nuclei. In order to fully erase it, 10 iterations, 10 repetitions and a boosting coefficient of 1.2 is necessary. In consequence, those parameters have been used in the results presented below.

Unfolding applied to water spectrum

Once optimal parameters for the deconvolution of the acquired spectra are obtained, the algorithm can be applied, in a first time, to the water spectrum. Its careful study allows to set a comparison point for the spectrum acquired for the several solutions. As can be seen in figure 4.10, several gamma lines are dominating the spectrum, on the higher energy range (above 4 MeV): they correspond to emission of prompt gamma from oxygen nuclei. At lower energies, the spectrum shows several lines, closely spaced in energy. Most of them have been identified as coming from nuclear reactions of the beam with the aluminum collimator.

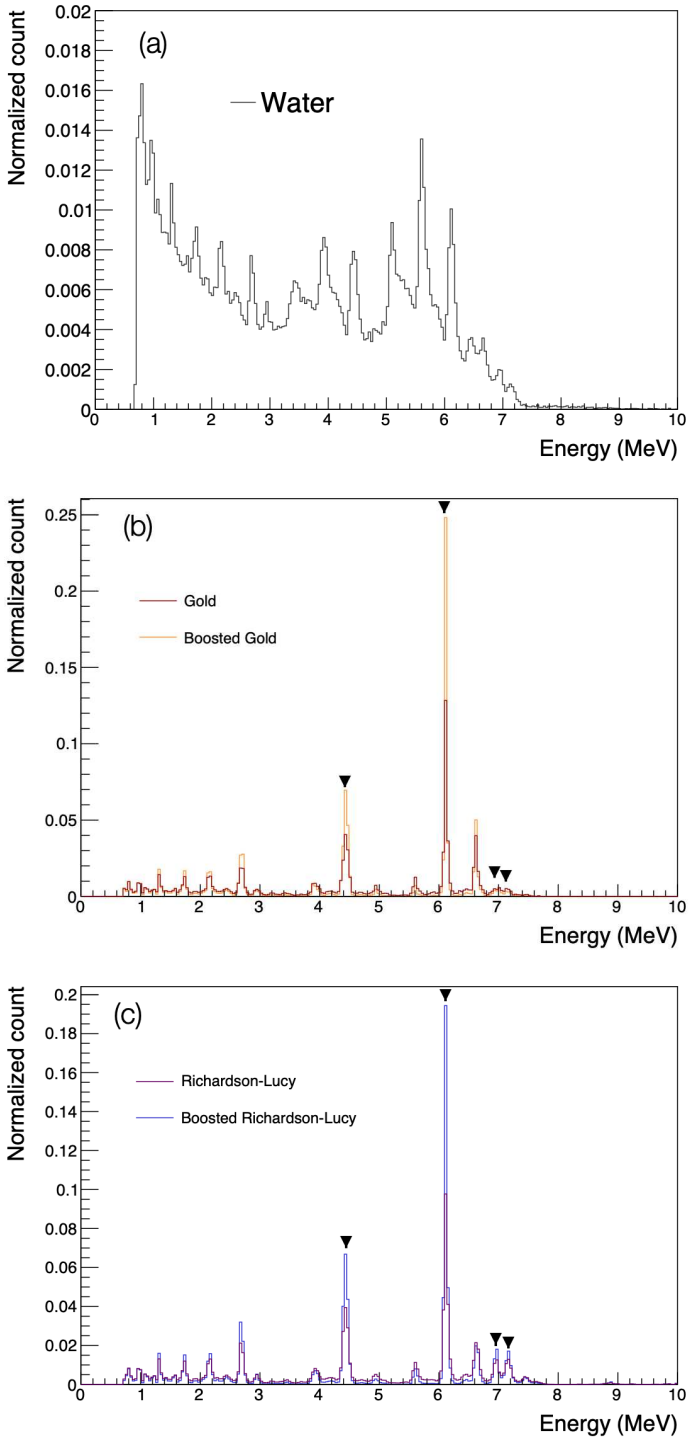

Figure 4.9: Calibrated spectrum for the target made uniquely of water in (a). The result of the application of the deconvolution algorithms to the water spectrum is also shown: (b) corresponds to the Gold and boosted Gold algorithms; while (c) corresponds to the Richardson-Lucy and boosted Richardson-Lucy algorithms. The down arrows are marking the main gamma lines produced after nuclear reaction between a proton and an oxygen nuclei, expected to be found in the unfolded spectrum: 7.12, 6.92, 6.13 and 4.44 MeV.

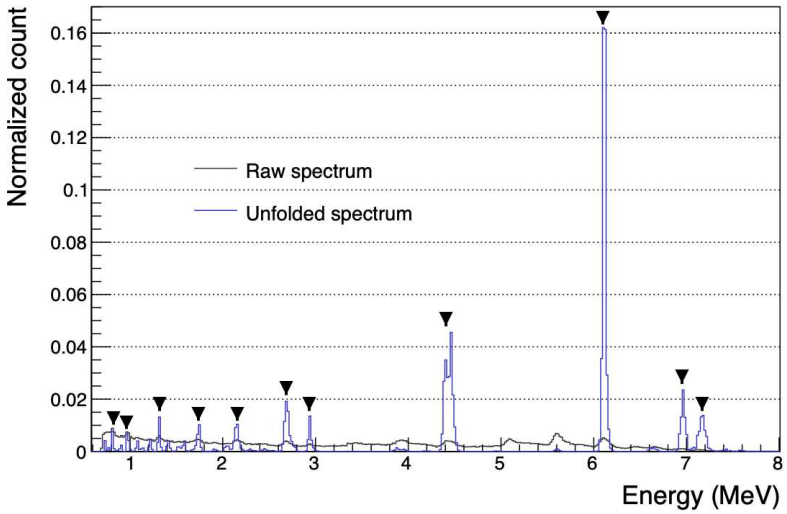

Figure 4.10: Raw water spectrum from acquired data (in black) and its unfolded version (in blue), resulting from the application of the boosted Richardson-Lucy algorithm, using optimal parameters. The down arrows correspond to the lines that have been investigated in this section.

With the result presented in figure 4.10, the energy of those specific gamma lines and their respective contribution to the acquired spectrum can be extracted. They have each been individually fitted with a Gaussian function: from this fit, the mean value of the line is directly available, and the proportion of counts of the line has been obtained by computing the integral of the function over three sigmas, from each side of the mean value. Only the principal visible lines have been investigated, and are denoted by the down arrows in the figure.

These results are reported in table 4.3, together with one or several possible reactions leading to the emission of the corresponding prompt gamma. These possible reactions are ranked in order of probability, and are based on the NNDC NuDat 3.0 database and [163]. By comparing the extracted energy of the line and the one found in the literature, it is obvious that the applied calibration is off by a few tens of keV.

Comparison between water and solutions

The comparison between the water spectrum and the one obtained for each solution is necessary in order to identify the gamma lines coming from the salts used in the solutions. This is especially crucial with the presence of the

Extracted energy (MeV)	Extracted proportion (%)	Proposed reactions	Energy (MeV)
7.1680 ± 0.0004	5.76 ± 0.07	$^{16}O(p,p'\gamma)^{16}O$	7.12
6.9625 ± 0.0003	6.34 ± 0.07	$^{16}O(p,p'\gamma)^{16}O$	6.92
6.11273 ± 0.00007	39.2 ± 0.1	$^{16}O(p,p'\gamma)^{16}O$	6.13
4.4395 ± 0.0003	18.3 ± 0.1	$^{16}O(p,p'\alpha\gamma)^{12}C$	4.44
2.9396 ± 0.0003	2.28 ± 0.04	$^{27}Al(p,p'\gamma)^{27}Al$	2.98
2.6923 ± 0.0003	5.72 ± 0.06	$^{16}O(p,p'\gamma)^{16}O$	2.74
2.1495 ± 0.0004	2.97 ± 0.06	$^{27}Al(p,p'\gamma)^{27}Al$	2.21
		$^{27}Al(p,n\gamma)^{27}Si$	2.16
1.7384 ± 0.0004	1.98 ± 0.03	$^{27}Al(p,\gamma)^{28}Si$	1.78
		$^{27}Al(p,p'\gamma)^{27}Al$	1.72
1.3162 ± 0.0002	1.82 ± 0.04	$^{27}Al(p,d\gamma)^{26}Al$	1.34
		$^{27}Al(p,\alpha\gamma)^{24}Mg$	1.37
0.9751 ± 0.0002	1.18 ± 0.05	$^{27}Al(p,p'\gamma)^{27}Al$	1.02
		$^{27}Al(p,d\gamma)^{26}Al$	1.06
		$^{27}Al(p,n\gamma)^{27}Si$	0.96
0.8124 ± 0.0003	0.89 ± 0.03	$^{27}Al(p,p'\gamma)^{27}Al$	0.84
		$^{27}Al(p,d\gamma)^{26}Al$	0.83
		$^{27}Al(p,n\gamma)^{27}Si$	0.78

Table 4.3: Summary of the main gamma lines found in the water spectrum after deconvolution. Their energy, and contribution to the total spectrum are reported. For each line, one or several reactions are proposed as origin of the line, together with the corresponding energy found in literature.

gamma lines from the aluminum collimator, which cover the energy region where the prompt gamma from the selected elements are emitted.

Yttrium solution

A comparison of the deconvoluted spectra from water and the $Y(NO_3)_3$ [1M] solution can be found in the top part of figure 4.11. The overall match between the shape of the two spectra is really good. On the bottom part of figure 4.11, a zoom on the 0.6 to 3 MeV region as been applied. A few disparities appears: the yttrium solution present two peaks that are not apparent in the water spectrum, respectively around 1.02 and 1.67 MeV; additionally, the peak at 1.57 is more pronounced than the one for the water spectrum.

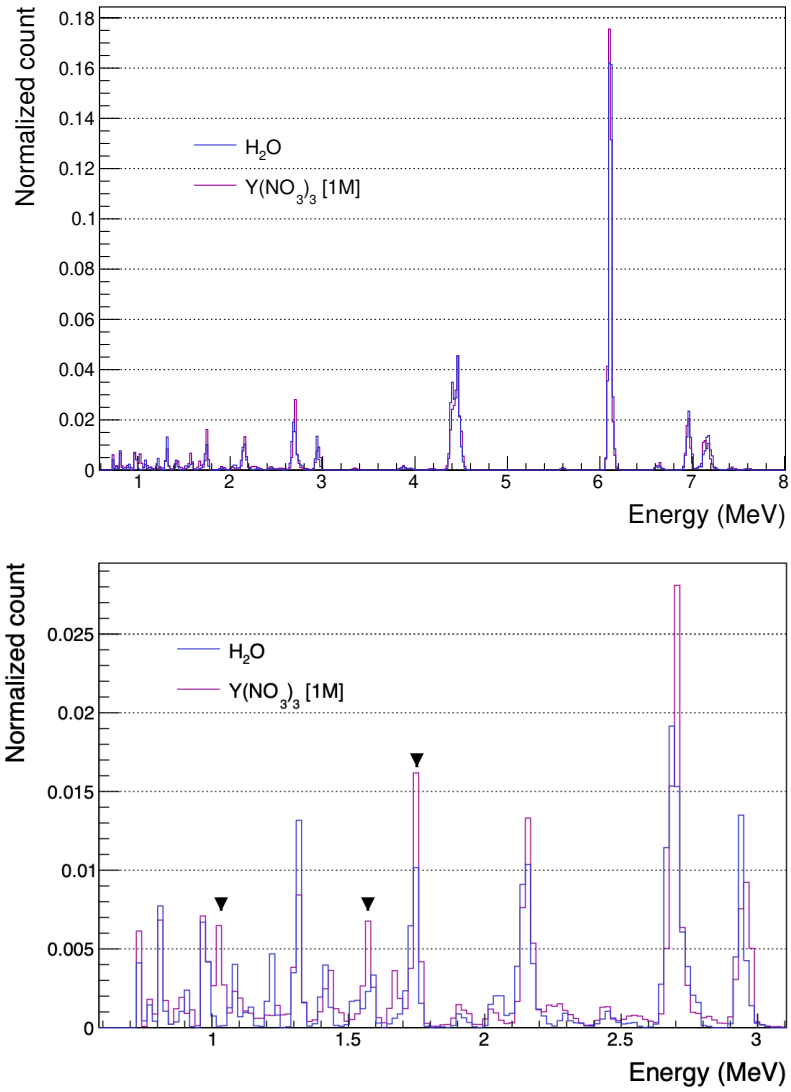

Figure 4.11: Top, comparison between the deconvoluted water spectrum and the one for the yttrium solution. On the bottom, a zoom on the energy range 0.6 to 3.1 MeV as been applied. The down arrows highlight the gamma lines where a difference between the water and the solution spectra is visible.

Extracted energy (MeV)	$Y(NO_3)_3$[1M] (%)	Water (%)	Proposed reactions	Energy (MeV)
1.6762 ± 0.0004	0.79 ± 0.01		$^{89}Y(p,p'\gamma)^{89}Y$	1.75
			$^{89}Y(p,p'\gamma)^{89}Y$	1.71
			$^{89}Y(p,\gamma)^{90}Zr$	1.76
1.5715 ± 0.0002	1.31 ± 0.02	0.76 ± 0.02	$^{14}N(p,p'\gamma)^{14}N$	1.64
1.0264 ± 0.0002	1.28 ± 0.02		$^{89}Y(p,2n\ \gamma)^{88}Zr$	1.06
			$^{89}Y(p,d\gamma)^{88}Y$	1.09

Table 4.4: Summary of the additional gamma lines found in the spectrum of the $Y(NO_3)_3$ [1 M] solution, by comparison with the water spectrum. The extracted energy, the contribution of the line to the total of the spectrum as well as the corresponding contribution in the water spectrum, when the line exists, are presented. Several possible nuclear reactions at the origin of the lines are also proposed, together with their corresponding energy.

Table 4.4 recapitulates the extracted energy of those peaks, as well as their contribution (in %) to the total spectrum. The corresponding contribution in the water spectrum is also indicated, when the peak is present. Several possible reactions from which those peaks could originate are also proposed: reactions on both ^{89}Y and ^{14}N have been included. The variations in the contribution to the total spectrum with respect to the spectrum from water is of the order of a few percents at most.

Copper solution

A comparison of the deconvoluted spectra from water and the $CuSO_4$ [1M] solution can be found in the top part of figure 4.12. Again, the overall match between the shape of the two spectra is really good. On the bottom part of figure 4.12, a zoom on the 0.6 to 3 MeV region as been applied. A few disparities, not as strong as the one for the yttrium solution, appears: the copper solution presents three peaks that are more marked than their respective counterpart in the water spectrum, around 0.96, 1.75 and 2.16 MeV; additionally, a extra peak can be found at 1.68 MeV.

Table 4.5 summarizes the extracted energy of those peaks, their contribution (in %) to the total spectrum, as well as their contribution in the water spectrum, when the peak is present. Several possible reactions from which those peaks could originate are also proposed: they include reactions on ^{63}Cu and ^{32}S. Once again, the variation in the contribution to the total spectrum with respect to the spectrum from water is of the order of a few percents at most.

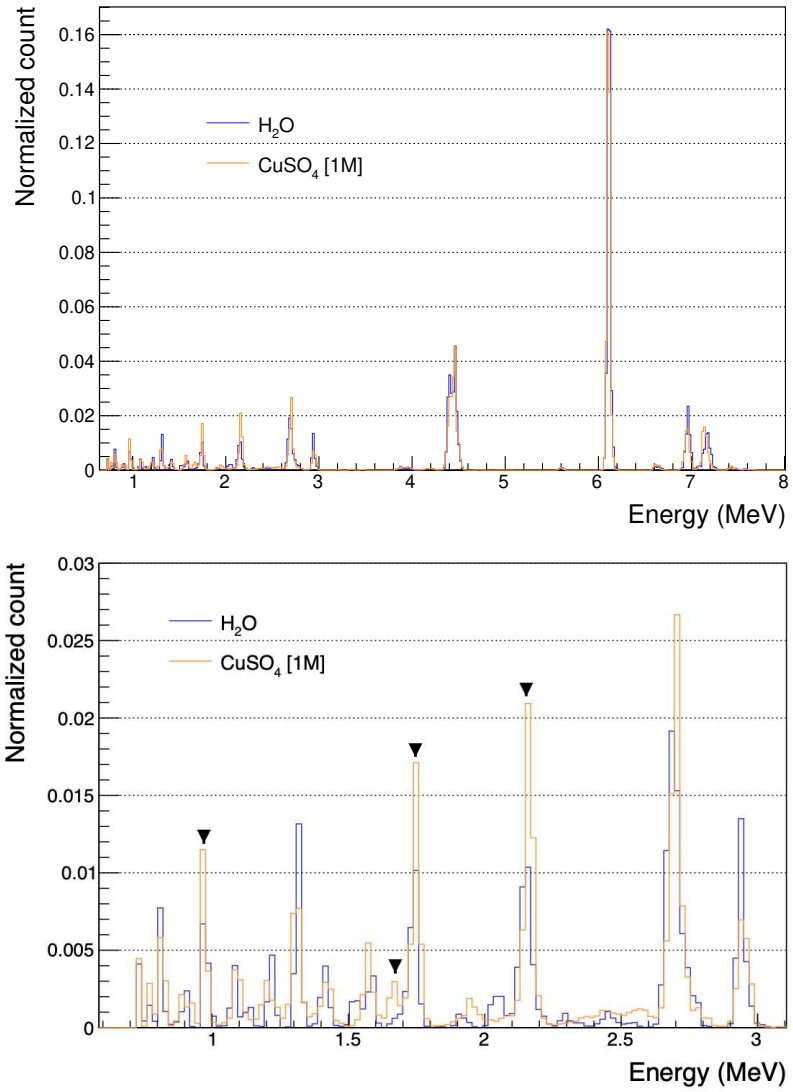

Figure 4.12: Top, comparison between the deconvoluted water spectrum and the one for the copper solution. On the bottom, a zoom on the energy range 0.6 to 3.1 MeV as been applied. The down arrows highlight the gamma lines where a difference between the water and the solution spectra is visible.

Extracted energy (MeV)	$CuSO_4$[1M] (%)	Water (%)	Proposed reactions	Energy (MeV)
2.1620 ± 0.0001	4.25 ± 0.04	2.97 ± 0.05	$^{32}S(p,p'\gamma)^{32}S$	2.23
			$^{32}S(p,d\gamma)^{31}S$	2.23
1.7478 ± 0.0002	2.87 ± 0.03	1.98 ± 0.04	$^{63}Cu(p,\gamma)^{64}Zn$	1.80
1.680 ± 0.002	1.06 ± 0.04		$^{63}Cu(p,n\gamma)^{63}Zn$	1.69
0.9720 ± 0.0001	1.54 ± 0.02	1.18 ± 0.05	$^{63}Cu(p,p'\gamma)^{63}Cu$	0.96
			$^{63}Cu(p,\gamma)^{64}Zn$	0.99

Table 4.5: Summary of the additional gamma lines found in the $CuSO_4$ [1M] solution, by comparison with the water spectrum. The extracted energy, the contribution of the line to the total of the spectrum, as well as the corresponding contribution in the water spectrum, when the line exists, are presented. Several possible nuclear reactions at the origin of the lines are also proposed, together with their corresponding energy.

Phosphorus solutions

A comparison of the deconvoluted spectra from water and the NaH_2PO_4 [2M] and NaH_2PO_4 [0.2M] solutions can be found in the top part of figures 4.13 and 4.14, respectively. For both of those, the overall match between the shape of the spectrum from the solution and the one of water is really good. On the bottom part of the figures 4.13 and 4.14, a zoom on the 0.6 to 3 MeV region as been applied.

The spectrum from the solution with the NaH_2PO_4 salt at a low concentration presents almost no differences with the spectrum from water. However, at a higher concentration, disparities emerge: the solution presents three peaks that distinguish from their counterpart in the water spectrum, around 1.22, 1.58 and 2.17 MeV. Table 4.6 reports the extracted energy of those peaks, their contribution (in %) to the total spectrum, as well as their contribution in the water spectrum. Several possible reactions from which those peaks could originate are also proposed: reactions on ^{23}Na as well as ^{31}P are considered. Once again, the variation in the contribution to the total spectrum with respect to the spectrum from water is of the order of a few percents at most.

Comparison between experiment and simulations

The second objective of this chapter is the evaluation of the hadronic models used in the GEANT4 toolkit. Again, only a comparison of the contribution

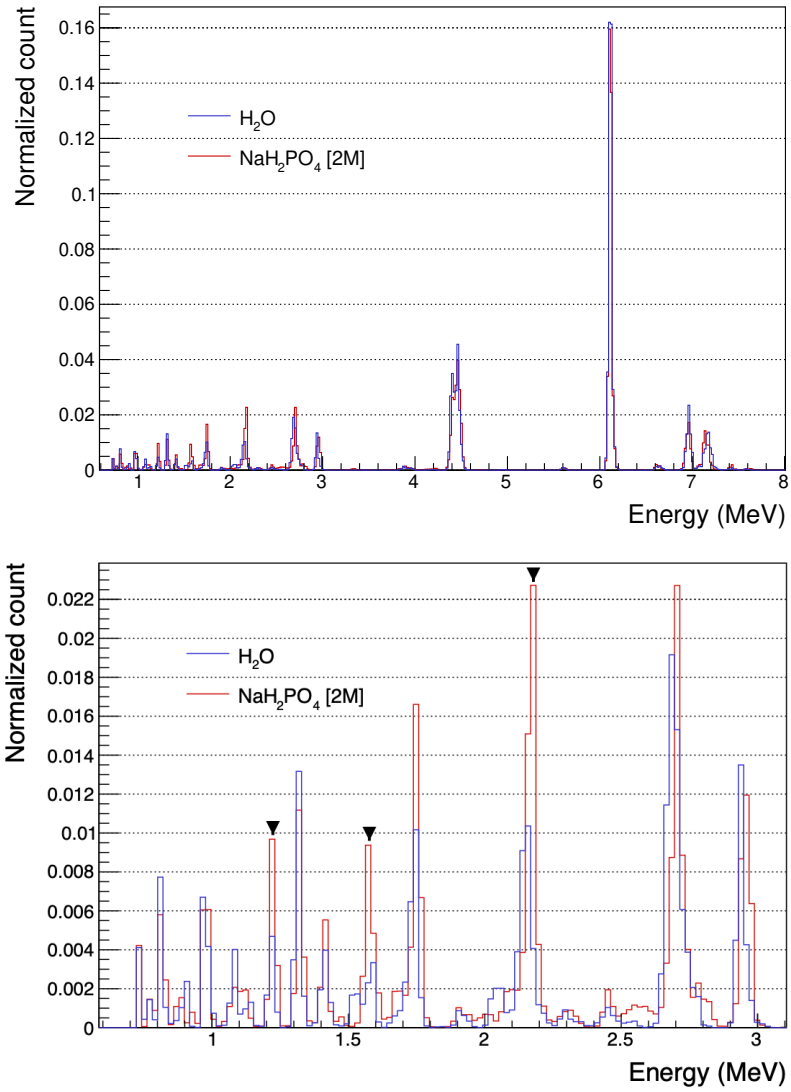

Figure 4.13: Top, comparison between the deconvoluted water spectrum and the one for the phosphorous solution, at a high concentration of 2 M. On the bottom, a zoom on the energy range 0.6 to 3.1 MeV as been applied. The down arrows highlight the gamma lines where a difference between the water and the solution spectra is visible.

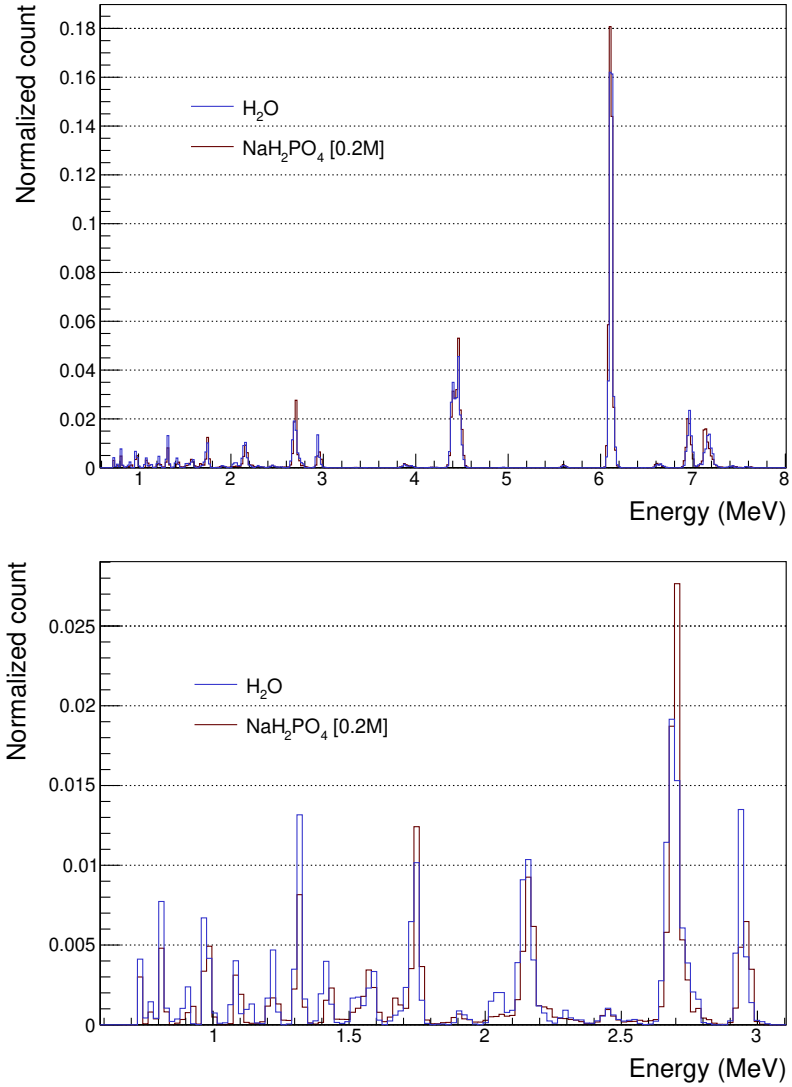

Figure 4.14: Top, comparison between the deconvoluted water spectrum and the one for the phosphorous solution, at a low concentration of 0.2 M. On the bottom, a zoom on the energy range 0.6 to 3.1 MeV as been applied.

Extracted energy (MeV)	NaH_2PO_4[2M] (%)	Water (%)	Proposed reactions	Energy (MeV)
2.1699 ± 0.0002	4.58 ± 0.04	2.97 ± 0.05	^{31}P(p, p'γ)^{31}P	2.23
			^{31}P(p, γ)^{32}S	2.23
			^{31}P(p, nγ)^{31}S	2.23
1.5779 ± 0.0003	1.83 ± 0.02	0.76 ± 0.02	^{23}Na(p, p'γ)^{23}Na	1.64
1.2233 ± 0.0002	1.44 ± 0.02	0.72 ± 0.02	^{31}P(p, p'γ)^{31}P	1.27
			^{31}P(p, nγ)^{31}S	1.25

Table 4.6: Summary of the additional gamma lines found in the NaH_2PO_4 [2M] solution, by comparison with the water spectrum. The extracted energy, the contribution of the line to the total of the spectrum, as well as the corresponding contribution in the water spectrum, when the line exists, are presented. Several possible nuclear reactions at the origin of the lines are also proposed, together with their corresponding energy.

of the respective gamma lines identified in the previous sections is possible, but it already provides a qualitative estimation of the said models.

For this comparison, only water has been simulated, following the description of the experimental setup detailed in section 4.2. Four different models have been used: the intranuclear cascade of Liège (INCL), the binary cascade model (BIC), a modified version of the previous model with a lowered initial energy threshold (BIClt) and the ParticleHP model (PHP).

Because of the form taken by the initial energy distribution of the prompt gamma produced by those models, the unfolding algorithm performances were not optimal. Indeed, they present large continuum range emission over the energy range of interest, which the deconvolution algorithms struggle to reconstruct. Furthermore, those continuum divert the overall statistics of the spectrum from the gamma lines, furthermore degrading the algorithms performances, especially for the lower energy region (down from 3 MeV). Therefore, the initial energy distribution of the prompt gamma are directly used as comparison point.

Figure 4.15 shows on the top the measured spectra from the experiment and the various models investigated. On the bottom, a comparison between the deconvoluted water spectrum and the initial energy distributions is presented. The spectra from the INCL, BIC and BIClt models are quite close to the one from the experiment. However, the differences with the PHP model is evident: none of the dominant gamma lines found in the experimental spectrum are well reproduced by the model; at best, very broad Gaussian distribution can be found. It has therefore been removed

Line (MeV)	INCL (MeV)	BIC (MeV)	BIClt (MeV)
7.12	7.1701 ± 0.0006	7.1769 ± 0.0006	7.1775 ± 0.0006
6.92	6.9672 ± 0.0004	6.9753 ± 0.0004	6.9754 ± 0.0004
6.13	6.1659 ± 0.0005	6.1703 ± 0.0005	6.1703 ± 0.0005
4.44	4.4855 ± 0.0007	4.4854 ± 0.0004	4.4849 ± 0.0005
2.98	3.011 ± 0.002	3.006 ± 0.003	3.009 ± 0.002
2.74			
2.21	2.2198 ± 0.0005	2.2243 ± 0.0005	2.2246 ± 0.0003
1.78			
1.34	1.370 ± 0.002	1.373 ± 0.001	1.375 ± 0.001
1.02	1.0229 ± 0.0003	1.0233 ± 0.0003	1.0230 ± 0.0003
0.84	0.8424 ± 0.0004	0.8412 ± 0.0004	0.8430 ± 0.0005

Table 4.7: Extracted energy for the models INCL, BIC and BIClt, for each of the gamma lines identified in the water spectrum acquired during the experiment.

from the more advanced evaluations presented below.

For all of the gamma lines identified in table 4.3, the same process has been applied to the simulated spectra, in order to retrieve the mean energy of the line and its contribution to the overall spectrum, when it was present. Table 4.7 reports on the mean value obtained for the simulation models for each line. Except for the 1.78 and 2.74 MeV, they can all be found in the initial energy spectrum of the INCL, BIC and BIClt models. However, their mean energy is overestimated, especially at high energies.

Table 4.8 summarizes the contribution to the overall spectrum from each of those lines, for the three different models. The corresponding contribution in the experimental spectrum is also presented.

With respect to the experimental spectrum, some of those lines are well represented by the models: notably the ones at 0.82, 1.02 and 7.12 MeV. The 6.92 MeV line is slightly overestimated, as well as the 2.98 MeV one. The lines at 1.34 and 2.21 MeV are underestimated. The biggest discrepancies are coming from the two main prompt gamma producing reactions: the contribution of the 6.13 MeV line is the same for all models but is largely underestimated; the two BIC models are underestimating the 4.44 MeV line, but less so than the INCL model. Interestingly, the two BIC models predict a more important 4.44 MeV line than the 6.13 MeV one, which is clearly incorrect when looking at the experimental spectra.

In this section, the application of deconvolution algorithms to acquired prompt gamma spectra has been shown to have several benefits. Indeed,

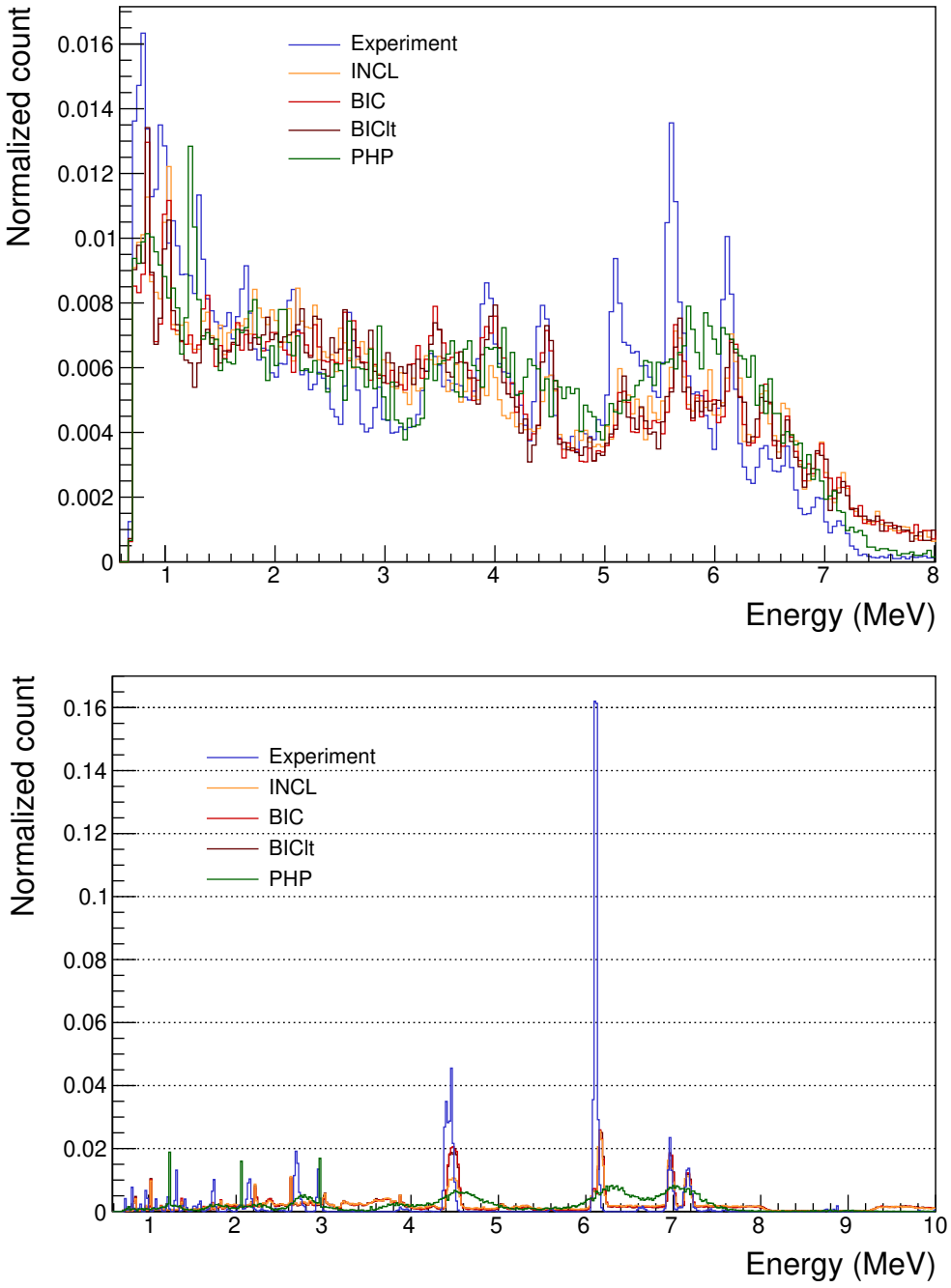

Figure 4.15: Top, comparison between the raw water spectrum from the experiment (in blue). The spectrum obtained by Monte-Carlo simulations are also shown: INCL (in yellow), BIC (in red), BIClt (in brown) and PHP (in green). Bottom, the unfolded spectrum for the experiment, and the corresponding initial spectra for the various models, with the same color scheme.

Line (MeV)	Experiment (%)	INCL (%)	BIC (%)	BIClt (%)
7.12	5.76 ± 0.07	5.59 ± 0.07	5.69 ± 0.08	5.53 ± 0.08
6.92	6.34 ± 0.07	7.29 ± 0.09	7.6 ± 0.1	7.63 ± 0.09
6.13	39.2 ± 0.1	10.3 ± 0.1	10.1 ± 0.1	10.2 ± 0.1
4.44	18.3 ± 0.1	8.24 ± 0.09	14.4 ± 0.1	13.7 ± 0.1
2.98	2.28 ± 0.04	3.1 ± 0.1	3.4 ± 0.2	2.59 ± 0.08
2.21	2.97 ± 0.06	1.94 ± 0.07	1.47 ± 0.04	1.57 ± 0.04
1.34	1.82 ± 0.04	0.95 ± 0.04	0.76 ± 0.03	0.63 ± 0.03
1.02	1.18 ± 0.05	1.26 ± 0.04	1.18 ± 0.04	1.26 ± 0.04
0.84	0.89 ± 0.03	0.74 ± 0.03	0.79 ± 0.03	0.80 ± 0.03

Table 4.8: Extracted contribution to the total of the spectrum for the INCL, BIC and BIClt models, for each gamma line identified in the water spectrum acquired during the experiment. The respective contribution of the lines for the water spectrum are also indicated.

it put into lights slight differences between the spectrum of water and the one of the solutions, highlighting the contribution of the salts to the overall spectrum. It was also useful to evaluate the relevance of various hadronic models used in the GEANT4 toolkit; and indeed, none of the models studied here were found to perform particularly well when looking at prompt gamma production: the energy of the lines are not well reproduced at high energies, and the contribution of each line to the total spectrum is wrong overall.

In the following section, a proof-of-concept of a novel range verification method, using the deconvolution algorithms presented in this chapter is exposed.

4.6 Unfolding as a range verification medium: proof-of-concept

As seen in the earlier section 4.1, the main modality of range verification using prompt gamma emission is based on the spatial distribution of the emitted photons. This is indeed the rationale behind the novel approach described in section 4.2: by enhancing the tumoral volume, a higher statistical count could be achieved in this region, and therefore the end range could be estimated with more precision.

However, other modalities have also been proposed; notably, the use of the energy of the incident photons as a mean to retrieve the corresponding range of the beam (PGS): the spatial distribution of some gamma lines are

more closely related to the incident ion depth than others.

In this section, we propose a proof-of-concept for an in between modality, combining the tumor enhancement and the study of the corresponding gamma line through the application of a deconvolution algorithm to the acquired spectrum. Indeed, the elements used to boost the tumor present a so-called signature in the spectrum: the set of gamma lines that could potentially be emitted by nuclear reaction of the beam in the medium.

The main idea behind this proof-of-concept is to show that by using a knife-edge collimator, for example, the contribution of the gamma lines, corresponding to the elements introduced in the tumor, will vary with the detector position along the beam axis. In this way, additional information on the tumor position can be obtained.

Simulated setup

In order to prove the feasibility of this method, a basic setup has been simulated, and is presented in the figure 4.16. The target is a 5×5×15 cm box of water. A knife-edge collimator, as described in [126], is placed along the body of water, at at distance of 15 cm. The same $CeBr_3$ crystal used in the central experiment of this chapter, is positioned 10 cm away from this collimator, and is rotated along the x axis, covering the solid angle allowed by the collimator.

For this proof of concept, only pure gamma lines were simulated as incident particles, instead of a proton beam. The main lines from the water spectrum, corresponding to nuclear reactions on ^{16}O, and summarized in section 4.5, were used as input, with their respective contribution to the total spectrum.

The effect of the irradiation on the tumoral volume was represented as an additional gamma line, at an energy of 2.23 MeV (taken from the phosphorous solution). Two configurations were considered: the volume of emission of this gamma line was a 5 × 5 × 1 cm box, centered either at a depth of 10 or 12 cm in the water target. This line was assigned an intensity of 2%, meaning that 2% of the incident photons had the corresponding energy.

Finally, $\times 10^7$ incident photons were simulated, in a $(2-\sqrt{2})\pi$ solid angle, in the direction of the collimator, i.e. with an angle θ varying between 0 to 45° from the y axis.

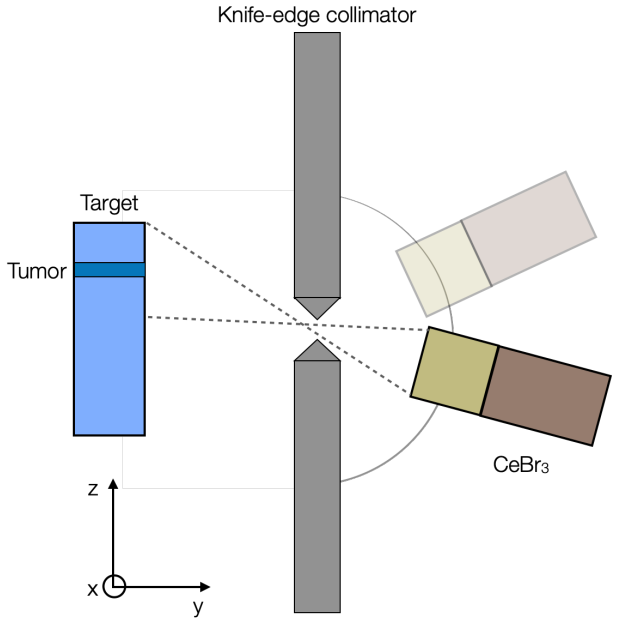

Figure 4.16: Schematic top view of the simulated setup for the proof-of-concept of a prompt gamma modality exploiting spectroscopic analysis of tumors enhanced for prompt gamma emission.

Result

Preliminary results for this proof-of-concept can be found in figure 4.17: the integral (normalized to the number of count in the spectrum) of the additional gamma line, after unfolding of the spectrum, is shown as a function of the angle of the $CeBr_3$ crystal, for the two "tumoral" positions.

Figure 4.17 clearly shows that the integral of the gamma line is correlated to the detector angle, and therefore to the position of the tumoral volume. Furthermore, the total number of counts in any of the spectrum considered does not exceed 12000, suggesting that this could be use even if low statistics are present.

This method, as presented here, is just a proof-of-concept. In order to proceed further, more advanced and realistics simulations would be necessary, as well as an experimental study of the feasability of such a technique. A careful selection of the detector as well as a proper optimization of the geometry would be required. Nonetheless, the utility of using unfolding as mean to extract additional information on beam range from prompt gamma spectra, in the context of an enhanced tumoral volume, was demonstrated.

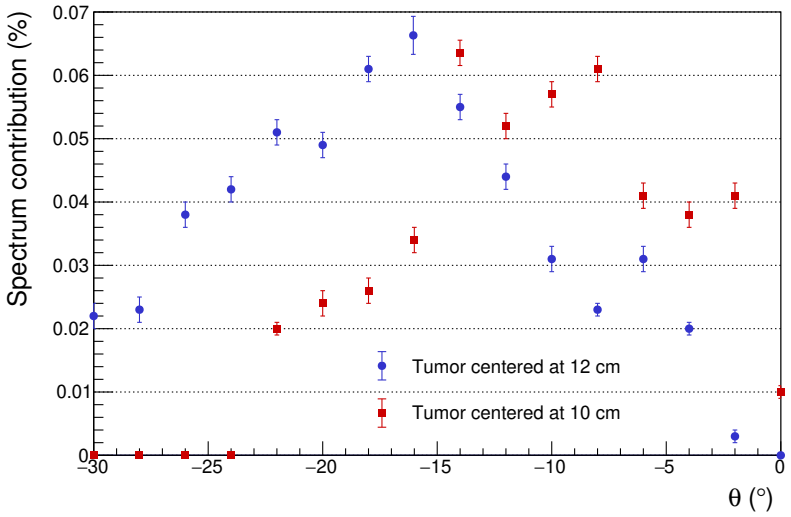

Figure 4.17: Contribution of the gamma line corresponding to the enhanced tumor (2.23 MeV) to the total of the spectrum, after unfolding, as a function of the angle of the crystal. The blue dots correspond to an emission volume centered at a depth of 12 cm, and the red squares at a depth of 10 cm.

4.7 Conclusion

A rapid overview of the prompt gamma field, as a mean to monitor the range of the beam during treatment, allowed the display of the difficulties it must face. The extensive use of Monte-Carlo simulations in the domain requires a proper evaluation of the models used during those simulations. Furthermore, the field is limited by the inherent low available statistics of the detected prompt gammas needed in order to perform its role.

A recent experiment, joining two research groups from Trento and Strasbourg, aimed to demonstrate a novel approach to range monitoring using prompt gammas: by enhancing the tumor with specific stable elements, the prompt gamma statistics would be artificially boosted in the corresponding volume. This experiment was also the perfect opportunity to evaluate the performances of the hadronic models, used in Monte-Carlo simulations, and involved in the production of prompt gammas.

Unfolding algorithms were applied to the several spectra acquired during the experiment, in order to extract the contribution of specific gamma line coming from the selected elements, and their contribution to the overall spectrum. For the several candidate studied, a net impact on the spectrum was observed, with respect to the spectrum of pure water.

A comparison between simulated spectrum, reproducing the experimental setup, and the acquired water spectrum, for four different hadronic model (INCL, BIC, BIClt, and PHP) showed that no model reproduced accurately the measured experimental spectrum. Furthermore, the position of the gamma lines identified in the experimental spectrum demonstrated that even the energy of the lines are inaccurate for those models.

The chapter ended with a proof-of-concept of the application of deconvolution algorithms to prompt gamma spectrum, in combination with tumor enhancement, as an additional information source for range verification during treatment.

Un aperçu rapide du champ de recherche des gammas prompts, en tant que moyen de contrôle du traitement, a permis de montrer les difficultés auxquelles il doit faire face. L'utilisation intensive des simulations de Monte-Carlo dans le domaine nécessite une évaluation correcte des modèles utilisés lors de ces simulations. En outre, le champ est limité par la faible quantité de gammas prompts détectables lors d'un traitement.

Une expérience récente, réunissant deux groupes de recherche de Trento et Strasbourg, visait à démontrer une nouvelle approche quant à l'utilisation de gammas prompts pour effectuer un contrôle du traitement : en dopant la tumeur avec certains éléments stables, le taux de gammas prompt produits pourrait y être artificiellement augmenté. Cette expérience a également été l'occasion d'évaluer les performances des modèles hadroniques, utilisés dans les simulations de Monte-Carlo, et impliqués dans la production de gammas prompts.

Des algorithmes de déconvolution ont été appliqués aux différents spectres acquis pendant l'expérience, afin d'extraire la contribution des raies gamma spécifiques provenant des éléments sélectionnés au spectre global. Pour les différents candidats étudiés, un impact net sur le spectre a été observé, par rapport au spectre de l'eau pure.

Une comparaison entre le spectre simulé, reproduisant la configuration expérimentale, et le spectre de l'eau acquis, pour quatre modèles hadroniques différents (INCL, BIC, BIClt, et PHP) a montré qu'aucun modèle ne reproduit avec précision le spectre expérimental mesuré. L'énergie des raies identifiées dans le spectre expérimental est également mal reproduite par ces modèles.

Le chapitre s'est terminé par une démonstration de l'utilité potentielle

de l'application d'algorithmes de déconvolution aux spectres de gammas prompt, en combinaison avec le dopage tumoral, pour le contrôle du traitement en hadronthérapie.

Conclusion and outlook

In this manuscript, the work carried out during this thesis was exposed in details: several algorithms (and associated software tools) have been developed in order to extract nuclear data of interest in the context of experiments dedicated to charged particle therapy.

The first chapter delineated the general context in which this thesis aligns itself. After a first overview of the health burden cancer imposes upon the world, the physical interactions encountered by either photons or charged particles, when crossing a medium, were exposed. From there, the resulting impact of ionizing radiations on living tissues, and particularly, the cells they are composed of, were described.

The benefits related to the use of charged particle therapy over conventional radiotherapy emerged upon this comparison: the localized energy deposition at fixed depth from incident ions, coupled with an increased lethal efficiency upon cells, allows for potential improvements in treatment, either through an increase of tumor control probability (TCP) for a same normal tissue complication probability (NTCP) or a decreased NTCP while conserving the same TCP.

However, on a physical point of view, charged particle therapy must handle the potential fragmentation of the incoming ion, or the target nuclei, a problem which finds no equivalence in classical radiotherapy. Indeed, because of this process, a mixed-field of particles is always produced during any treatment involving ions as the beam type. This mixed-field, because of its impact on the energy deposition, needs to be taken into account while planning the treatment through nuclear models.

Another downside of the use of charged particle therapy lies in the difficulty to monitor treatment, because the particles from the beam do not exit the patient. To this day, several techniques have been developed, but they are limited by the inherent low statistics available, and also rely on the use of models for nuclear reaction simulation.

Thus, it appears that the models used to reproduce the effects of nuclear reactions on treatment, in the charged particle therapy context, need

further evaluations of their accuracy in order for the field to continue to improve. This evaluation needs to be based on experimental cross-section measurements, which are currently incomplete in the therapeutic energy range.

The FOOT experiment was proposed to extend the set of isotopic double differential cross-sections data available for charged particle therapy. The experiment aims to cover a large angular acceptance as well as the largest variety of fragments possible. The description of the electronic setup was the object of the second chapter. It is composed of three regions: the upstream region, which main purpose is to monitor the beam; a magnetic tracking region, aiming to provide tracking information of the fragments; and finally a downstream region, to supply charge identification as well as time-of-flight and energy measurements.

In order to achieve isotopic double differential cross-sections measurements, a robust fragment identification needs to be used. The envisioned method makes use of three different mass number evaluations. A few fundamental quantities needed to compute the mass numbers can be directly extracted from the local reconstruction of the detectors. However, quantities such as momentum or path length need the use of a global reconstruction of the event in order to be extracted.

The tracking of ejectiles algorithm (TOE), exposed in full details in the chapter three, is a global reconstruction method aiming for track recognition and reconstruction, developed for the FOOT experiment. The algorithm is divided into two parts: a first procedure allows for the recognition of the tracks, while the second establishes the parameters of interest, i.e. the momentum and length of the reconstructed track.

The recognition procedure is based around three components: a propagator, a filtering and selection procedure, and finally, an arborescence or history of the reconstruction. In order to start the reconstruction, several hypothesis, divided between a static and dynamic part, need to be formed. Developed to be computationally efficient, the algorithm does not take into account multiple scattering or energy loss during propagation; this role is achieved by a set of geometrical cuts computed separately by a dedicated tool.

The performances of the recognition procedure were evaluated with respect to two main criteria, with a charge reconstruction taken from Monte-Carlo truth: the efficiency of the reconstruction and the cluster purity. For the four data sets studied, the overall recognition efficiency of the algorithm was found to be around 99%, for a cluster purity above 94%. The cluster purity is impacted mainly by light fragments, with $Z = 1$ or $Z = 2$:

their broad angle and energy distributions, and the fact that they are more prone to multiple Coulomb scattering makes their reconstruction difficult. The fake yield and clone multiplicity were also evaluated: they represent the number of false reconstruction and the average number of concurrent history needed to reconstruct a track. The performances of the algorithm, for both of those quantities, are also diminished for the reconstruction of lighter fragments, for the same reasons.

The second part of the algorithm, the track fitting procedure, makes use of a two step procedure to extract the parameters of interest after reconstruction: first, a basic polynomial model is fitted to the selected measurements for any given track; from there, a scanning algorithm is used to refine the extracted values.

The quality of the track parameters evaluation was assessed through the determination of the momentum resolution achievable by the algorithm. In order to achieve its goals, the FOOT experiment need to achieve a momentum resolution of the reconstructed track lower than 5%. For the heavier incident particles ($Z > 2$), this objective is achieved by the TOE algorithm. For lighter fragments, the performances are more nuanced: the resolution on the momentum ranges between 5% to 40%.

As it stands now, the TOE algorithm is ready to be used on any input data, but has not yet been tested on experimental input acquired with a complete setup. While it was developed to be robust even in experimental conditions, its status can only be confirmed then. Furthermore, some possible modifications have not been implemented yet: for example, the algorithm could also be used to correct any erroneous charge identification coming from the local reconstruction, through the inclusion of more hypotheses in the dedicated pool, but at the cost of computational time.

The final chapter of this thesis was based on a feasibility study of a novel technique looking at monitoring through the use of prompt gammas: in order to remedy the low statistics available during treatment, it was proposed to enhanced the tumoral volume with specific elements in order to boost the production of prompt gammas.

The chapter had several objectives: through the use of unfolding algorithms, it was possible to extract relative contribution of specific gamma lines from the data acquired during the experiment; those algorithms were also used to evaluate the accuracy of Monte-Carlo hadronic models from the GEANT4 toolkit for prompt-gamma production; finally, a proof-of-concept of a complementary range verification technique using an unfolding algorithm coupled to an enhanced tumor was proposed.

Thus, the most important gamma lines composing the acquired water

spectrum, together with the lines coming from contamination by the aluminum collimator used in the experiment, were identified after the application of the boosted Richardson-Lucy unfolding algorithm. The contribution of each line to the total spectrum was also retrieved. The comparison of this control spectrum with the one from several solutions (with respectively $CuSO_4$, $Y(NO_3)_3$ and NaH_2PO_4 salts) allowed for the identification of the nuclear reactions involved in the production of the prompt gammas for the corresponding elements.

A comparison between simulated spectrum, reproducing the experimental setup, and the acquired water spectrum, for four different hadronic models (INCL, BIC, BIClt, and PHP) showed that no model reproduced accurately the measured experimental spectrum. For example, the main gamma line found in the experimental water spectrum (6.13 MeV) is largely underestimated by all models. Furthermore, the position of the gamma lines identified in the experimental spectrum demonstrated that even the energy of the lines are inaccurate for those models.

The chapter ended with a basic proof-of-concept of the application of deconvolution algorithms to prompt gamma spectrum, in combination with tumor enhancement, as an additional information source for range verification during treatment. Using a knife-edge collimator, it was shown that the integral of the gamma line corresponding to the selected element was correlated to the angle of the detection setup. This proof-of-concept, needs refinements in order to be taken further: a more accurate simulation with a proton beam, and modified evaporation models would demonstrate the feasibility of this method with more weight. A basic experimental study could also be envisioned, with a minimalist setup.

In conclusion, the charged particle therapy field would benefit greatly from the extension of the available isotopic double differential cross-sections measurements at the corresponding energy range. Indeed, they would allow for proper constrains to be put on Monte-Carlo simulation models, which is of particular relevance because of the extensive use of those models in the field. The FOOT experiment, and with it the TOE algorithm, are aiming to solve this issue in the near future. At the same time, other aspects of the field needs to be improved as well, in order to fully exploit the potential of hadrontherapy: notably, the efforts towards treatment monitoring need to be pursued.

Dans ce manuscrit, le travail réalisé au cours de cette thèse a été exposé en détails : plusieurs algorithmes (et outils logiciels associés) ont été développés afin d'extraire des données nucléaires d'intérêt dans le cadre d'expériences dédiées à la thérapie par particules chargées.

Le premier chapitre a délimité le contexte général dans lequel s'inscrit cette thèse. Après un premier aperçu du fardeau sanitaire que le cancer impose au monde, les interactions physiques subies par les photons ou les particules chargées, lorsqu'ils traversent un milieu, ont été exposées. Ensuite, l'impact des radiations ionisantes sur les tissus vivants, et en particulier sur les cellules qui les composent, a été décrit.

Les avantages liés à l'utilisation de la thérapie par particules chargées par rapport à la radiothérapie conventionnelle sont apparus lors de cette comparaison : leur dépôt d'énergie localisé à une profondeur déterminée, associé à une efficacité létale accrue sur les cellules, permet des améliorations potentielles du traitement, soit par une augmentation de la probabilité de contrôle de la tumeur (TCP) pour une même probabilité de complication des tissus sains (NTCP), soit par une diminution de la NTCP tout en conservant la même TCP.

Cependant, d'un point de vue physique, la thérapie par particules chargées doit faire face à la fragmentation potentielle de l'ion entrant, ou des noyaux cibles, un problème qui ne trouve pas d'équivalent en radiothérapie classique. En effet, en raison de ce processus, un champ mixte de particules est toujours produit lors de tout traitement impliquant des ions comme type de faisceau. Ce champ mixte vient modifier le dépôt d'énergie initialement planifié, et doit donc être pris en compte lors de la planification du traitement, au moyen notamment de modèles nucléaires.

Un autre inconvénient de l'utilisation de la thérapie par particules chargées réside dans la difficulté de contrôle du traitement, car les particules du faisceau ne sortent pas du patient. A ce jour, plusieurs techniques ont été développées, mais elles sont limitées par la faible quantité de particule secondaires produites lors du traitement, et reposent également sur l'utilisation de modèles de simulation des réactions nucléaires.

Ainsi, il semble que la précision des modèles utilisés pour reproduire les effets des réactions nucléaires sur le traitement, dans le contexte de la thérapie par particules chargées, doit être rigoureusement évaluée pour que le domaine puisse continuer à s'améliorer. Cette évaluation doit être basée sur des mesures expérimentales de sections efficaces, qui sont actuellement incomplètes dans la gamme d'énergie thérapeutique.

L'expérience FOOT a été proposée pour compléter les données de sections efficaces doublement différentielles isotopiques disponibles pour la

thérapie par particules chargées. L'expérience vise à couvrir une large ouverture angulaire ainsi que la plus grande variété de fragments possible. La description du dispositif électronique de l'expérience a fait l'objet du deuxième chapitre de ce manuscrit. Il est composé de trois régions : la région amont, dont le but principal est le suivi du faisceau ; une région de trajectographie magnétique, visant à fournir des informations sur les trajectoires des fragments ; et enfin une région aval, permettant de fournir une identification de la charge ainsi que des mesures de temps de vol et d'énergie des fragments produits.

Afin d'obtenir des mesures de sections efficaces isotopiques doublement différentielles, une identification robuste des fragments doit être utilisée. La méthode envisagée fait appel à trois évaluations différentes du nombre de nucléons. Quelques quantités fondamentales nécessaires à ces calculs peuvent être directement extraites de la reconstruction locale des détecteurs. Cependant, des quantités telles que la quantité de mouvement ou la longueur de la trajectoire nécessitent l'utilisation d'une reconstruction globale de l'événement afin d'être extraites.

L'algorithme TOE, exposé en détails dans le chapitre trois, est une méthode de reconstruction globale visant la reconnaissance et la reconstruction des trajectoires des fragments, développée pour l'expérience FOOT. L'algorithme est divisé en deux parties : une première procédure permet la reconnaissance des traces, tandis que la seconde établit les paramètres d'intérêt, c'est-à-dire la quantité de mouvement et la longueur de la trajectoire reconstruite.

La procédure de reconnaissance s'articule autour de trois composantes : un propagateur, une procédure de filtrage et de sélection, et enfin, une arborescence ou historique de la reconstruction. Afin de commencer la reconstruction, plusieurs hypothèses, divisées entre une partie statique et une partie dynamique, doivent être formulées. Développé pour être efficace en termes de temps de calcul, l'algorithme ne tient pas compte de la diffusion multiple ou de la perte d'énergie pendant la propagation ; ce rôle est rempli par un ensemble de coupures géométriques calculées séparément par un outil dédié.

Les performances de la procédure de reconnaissance ont été évaluées en fonction de deux critères principaux, avec une reconstruction de charge issue de la simulation de Monte-Carlo : l'efficacité de la reconstruction et la pureté. Pour les quatre jeux de données étudiés, l'efficacité globale de reconnaissance de l'algorithme s'est avérée être d'environ 99%, pour une pureté supérieure à 94%. La pureté est principalement impactée par les fragments légers, avec $Z = 1$ ou $Z = 2$: leurs larges distributions en angle

et en énergie, et le fait qu'ils soient plus enclins à la diffusion coulombienne multiple rendent leur reconstruction difficile. Le taux de faux et la multiplicité des clones ont également été évalués : ils représentent la proportion de fausses reconstructions et le nombre moyen d'historiques simultanés nécessaires pour reconstruire une trajectoire. Les performances de l'algorithme, pour ces deux quantités, sont plus mauvaises lors de la reconstruction de fragments plus légers, pour des raisons similaires à celles déjà évoquées.

La deuxième partie de l'algorithme, la procédure d'ajustement de la trajectoire, utilise une procédure en deux étapes pour extraire les paramètres d'intérêt après la reconstruction : tout d'abord, un modèle polynomial est ajusté aux amas sélectionnés par la procédure de reconnaissance ; à partir de là, un algorithme de balayage est utilisé pour affiner les valeurs des paramètres d'intérêt.

La qualité de l'estimation de ces paramètres a été évaluée par la détermination de la résolution sur la quantité de mouvement reconstruite par l'algorithme. Afin d'atteindre ses objectifs, l'expérience FOOT doit obtenir une résolution sur l'impulsion reconstruite inférieure à 5%. Pour les particules incidentes les plus lourdes (Z ¿ 2), cet objectif est atteint par l'algorithme TOE. Pour les fragments plus légers, les performances sont plus nuancées : la résolution sur la quantité de mouvement varie entre 5% et 40%.

Dans l'état actuel des choses, l'algorithme TOE est prêt à être utilisé sur tout type de données d'entrée, mais n'a pas encore été testé sur des données expérimentales acquises avec une configuration complète. Bien qu'il ait été développé pour être robuste même dans des conditions expérimentales, son statut ne pourra être confirmé qu'alors. En outre, certaines modifications possibles n'ont pas encore été mises en place : par exemple, l'algorithme pourrait également être utilisé pour corriger toute identification erronée de la charge provenant de la reconstruction locale, par l'inclusion de plus d'hypothèses dans l'ensemble de départ mais ce au prix du temps de calcul.

Le dernier chapitre de cette thèse était basé sur une étude de faisabilité d'une nouvelle technique de contrôle de traitement par détection de gammas prompts : afin de remédier aux faibles statistiques disponibles pendant le traitement, il a été proposé d'enrichir le volume tumoral avec des éléments spécifiques afin de stimuler la production de gammas prompts.

Ce chapitre avait plusieurs objectifs : grâce à l'utilisation d'algorithmes de déconvolution, il a été possible d'extraire la contribution relative de raies gamma spécifiques à partir des données acquises pendant l'expérience ; ces algorithmes ont également été utilisés pour évaluer la précision des modèles

hadroniques Monte-Carlo du progiciel GEANT4 pour la production de gammas prompts ; enfin, une preuve de concept d'une technique complémentaire de contrôle de traitement utilisant un algorithme de déconvolution couplé à une tumeur dopée a été proposée.

Ainsi, les raies gamma les plus importantes composant le spectre de l'eau acquis lors de l'expérience, ainsi que les raies provenant de la contamination par le collimateur en aluminium utilisé, ont été identifiées après l'application de l'algorithme de déconvolution de Richardson-Lucy. La contribution de chaque raie au spectre total a également été extraite. La comparaison de ce spectre de contrôle avec celui de plusieurs solutions (avec respectivement des sels de $CuSO_4$, $Y(NO_3)_3$ et NaH_2PO_4) a permis d'identifier les réactions nucléaires impliquées dans la production des gammas prompts pour les éléments correspondants.

Une comparaison entre le spectre simulé, reproduisant le montage expérimental, et le spectre de l'eau acquis lors de l'experience, pour quatre modèles hadroniques différents (INCL, BIC, BIClt, et PHP) a montré qu'aucun modèle ne reproduit fidèlement les données expérimentales. Par exemple, la principale raie gamma trouvée dans le spectre expérimental de l'eau (6,13 MeV) est largement sous-estimée par tous les modèles. L'énergie de ces raies est également imprécise pour tous les modèles considérés.

La dernière section du chapitre a permis la démonstration du principe de base d'une technique complémentaire pour le contrôle de traitement, lors de l'utilisation d'une tumeur dopée. En utilisant un collimateur knife-edge, il a été montré que l'intégrale de la raie gamma correspondant à l'élément inséré dans la tumeur était corrélée à l'angle du dispositif de détection. Cette validation de principe doit être affinée afin d'aller plus loin : une simulation plus précise avec un faisceau de protons et des modèles d'évaporation modifiés permettraient de démontrer la faisabilité de cette méthode avec plus de poids. Une étude expérimentale de base pourrait également être envisagée, avec une installation minimaliste.

En conclusion, le domaine de la thérapie par particules chargées bénéficierait grandement de l'extension des mesures de sections efficaces isotopique doublement différentielles disponibles dans la gamme d'énergie correspondante. En effet, ces mesures permettraient d'imposer des contraintes appropriées aux modèles de simulation de Monte-Carlo, ce qui est particulièrement important en raison de l'utilisation intensive de ces modèles dans ce domaine. L'expérience FOOT, et avec elle l'algorithme TOE, visent à résoudre ce problème dans un avenir proche. Parallèlement, d'autres aspects du domaine doivent également être améliorés afin d'exploiter pleinement le potentiel de la hadronthérapie : notamment, les efforts en matière

de contrôle du traitement doivent être poursuivis.

APPENDIX A

Butcher arrays for GRKN-5(6)

i	0	1	2	3	4	5	6	7	8
$\dot{w}_i$	$\frac{7}{81}$	0	0	0	$-\frac{250}{3483}$	$\frac{160}{351}$	$\frac{2401}{5590}$	$\frac{1}{10}$	0
w_i	$\frac{31}{360}$	0	0	0	0	$\frac{64}{588}$	$\frac{2401}{7800}$	$-\frac{1}{300}$	0
$\hat{w}_i$	$\frac{31}{360}$	0	0	0	0	$\frac{64}{588}$	$\frac{2401}{7800}$	0	$-\frac{1}{300}$

Table A.1: Values for the weights used in the computation of the solutions of the differential equations, taken from [112].

	α_i	β_{ij}							
i \ j		0	1	2	3	4	5	6	7
0	0								
1	$\frac{4}{15}$	$\frac{4}{15}$							
2	$\frac{2}{5}$	$\frac{1}{10}$	$\frac{3}{10}$						
3	$\frac{3}{5}$	$\frac{3}{20}$	0	$\frac{9}{20}$					
4	$\frac{9}{10}$	$\frac{9}{40}$	0	0	$\frac{27}{40}$				
5	$\frac{3}{4}$	$\frac{11}{48}$	0	0	$\frac{5}{8}$	$-\frac{5}{48}$			
6	$\frac{2}{7}$	$\frac{27112}{194481}$	0	0	$\frac{56450}{64827}$	$\frac{80000}{194181}$	$-\frac{24544}{21609}$		
7	1	$-\frac{26033}{41796}$	0	0	$-\frac{236575}{38313}$	$-\frac{14500}{10449}$	$\frac{275936}{45279}$	$\frac{228095}{73788}$	
8	1	$\frac{7}{81}$	0	0	0	$-\frac{250}{3483}$	$\frac{160}{351}$	$\frac{2401}{5590}$	$\frac{1}{10}$

Table A.2: Butcher array for the β factors, taken from [112].

	α_i	γ_{ij}							
i \ j		0	1	2	3	4	5	6	7
0	0								
1	$\frac{4}{15}$	$\frac{4}{15}$							
2	$\frac{2}{5}$	$\frac{1}{10}$	$\frac{3}{10}$						
3	$\frac{3}{5}$	$\frac{3}{20}$	0	$\frac{9}{20}$					
4	$\frac{9}{10}$	$\frac{9}{40}$	0	0	$\frac{27}{40}$				
5	$\frac{3}{4}$	$\frac{11}{48}$	0	0	$\frac{5}{8}$	$-\frac{5}{48}$			
6	$\frac{2}{7}$	$\frac{27112}{194481}$	0	0	$\frac{56450}{64827}$	$\frac{80000}{194181}$	$-\frac{24544}{21609}$		
7	1	$-\frac{26033}{41796}$	0	0	$-\frac{236575}{38313}$	$-\frac{14500}{10449}$	$\frac{275936}{45279}$	$\frac{228095}{73788}$	
8	1	$\frac{7}{81}$	0	0	0	$-\frac{250}{3483}$	$\frac{160}{351}$	$\frac{2401}{5590}$	$\frac{1}{10}$

Table A.3: Butcher array for the γ factors, taken from [112].

Bibliography

[1] H. Sung, J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, and F. Bray, "Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries," *CA A Cancer J Clin*, vol. 71, pp. 209–249, May 2021.

[2] R. R. Wilson, "Radiological Use of Fast Protons," *Radiology*, vol. 47, pp. 487–491, Nov. 1946. Publisher: Radiological Society of North America.

[3] A. C. Society, "Global Cancer Facts & Figures 4th Edition," 2018.

[4] W. . IARC, 2020.

[5] A. C. Society, "Cancer Treatment & Survivorship Facts & Figures 2019-2021," 2019.

[6] PTCOG, Particle therapy co-operative group.

[7] M. Durante and H. Paganetti, "Nuclear physics in particle therapy: a review," *Rep. Prog. Phys.*, vol. 79, p. 096702, Sept. 2016.

[8] H. Ishikawa, Tadashi Kamada, Koichiro Akakura, Hiroyoshi Suzuki, Jun Shimazaki, Hirohiko Tsuji, and H. Tsuji, "Carbon-ion radiation therapy for prostate cancer," *I. Jr. Of Uro.*, vol. 19, no. 4, pp. 296–305, 2012.

[9] S. C. Darby, M. Ewertz, P. McGale, A. M. Bennet, U. Blom-Goldman, D. Brønnum, C. Correa, D. Cutter, G. Gagliardi, B. Gigante, M.-B. Jensen, A. Nisbet, R. Peto, K. Rahimi, C. Taylor, and P. Hall, "Risk of Ischemic Heart Disease in Women after Radiotherapy for Breast Cancer," *N Engl J Med*, vol. 368, pp. 987–998, Mar. 2013.

[10] T. Kamada, H. Tsujii, E. A. Blakely, J. Debus, W. De Neve, M. Durante, O. Jäkel, R. Mayer, R. Orecchia, R. Pötter, S. Vatnitsky, and

W. T. Chu, "Carbon ion radiotherapy in Japan: an assessment of 20 years of clinical experience," *The Lancet Oncology*, vol. 16, pp. e93–e100, Feb. 2015.

[11] M. Riboldi, R. Orecchia, and G. Baroni, "Real-time tumour tracking in particle therapy: technological developments and future perspectives," *The Lancet Oncology*, vol. 13, pp. e383–e391, Sept. 2012.

[12] D. Hanahan and R. A. Weinberg, "The Hallmarks of Cancer," *Cell*, vol. 100, pp. 57–70, Jan. 2000.

[13] D. Hanahan and R. Weinberg, "Hallmarks of Cancer: The Next Generation," *Cell*, vol. 144, pp. 646–674, Mar. 2011.

[14] N. Pavlova and C. Thompson, "The Emerging Hallmarks of Cancer Metabolism," *Cell Metabolism*, vol. 23, pp. 27–47, Jan. 2016.

[15] R. D. Evans, *The atomic nucleus*. Tata McGraw-Hill Publishing Company, 1955.

[16] H. Bethe, "Zur Theorie des Durchgangs schneller Korpuskularstrahlen durch Materie," *Ann. Phys.*, vol. 397, no. 3, pp. 325–400, 1930.

[17] W. Ulmer, "Theoretical aspects of energy–range relations, stopping power and energy straggling of protons," *Radiation Physics and Chemistry*, vol. 76, pp. 1089–1107, July 2007.

[18] D. Schardt, T. Elsässer, and D. Schulz-Ertner, "Heavy-ion tumor therapy: Physical and radiobiological benefits," *Rev. Mod. Phys.*, vol. 82, pp. 383–425, Feb. 2010.

[19] M. Krämer, E. Scifoni, C. Schuy, M. Rovituso, W. Tinganelli, A. Maier, R. Kaderka, W. Kraft-Weyrather, S. Brons, T. Tessonnier, K. Parodi, and M. Durante, "Helium ions for radiotherapy? Physical and biological verifications of a novel treatment modality: Helium ions for radiotherapy?," *Med. Phys.*, vol. 43, pp. 1995–2004, Mar. 2016.

[20] Marie Vanstalle, *Amélioration des plans de traitement en hadronthérapie*. HDR, University of Strasbourg, 2019.

[21] K. Kosaki, S. Ecker, D. Habermehl, S. Rieken, O. Jäkel, K. Herfarth, J. Debus, and S. E. Combs, "Comparison of intensity modulated radiotherapy (IMRT) with intensity modulated particle therapy

(IMPT) using fixed beams or an ion gantry for the treatment of patients with skull base meningiomas," *Radiat Oncol*, vol. 7, p. 44, Dec. 2012.

[22] ICRU, "Quantities and Units in Radiation Protection Dosimetry," 1993.

[23] ICRU, "Linear Energy Transfer," 1970.

[24] F. Tommasino, E. Scifoni, and M. Durante, "New Ions for Therapy," *International Journal of Particle Therapy*, vol. 2, pp. 428–438, Dec. 2015.

[25] G. R. Lynch and O. I. Dahl, "Approximations to multiple Coulomb scattering," *Nuclear Instruments and Methods in Physics Research B*, pp. 6–10, 1991.

[26] J. Hüfner, K. Schäfer, and B. Schürmann, "Abrasion-ablation in reactions between relativistic heavy ions," *Phys. Rev. C*, vol. 12, pp. 1888–1898, Dec. 1975.

[27] Wilson J. W., Tripathi R. K., Cucinotta F. A., Shinn J. L., Badavi F. F., Chun S. Y., Norbury J. W., Zeitlin C. J., Heilbronn L., and Miller J., "NUCFRG2: An evaluation of the semiempirical nuclear fragmentation database," Technical paper 3533, NASA, 1995.

[28] K. Gunzert-Marx, H. Iwase, D. Schardt, and R. S. Simon, "Secondary beam fragments produced by 200 MeV/u ^{12} C ions in water and their dose contributions in carbon ion radiotherapy," *New J. Phys.*, vol. 10, p. 075003, July 2008.

[29] J. Cugnon and R. Sartor, "Factorization in high-energy nucleus-nucleus fragmentation cross sections," *Phys. Rev. C*, vol. 21, pp. 2342–2344, June 1980.

[30] D. L. Olson, B. L. Berman, D. E. Greiner, H. H. Heckman, P. J. Lindstrom, and H. J. Crawford, "Factorization of fragment-production cross sections in relativistic heavy-ion collisions," *Phys. Rev. C*, vol. 28, pp. 1602–1613, Oct. 1983.

[31] C. La Tessa, L. Sihver, C. Zeitlin, J. Miller, S. Guetersloh, L. Heilbronn, D. Mancusi, Y. Iwata, and T. Murakami, "Test of weak and strong factorization in nucleus–nucleus collisions at several hundred MeV/nucleon," *Nuclear Physics A*, vol. 791, pp. 434–450, July 2007.

[32] E. Haettner, H. Iwase, M. Krämer, G. Kraft, and D. Schardt, "Experimental study of nuclear fragmentation of 200 and 400 MeV/ u ^{12} C ions in water for applications in particle therapy," *Phys. Med. Biol.*, vol. 58, pp. 8265–8279, Dec. 2013.

[33] J. W. Norbury, G. Battistoni, J. Besuglow, L. Bocchini, D. Boscolo, A. Botvina, M. Clowdsley, W. de Wet, M. Durante, M. Giraudo, T. Haberer, L. Heilbronn, F. Horst, M. Krämer, C. La Tessa, F. Luoni, A. Mairani, S. Muraro, R. B. Norman, V. Patera, G. Santin, C. Schuy, L. Sihver, T. C. Slaba, N. Sobolevsky, A. Topi, U. Weber, C. M. Werneth, and C. Zeitlin, "Are Further Cross Section Measurements Necessary for Space Radiation Protection or Ion Therapy Applications? Helium Projectiles," *Front. Phys.*, vol. 8, p. 565954, Nov. 2020.

[34] J. M. Kidd, P. J. Lindstrom, H. J. Crawford, and G. Woods, "Fragmentation of carbon ions at 250 MeV/nucleon," *Phys. Rev. C*, vol. 37, pp. 2613–2623, June 1988.

[35] H. D. MacCabee and M. A. Ritter, "Fragmentation of High-Energy Oxygen-Ion Beams in Water," *Radiation Research*, vol. 60, p. 409, Dec. 1974.

[36] W. R. Webber, J. C. Kish, and D. A. Schrier, "Total charge and mass changing cross sections of relativistic nuclei in hydrogen, helium, and carbon targets," *Phys. Rev. C*, vol. 41, pp. 520–532, Feb. 1990.

[37] D. Schardt, I. Schall, H. Geissel, H. Irnich, G. Kraft, A. Magel, M. Mohar, G. Münzenberg, F. Nickel, C. Scheidenberger, W. Schwab, and L. Sihver, "Nuclear fragmentation of high-energy heavy-ion beams in water," *Advances in Space Research*, vol. 17, pp. 87–94, Jan. 1996.

[38] I. Schall, D. Schardt, H. Geissel, H. Irnich, E. Kankeleit, G. Kraft, A. Magel, M. Mohar, G. Münzenberg, F. Nickel, C. Scheidenberger, and W. Schwab, "Charge-changing nuclear reactions of relativistic light-ion beams (5 $\leq$Z $\leq$ 10) passing through thick absorbers," *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, vol. 117, pp. 221–234, Sept. 1996.

[39] G. M. S., Aleksandrov D. V., Chulkov L. V., Kraus G., and Schardt D., "Fragmentation of 270 MeV/u carbon ions in water," GSI Preprint 97-08, 1997.

[40] B. Alpat, E. Pilicer, S. Blasko, D. Caraffini, F. D. Capua, V. Postolache, G. Saltanocchi, M. Menichelli, L. Desorgher, M. Durante, R. Pleskac, and C. L. Tessa, "Total and Partial Fragmentation Cross-Section of 500 MeV/nucleon Carbon Ions on Different Target Materials," *IEEE Trans. Nucl. Sci.*, vol. 60, pp. 4673–4682, Dec. 2013.

[41] A. N. Golovchenko, J. Skvarč, N. Yasuda, M. Giacomelli, S. P. Tretyakova, R. Ilić, R. Bimbot, M. Toulemonde, and T. Murakami, "Total charge-changing and partial cross-section measurements in the reactions of $\sim$ 110 – 250 M e V / n u c l e o n 12 C in carbon, paraffin, and water," *Phys. Rev. C*, vol. 66, p. 014609, July 2002.

[42] T. Toshito, K. Kodama, L. Sihver, K. Yusa, M. Ozaki, K. Amako, S. Kameoka, K. Murakami, T. Sasaki, S. Aoki, T. Ban, T. Fukuda, M. Komatsu, H. Kubota, N. Naganawa, T. Nakamura, T. Nakano, M. Natsume, K. Niwa, S. Takahashi, J. Yoshida, H. Yoshida, M. Kanazawa, N. Kanematsu, M. Komori, S. Sato, M. Asai, T. Koi, C. Fukushima, S. Ogawa, M. Shibasaki, and H. Shibuya, "Measurements of total and partial charge-changing cross sections for 200- to 400-MeV/nucleon C 12 on water and polycarbonate," *Phys. Rev. C*, vol. 75, p. 054606, May 2007.

[43] G. De Lellis, S. Buontempo, F. Di Capua, A. Di Crescenzo, P. Migliozzi, Y. Petukhov, C. Pistillo, A. Russo, P. Strolin, V. Tioukov, M. Durante, Y. Furusawa, T. Toshito, N. Yasuda, A. Ariga, and N. Naganawa, "Measurement of the fragmentation of Carbon nuclei used in hadron-therapy," *Nuclear Physics A*, vol. 853, pp. 124–134, Mar. 2011.

[44] C. Zeitlin, S. Guetersloh, L. Heilbronn, J. Miller, A. Fukumura, Y. Iwata, and T. Murakami, "Fragmentation cross sections of 290 and 400 MeV/nucleon C 12 beams on elemental targets," *Phys. Rev. C*, vol. 76, p. 014911, July 2007.

[45] C. Zeitlin, J. Miller, S. Guetersloh, L. Heilbronn, A. Fukumura, Y. Iwata, T. Murakami, S. Blattnig, R. Norman, and S. Mashnik, "Fragmentation of N 14 , O 16 , Ne 20 , and Mg 24 nuclei at 290 to 1000 MeV/nucleon," *Phys. Rev. C*, vol. 83, p. 034909, Mar. 2011.

[46] T. Ogawa, T. Sato, S. Hashimoto, D. Satoh, S. Tsuda, and K. Niita, "Energy-dependent fragmentation cross sections of relativistic C 12," *Phys. Rev. C*, vol. 92, p. 024614, Aug. 2015.

[47] M. Rovituso, C. Schuy, U. Weber, S. Brons, M. A. Cortés-Giraldo, C. La Tessa, E. Piasetzky, D. Izraeli, D. Schardt, M. Toppi, E. Scifoni, M. Krämer, and M. Durante, "Fragmentation of 120 and 200 MeV/u ^{4} He ions in water and PMMA targets," *Phys. Med. Biol.*, vol. 62, pp. 1310–1326, Feb. 2017.

[48] M. Marafini, R. Paramatti, D. Pinci, G. Battistoni, F. Collamati, E. De Lucia, R. Faccini, P. M. Frallicciardi, C. Mancini-Terracciano, I. Mattei, S. Muraro, L. Piersanti, M. Rovituso, A. Rucinski, A. Russomando, A. Sarti, A. Sciubba, E. Solfaroli Camillocci, M. Toppi, G. Traini, C. Voena, and V. Patera, "Secondary radiation measurements for particle therapy applications: nuclear fragmentation produced by ^{4} He ion beams in a PMMA target," *Phys. Med. Biol.*, vol. 62, pp. 1291–1309, Feb. 2017.

[49] A. Rucinski, G. Battistoni, F. Collamati, E. De Lucia, R. Faccini, P. M. Frallicciardi, C. Mancini-Terracciano, M. Marafini, I. Mattei, S. Muraro, R. Paramatti, L. Piersanti, D. Pinci, A. Russomando, A. Sarti, A. Sciubba, E. Solfaroli Camillocci, M. Toppi, G. Traini, C. Voena, and V. Patera, "Secondary radiation measurements for particle therapy applications: charged particles produced by ^{4} He and ^{12} C ion beams in a PMMA target at large angle," *Phys. Med. Biol.*, vol. 63, p. 055018, Mar. 2018.

[50] F. Horst, C. Schuy, U. Weber, K.-T. Brinkmann, and K. Zink, "Measurement of charge- and mass-changing cross sections for He 4 + C 12 collisions in the energy range 80–220 MeV/u for applications in ion beam therapy," *Phys. Rev. C*, vol. 96, p. 024624, Aug. 2017.

[51] F. Horst, G. Aricò, K.-T. Brinkmann, S. Brons, A. Ferrari, T. Haberer, A. Mairani, K. Parodi, C.-A. Reidel, U. Weber, K. Zink, and C. Schuy, "Measurement of He 4 charge- and mass-changing cross sections on H, C, O, and Si targets in the energy range 70–220 MeV/u for radiation transport calculations in ion-beam therapy," *Phys. Rev. C*, vol. 99, p. 014603, Jan. 2019.

[52] J. Mougey, R. Ost, M. Buenerd, A. Cole, C. Guet, D. Lebrun, J. Loiseaux, P. Martin, M. Maurel, E. Monnand, H. Nifenecker, P. Perrin, J. Pinston, C. Ristori, P. de Saintignon, F. Schussler, L. Carlén, B. Jakobsson, A. Oskarsson, I. Otterlund, B. Schroder, H. Gustafsson, T. Johansson, H. Ryde, J. Bondorf, O. Nielsen, and G. Tibell, "Projectile fragments from 86 MeV/nucleon 12C induced reactions," *Physics Letters B*, vol. 105, pp. 25–29, Sept. 1981.

[53] M. De Napoli, C. Agodi, G. Battistoni, A. A. Blancato, G. A. P. Cirrone, G. Cuttone, F. Giacoppo, M. C. Morone, D. Nicolosi, L. Pandola, V. Patera, G. Raciti, E. Rapisarda, F. Romano, D. Sardina, A. Sarti, A. Sciubba, V. Scuderi, C. Sfienti, and S. Tropea, "Carbon fragmentation measurements and validation of the Geant4 nuclear reaction models for hadrontherapy," *Phys. Med. Biol.*, vol. 57, pp. 7651–7671, Nov. 2012.

[54] J. Dudouet, D. Juliani, M. Labalme, D. Cussol, J. C. Angélique, B. Braunn, J. Colin, C. Finck, J. M. Fontbonne, H. Guérin, P. Henriquet, J. Krimmer, M. Rousseau, M. G. Saint-Laurent, and S. Salvador, "Double-differential fragmentation cross-section measurements of 95 MeV/nucleon 12 C beams on thin targets for hadron therapy," *Phys. Rev. C*, vol. 88, p. 024606, Aug. 2013.

[55] C. Divay, J. Colin, D. Cussol, C. Finck, Y. Karakaya, M. Labalme, M. Rousseau, S. Salvador, and M. Vanstalle, "Differential cross section measurements for hadron therapy: 50 MeV/nucleon $\^{}{12}\mathrm{C}$ reactions on H, C, O, Al, and $\^{}{\mathrm{nat}}\mathrm{Ti}$ targets," *Phys. Rev. C*, vol. 95, p. 044602, Apr. 2017. Publisher: American Physical Society.

[56] S. Agostinelli, J. Allison, K. Amako, J. Apostolakis, H. Araujo, P. Arce, M. Asai, D. Axen, S. Banerjee, G. Barrand, F. Behner, L. Bellagamba, J. Boudreau, L. Broglia, A. Brunengo, H. Burkhardt, S. Chauvie, J. Chuma, R. Chytracek, G. Cooperman, G. Cosmo, P. Degtyarenko, A. Dell'Acqua, G. Depaola, D. Dietrich, R. Enami, A. Feliciello, C. Ferguson, H. Fesefeldt, G. Folger, F. Foppiano, A. Forti, S. Garelli, S. Giani, R. Giannitrapani, D. Gibin, J. Gómez Cadenas, I. González, G. Gracia Abril, G. Greeniaus, W. Greiner, V. Grichine, A. Grossheim, S. Guatelli, P. Gumplinger, R. Hamatsu, K. Hashimoto, H. Hasui, A. Heikkinen, A. Howard, V. Ivanchenko, A. Johnson, F. Jones, J. Kallenbach, N. Kanaya, M. Kawabata, Y. Kawabata, M. Kawaguti, S. Kelner, P. Kent, A. Kimura, T. Kodama, R. Kokoulin, M. Kossov, H. Kurashige, E. Lamanna, T. Lampén, V. Lara, V. Lefebure, F. Lei, M. Liendl, W. Lockman, F. Longo, S. Magni, M. Maire, E. Medernach, K. Minamimoto, P. Mora de Freitas, Y. Morita, K. Murakami, M. Nagamatu, R. Nartallo, P. Nieminen, T. Nishimura, K. Ohtsubo, M. Okamura, S. O'Neale, Y. Oohata, K. Paech, J. Perl, A. Pfeiffer, M. Pia, F. Ranjard, A. Rybin, S. Sadilov, E. Di Salvo, G. Santin, T. Sasaki, N. Savvas, Y. Sawada, S. Scherer, S. Sei, V. Sirotenko, D. Smith,

N. Starkov, H. Stoecker, J. Sulkimo, M. Takahata, S. Tanaka, E. Tcherniaev, E. Safai Tehrani, M. Tropeano, P. Truscott, H. Uno, L. Urban, P. Urban, M. Verderi, A. Walkden, W. Wander, H. Weber, J. Wellisch, T. Wenaus, D. Williams, D. Wright, T. Yamada, H. Yoshida, and D. Zschiesche, "Geant4—a simulation toolkit," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 506, pp. 250–303, July 2003.

[57] 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. v. d. 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. V. d. Walle, E. Wieërs, and C. Morel, "GATE: a simulation toolkit for PET and SPECT," *Phys. Med. Biol.*, vol. 49, pp. 4543–4561, Sept. 2004. Publisher: IOP Publishing.

[58] J. Perl, J. Shin, J. Schümann, B. Faddegon, and H. Paganetti, "TOPAS: An innovative proton Monte Carlo platform for research and clinical applications," *Med Phys*, vol. 39, pp. 6818–6837, Nov. 2012.

[59] A. Ferrari, P. Sala, A. Fasso, and J. Ranft, "FLUKA: A Multi-Particle Transport Code," Tech. Rep. SLAC-R-773, 877507, Dec. 2005.

[60] K. Niita, T. Sato, H. Iwase, H. Nose, H. Nakashima, and L. Sihver, "PHITS—a particle and heavy ion transport code system," *Radiation Measurements*, vol. 41, pp. 1080–1090, Oct. 2006.

[61] T. Goorley, M. James, T. Booth, F. Brown, J. Bull, L. J. Cox, J. Durkee, J. Elson, M. Fensin, R. A. Forster, J. Hendricks, H. G. Hughes, R. Johns, B. Kiedrowski, R. Martz, S. Mashnik, G. McKinney, D. Pelowitz, R. Prael, J. Sweezy, L. Waters, T. Wilcox, and T. Zukaitis, "Initial MCNP6 Release Overview," *Nuclear Technology*, vol. 180, pp. 298–315, Dec. 2012. Publisher: Taylor & Francis _eprint: https://doi.org/10.13182/NT11-135.

[62] A. V. Dementyev and N. M. Sobolevsky, "shield — universal Monte Carlo hadron transport code: scope and applications," *Radiation Measurements*, vol. 30, pp. 553–557, Oct. 1999.

[63] G. Dedes and K. Parodi, "Monte Carlo Simulations of Particle Interactions with Tissue in Carbon Ion Therapy," *International Journal of Particle Therapy*, vol. 2, pp. 447–458, Dec. 2015.

[64] T. T. Böhlen, F. Cerutti, M. Dosanjh, A. Ferrari, I. Gudowska, A. Mairani, and J. M. Quesada, "Benchmarking nuclear models of FLUKA and GEANT4 for carbon ion therapy," *Phys. Med. Biol.*, vol. 55, pp. 5833–5847, Oct. 2010.

[65] I. Pshenichnov, A. Botvina, I. Mishustin, and W. Greiner, "Nuclear fragmentation reactions in extended media studied with Geant4 toolkit," *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, vol. 268, pp. 604–615, Mar. 2010.

[66] M. De Napoli, F. Romano, D. D'Urso, T. Licciardello, C. Agodi, G. Candiano, F. Cappuzzello, G. A. P. Cirrone, G. Cuttone, A. Musumarra, L. Pandola, and V. Scuderi, "Nuclear reaction measurements on tissue-equivalent materials and GEANT4 Monte Carlo simulations for hadrontherapy," *Phys. Med. Biol.*, vol. 59, pp. 7643–7652, Dec. 2014.

[67] J. Dudouet, D. Cussol, D. Durand, and M. Labalme, "Benchmarking geant4 nuclear models for hadron therapy with 95 MeV/nucleon carbon ions," *Phys. Rev. C*, vol. 89, p. 054616, May 2014.

[68] M. Puchalska, T. Tessonnier, K. Parodi, and L. Sihver, "Benchmarking of PHITS for Carbon Ion Therapy," *Int J Part Ther*, vol. 4, no. 3, pp. 48–55, 2018.

[69] A. F. Resch, H. Fuchs, and D. Georg, "Benchmarking GATE/Geant4 for16O ion beam therapy," *Phys. Med. Biol.*, vol. 62, pp. N474–N484, Sept. 2017. Publisher: IOP Publishing.

[70] J. F. Ward, "DNA Damage Produced by Ionizing Radiation in Mammalian Cells: Identities, Mechanism of Formation, and Reparability," vol. 35, pp. 95–125, 1988.

[71] M. Lomax, L. Folkes, and P. O'Neill, "Biological Consequences of Radiation-induced DNA Damage: Relevance to Radiotherapy," *Clinical Oncology*, vol. 25, pp. 578–585, Oct. 2013.

[72] Z. Nikitaki, V. Nikolov, I. V. Mavragani, E. Mladenov, A. Mangelis, D. A. Laskaratou, G. I. Fragkoulis, C. E. Hellweg, O. A. Martin, D. Emfietzoglou, V. I. Hatzi, G. I. Terzoudi, G. Iliakis, and A. G. Georgakilas, "Measurement of complex DNA damage induction and repair in human cellular systems after exposure to ionizing radiations of varying linear energy transfer (LET)," *Free Radical Research*, vol. 50, pp. S64–S78, Nov. 2016.

[73] G. van de Kamp, T. Heemskerk, R. Kanaar, and J. Essers, "DNA Double Strand Break Repair Pathways in Response to Different Types of Ionizing Radiation," *Front. Genet.*, vol. 12, p. 738230, Sept. 2021.

[74] I. Dokic, A. Mairani, M. Niklas, F. Zimmermann, N. Chaudhri, D. Krunic, T. Tessonnier, A. Ferrari, K. Parodi, O. Jäkel, J. Debus, T. Haberer, and A. Abdollahi, "Next generation multi-scale biophysical characterization of high precision cancer particle radiotherapy using clinical proton, helium-, carbon- and oxygen ion beams," *Oncotarget*, vol. 7, pp. 56676–56689, Aug. 2016.

[75] T. Friedrich, U. Scholz, T. ElsäSser, M. Durante, and M. Scholz, "Systematic analysis of RBE and related quantities using a database of cell survival experiments with ion beam irradiation," *Journal of Radiation Research*, vol. 54, pp. 494–514, May 2013.

[76] N. Amornwichet, T. Oike, A. Shibata, H. Ogiwara, N. Tsuchiya, M. Yamauchi, Y. Saitoh, R. Sekine, M. Isono, Y. Yoshida, T. Ohno, T. Kohno, and T. Nakano, "Carbon-Ion Beam Irradiation Kills X-Ray-Resistant p53-Null Cancer Cells by Inducing Mitotic Catastrophe," *PLoS ONE*, vol. 9, p. e115121, Dec. 2014.

[77] E. Cohen–Jonathan, E. J. Bernhard, and W. G. McKenna, "How does radiation kill cells?," *Current Opinion in Chemical Biology*, vol. 3, pp. 77–83, Feb. 1999.

[78] S. R. McKeown, "Defining normoxia, physoxia and hypoxia in tumours—implications for treatment response," *BJR*, vol. 87, p. 20130676, Mar. 2014.

[79] W. Tinganelli, M. Durante, R. Hirayama, M. Krämer, A. Maier, W. Kraft-Weyrather, Y. Furusawa, T. Friedrich, and E. Scifoni, "Kill-painting of hypoxic tumours in charged particle therapy," *Sci Rep*, vol. 5, p. 17016, Dec. 2015.

[80] H. Paganetti, "Nuclear interactions in proton therapy: dose and relative biological effect distributions originating from primary and secondary particles," *Phys. Med. Biol.*, vol. 47, pp. 747–764, Mar. 2002.

[81] M. Scholz, A. M. Kellerer, W. Kraft-Weyrather, and G. Kraft, "Computation of cell survival in heavy ion beams for therapy," *Radiation and Environmental Biophysics*, vol. 36, pp. 59–66, Mar. 1997.

[82] T. Inaniwa, T. Furukawa, Y. Kase, N. Matsufuji, T. Toshito, Y. Matsumoto, Y. Furusawa, and K. Noda, "Treatment planning for a scanned carbon beam with a modified microdosimetric kinetic model," *Phys. Med. Biol.*, vol. 55, pp. 6721–6737, Oct. 2010. Publisher: IOP Publishing.

[83] A. Mairani, T. T. Böhlen, A. Schiavi, T. Tessonnier, S. Molinelli, S. Brons, G. Battistoni, K. Parodi, and V. Patera, "A Monte Carlo-based treatment planning tool for proton therapy," *Phys. Med. Biol.*, vol. 58, pp. 2471–2490, Apr. 2013.

[84] T. T. Bohlen, J. Bauer, M. Dosanjh, A. Ferrari, T. Haberer, K. Parodi, V. Patera, and A. Mairani, "A Monte Carlo-based treatment-planning tool for ion beam therapy," *Journal of Radiation Research*, vol. 54, pp. i77–i81, July 2013.

[85] J. Schuemann, D. Giantsoudi, C. Grassberger, M. Moteabbed, C. H. Min, and H. Paganetti, "Assessing the Clinical Impact of Approximations in Analytical Dose Calculations for Proton Therapy," *International Journal of Radiation Oncology*Biology*Physics*, vol. 92, pp. 1157–1164, Aug. 2015.

[86] C. Grassberger and H. Paganetti, "Elevated LET components in clinical proton beams," *Phys. Med. Biol.*, vol. 56, pp. 6677–6691, Oct. 2011.

[87] Y. Kase, N. Kanematsu, T. Kanai, and N. Matsufuji, "Biological dose calculation with Monte Carlo physics simulation for heavy-ion radiotherapy," *Phys. Med. Biol.*, vol. 51, pp. N467–N475, Dec. 2006.

[88] G. Fattori, M. Riboldi, E. Scifoni, M. Krämer, A. Pella, M. Durante, S. Ronchi, M. Bonora, R. Orecchia, and G. Baroni, "Dosimetric effects of residual uncertainties in carbon ion treatment of head chordoma," *Radiother Oncol*, vol. 113, pp. 66–71, Oct. 2014.

[89] A. Lühr, D. C. Hansen, R. Teiwes, N. Sobolevsky, O. Jäkel, and N. Bassler, "The impact of modeling nuclear fragmentation on delivered dose and radiobiology in ion therapy," *Phys. Med. Biol.*, vol. 57, pp. 5169–5185, Aug. 2012.

[90] M. Krämer and M. Scholz, "Treatment planning for heavy-ion radiotherapy: calculation and optimization of biologically effective dose," *Phys. Med. Biol.*, vol. 45, pp. 3319–3330, Nov. 2000.

[91] ICRU, "Prescribing, Recording and Reporting Photon Beam Therapy," 1999.

[92] P. Fossati, N. Matsufuji, T. Kamada, and C. P. Karger, "Radiobiological issues in prospective carbon ion therapy trials," *Med. Phys.*, vol. 45, Nov. 2018.

[93] A.-C. Knopf and A. Lomax, "In vivo proton range verification: a review," *Phys. Med. Biol.*, vol. 58, pp. R131–R160, Aug. 2013.

[94] 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," *Phys. Med. Biol.*, vol. 56, pp. 1063–1082, Jan. 2011. Publisher: IOP Publishing.

[95] H. Rohling, W. Enghardt, L. Sihver, M. Priegnitz, and F. Fiedler, "Comparison of PHITS, GEANT4, and HIBRAC simulations of depth-dependent yields of {beta}{sup +}-emitting nuclei during therapeutic particle irradiation to measured data," *Physics in Medicine and Biology*, vol. 58, Sept. 2013.

[96] J. M. Verburg, H. A. Shih, and J. Seco, "Simulation of prompt gamma-ray emission during proton radiotherapy," vol. 57, p. 15, 2012.

[97] 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," *Phys. Med. Biol.*, vol. 59, pp. 1747–1772, Apr. 2014.

[98] J. Jeyasugiththan and S. W. Peterson, "Evaluation of proton inelastic reaction models in Geant4 for prompt gamma production during proton radiotherapy," *Phys. Med. Biol.*, p. 20, 2015.

[99] M. Vanstalle, I. Mattei, A. Sarti, F. Bellini, F. Bini, F. Collamati, E. D. Lucia, M. Durante, R. Faccini, F. Ferroni, C. Finck, S. Fiore, M. Marafini, V. Patera, L. Piersanti, M. Rovituso, C. Schuy, A. Sciubba, G. Traini, C. Voena, and C. L. Tessa, "Benchmarking Geant4 hadronic models for prompt-gamma monitoring in carbon ion therapy," *Medical Physics*, vol. 44, no. 8, pp. 4276–4286, 2017. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/mp.12348.

[100] F. Tommasino and M. Durante, "Proton Radiobiology," *Cancers*, vol. 7, pp. 353–381, Feb. 2015.

[101] G. Battistoni, M. Toppi, V. Patera, and The FOOT Collaboration, "Measuring the Impact of Nuclear Interaction in Particle Therapy and in Radio Protection in Space: the FOOT Experiment," *Front. Phys.*, vol. 8, p. 568242, Feb. 2021.

[102] L. Galli, A. M. Baldini, F. Cei, M. Chiappini, M. Francesconi, M. Grassi, U. Hartmann, M. Meucci, F. Morsani, D. Nicolò, A. Papa, S. Ritt, E. Schmid, and G. Signorelli, "WaveDAQ: An highly integrated trigger and data acquisition system," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 936, pp. 399–400, Aug. 2019.

[103] Y. Dong, S. Gianluigi, C. Sofia, A. Andrey, A. Behcet, A. Giovanni, A. Stefano, R. Arteche Diaz, B. Mattia, B. Nazar, B. Nicola, B. Elettra, B. Silvia, M. G. Bisogni, B. Graziano, C. Pietro, C. Piergiorgio, C. Esther, C. Alberto, G. De Lellis, A. Del Guerra, M. De Simoni, A. Di Crescenzo, B. Di Ruzza, D. Marco, D. Marco, F. Veronica, F. Emanuele, F. Christian, F. Elisa, F. Marta, F. Marco, F. Matteo, F. Gaia, G. Giuliana, G. Luca, G. Valerio, G. Giuseppe, H. Ronja, I. Enzo, I. Maria, K. Keida, A. C. Kraan, L. Valeria, C. La Tessa, L. Martina, L. Adele, E. Lopez Torres, M. Michela, M. Cristian, M. Ilaria, M. Alberto, M. Andrea, M. C. Montesi, M. C. Morone, M. Matteo, M. Silvia, N. Livio, P. Alessandra, P. Nadia, V. Patera, P. Francesco, P. Pisana, P. Marco, R. Fabrizio, R. Luciano, R. Riccardo, R. Valeria, S. Claudio, S. Alessio, S. Gabriella, S. Osamu, S. Simone, S. Lorenzo, S. Angelo, S. Christoph, S. Emanuele, S. Adalberto, S. Alexandre, S. Marco, S. Mario, S. Roberto, S. Eleuterio, S. Giancarlo, S. Achim, T. Sandro, T. Marco, T. Giacomo, V. Tioukov, S. M. Valle, M. Vanstalle, M. Villa, W. Ulrich, Z. Antonio, G. Battistoni, L. Servoli, and F. Tommasino, "The Drift Chamber detector

of the FOOT experiment: Performance analysis and external calibration," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 986, p. 164756, Jan. 2021.

[104] A. Nomerotski, O. Bachynska, J. Baudot, N. Chon-Sen, G. Claus, R. De Masi, M. Deveaux, W. Dulinski, R. Gauld, M. Goffe, J. Goldstein, I. M. Gregor, C. Hu-Guo, M. Imhoff, U. Koetz, W. Lau, C. Muntz, C. Santos, C. Schrader, M. Specht, J. Stroth, M. Winter, and S. Yang, "PLUME collaboration: Ultra-light ladders for linear collider vertex detector," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 650, pp. 208–212, Sept. 2011.

[105] S. Biondi, A. Alexandrov, B. Alpat, G. Ambrosi, S. Argirò, R. A. Diaz, N. Bartosik, G. Battistoni, N. Belcari, E. Bellinzona, M. G. Bisogni, G. Bruni, P. Carra, P. Cerello, E. Ciarrocchi, A. Clozza, S. Colombi, G. De Lellis, A. Del Guerra, M. De Simoni, A. Di Crescenzo, B. Di Ruzza, M. Donetti, Y. Dong, M. Durante, V. Ferrero, E. Fiandrini, C. Finck, E. Fiorina, M. Fischetti, M. Francesconi, M. Franchini, G. Franciosini, G. Galati, L. Galli, V. Gentile, G. Giraudo, R. Hetzel, E. Iarocci, M. Ionica, K. Kanxheri, A. C. Kraan, C. La Tessa, M. Laurenza, A. Lauria, E. L. Torres, M. Marafini, C. Massimi, I. Mattei, A. Mengarelli, A. Moggi, M. C. Montesi, M. C. Morone, M. Morrocchi, S. Muraro, L. Narici, A. Pastore, N. Pastrone, V. Patera, F. Pennazio, P. Placidi, M. Pullia, F. Raffaelli, L. Ramello, R. Ridolfi, V. Rosso, C. Sanelli, A. Sarti, G. Sartorelli, O. Sato, S. Savazzi, L. Scavarda, A. Schiavi, C. Schuy, E. Scifoni, A. Sciubba, A. Sécher, M. Selvi, L. Servoli, G. Silvestre, M. Sitta, E. Spiriti, G. Sportelli, A. Stahl, S. Tomassini, F. Tommasino, M. Toppi, G. Traini, T. Valeri, S. M. Valle, M. Vanstalle, U. Weber, A. Zoccoli, R. Spighi, and M. Villa, "The FragmentatiOn Of Target Experiment (FOOT) and its DAQ system," *arXiv:2010.16251 [nucl-ex, physics:physics]*, Oct. 2020. arXiv: 2010.16251.

[106] S. Colombi, A. Alexandrov, B. Alpat, G. Ambrosi, S. Argir, R. Arteche Diaz, N. Bartosik, G. Battistoni, N. Belcari, E. Bellinzona, S. Biondi, M. G. Bisogni, G. Bruni, P. Carra, P. Cerello, E. Ciarrocchi, A. Clozza, G. De Lellis, A. Del Guerra, M. De Simoni, A. Di Crescenzo, B. Di Ruzza, M. Donetti, Y. Dong, M. Durante, R. Faccini, V. Ferrero, E. Fiandri, C. Finck, E. Fiorina, M. Fischetti, M. Francesconi, M. Franchini, G. Franciosini, G. Galati, L. Galli,

V. Gentile, G. Giraudo, R. Hetzel, E. Iarocci, M. Ionica, A. Iuliano, K. Kanxheri, A. C. Kraan, V. Lante, C. La Tessa, M. Laurenza, A. Lauria, E. Lopez Torres, M. Marafini, C. Massimi, I. Mattei, A. Mengarelli, A. Moggi, M. C. Montesi, M. C. Morone, M. Morrocchi, S. Muraro, F. Murtas, A. Pastore, N. Pastrone, V. Patera, F. Pennazio, P. Placidi, M. Pullia, F. Raffaelli, L. Ramello, R. Ridolfi, V. Rosso, C. Sanelli, A. Sarti, G. Sartorelli, O. Sato, S. Savazzi, L. Scavarda, A. Schiavi, C. Schuy, E. Scifoni, A. Sciubba, A. Sécher, M. Selvi, L. Servoli, G. Silvestre, M. Sitta, R. Spighi, E. Spiriti, G. Sportelli, A. Stahl, V. Tioukov, S. Tommasini, F. Tommasino, M. Toppi, G. Traini, S. M. Valle, M. Vanstalle, M. Villa, U. Weber, R. Zarrella, and A. Zoccoli, "Enhancing the understanding of fragmentation processes in hadrontherapy and radioprotection in space with the FOOT experiment," *Phys. Scr.*, vol. 96, p. 114013, Nov. 2021.

[107] G. Traini and The FOOT Collaboration, "Performance of the ToF detectors in the FOOT experiment," *Il Nuovo Cimento C*, vol. 43, pp. 1–7, May 2020.

[108] A. Kraan, R. Zarrella, A. Alexandrov, B. Alpat, G. Ambrosi, S. Argirò, R. Arteche Diaz, N. Bartosik, G. Battistoni, N. Belcari, E. Bellinzona, S. Biondi, G. Bruni, P. Carra, P. Cerello, E. Ciarrocchi, A. Clozza, S. Colombi, G. De Lellis, A. Del Guerra, M. De Simoni, A. Di Crescenzo, B. Di Ruzza, M. Donetti, Y. Dong, M. Durante, R. Faccini, V. Ferrero, E. Fiandrini, C. Finck, E. Fiorina, M. Fischetti, M. Francesconi, M. Franchini, G. Franciosini, G. Galati, L. Galli, V. Gentile, G. Giraudo, R. Hetzel, E. Iarocci, M. Ionica, A. Iuliano, K. Kanxheri, V. Lante, C. La Tessa, M. Laurenza, A. Lauria, E. Lopez Torres, M. Marafini, C. Massimi, I. Mattei, A. Mengarelli, A. Moggi, M. Montesi, M. Morone, M. Morrocchi, S. Muraro, F. Murtas, A. Pastore, N. Pastrone, V. Patera, F. Pennazio, P. Placidi, M. Pullia, F. Raffaelli, L. Ramello, R. Ridolfi, V. Rosso, C. Sanelli, A. Sarti, G. Sartorelli, O. Sato, S. Savazzi, L. Scavarda, A. Schiavi, C. Schuy, E. Scifoni, A. Sciubba, A. Sécher, M. Selvi, L. Servoli, G. Silvestre, M. Sitta, R. Spighi, E. Spiriti, G. Sportelli, A. Stahl, V. Tioukov, S. Tomassini, F. Tommasino, M. Toppi, G. Traini, S. Valle, M. Vanstalle, M. Villa, U. Weber, A. Zoccoli, and M. Bisogni, "Charge identification of nuclear fragments with the FOOT Time-Of-Flight system," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrome-*

ters, Detectors and Associated Equipment, vol. 1001, p. 165206, June 2021.

[109] R. Rescigno, C. Finck, D. Juliani, E. Spiriti, J. Baudot, Z. Abou-Haidar, C. Agodi, M. Alvarez, T. Aumann, G. Battistoni, A. Bocci, T. Böhlen, A. Boudard, A. Brunetti, M. Carpinelli, G. Cirrone, M. Cortes-Giraldo, G. Cuttone, M. De Napoli, M. Durante, M. Gallardo, B. Golosio, E. Iarocci, F. Iazzi, G. Ickert, R. Introzzi, J. Krimmer, N. Kurz, M. Labalme, Y. Leifels, A. Le Fevre, S. Leray, F. Marchetto, V. Monaco, M. Morone, P. Oliva, A. Paoloni, V. Patera, L. Piersanti, R. Pleskac, J. Quesada, N. Randazzo, F. Romano, D. Rossi, M. Rousseau, R. Sacchi, P. Sala, A. Sarti, C. Scheidenberger, C. Schuy, A. Sciubba, C. Sfienti, H. Simon, V. Sipala, S. Tropea, M. Vanstalle, and H. Younis, "Performance of the reconstruction algorithms of the FIRST experiment pixel sensors vertex detector," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 767, pp. 34–40, Dec. 2014.

[110] C. A. Reidel, C. Finck, C. Schuy, M. Rovituso, and U. Weber, "Alignment procedure of silicon pixel detectors for ion-beam therapy applications," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 931, pp. 142–150, July 2019.

[111] L. Bugge and J. Myrheim, "Tracking and track fitting," *Nuclear Instruments and Methods*, vol. 179, pp. 365–381, Jan. 1981.

[112] Erwin Fehlberg, "Classical Seventh-, sixth- and fifth-order Runge-Kutta-Nyström formulas with stepsize control for general second-order differential equations," Nasa Technical Report R-432, NASA, 1974.

[113] C. Tsitouras, G. Papageorgiou, and T. Kalvouridis, "ON THE EFFICIENCY OF RUNGE-KUTTA-NYSTROM METHODS WITH INTERPOLANTS FOR SOLVING EQUATIONS OF THE FORM Y" = F(T, Y, yi) OVER SHORT TIMESPANS," vol. 53, pp. 329–346, 1992.

[114] R. Frühwirth, "Application of Kalman filtering to track and vertex fitting," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 262, pp. 444–450, Dec. 1987.

[115] P. Billoir, "PROGRESSIVE TRACK RECOGNITION WITh A KALMAN-LIKE FITI1NG PROCEDURE," vol. 57, pp. 390–394, 1989.

[116] R. E. Kalman and R. S. Bucy, "New Results in Linear Filtering and Prediction Theory," *Journal of Basic Engineering*, vol. 83, pp. 95–108, Mar. 1961.

[117] S. J. Julier and J. K. Uhlmann, "A New Extension of the Kalman Filter to Nonlinear Systems," p. 12.

[118] R. Mankel, "Pattern recognition and event reconstruction in particle physics experiments," *Rep. Prog. Phys.*, vol. 67, pp. 553–622, Apr. 2004.

[119] H. Paganetti, "Range uncertainties in proton therapy and the role of Monte Carlo simulations," *Phys. Med. Biol.*, vol. 57, pp. R99–R117, June 2012.

[120] C.-H. Min, C. H. Kim, M.-Y. Youn, and J.-W. Kim, "Prompt gamma measurements for locating the dose falloff region in the proton therapy," *Appl. Phys. Lett.*, vol. 89, p. 4, 2006.

[121] J. M. Verburg, K. Riley, T. Bortfeld, and J. Seco, "Energy- and time-resolved detection of prompt gamma-rays for proton range verification," *Phys. Med. Biol.*, vol. 58, pp. L37–L49, Oct. 2013.

[122] J. C. Polf, S. Peterson, G. Ciangaru, M. Gillin, and S. Beddar, "Prompt gamma-ray emission from biological tissues during proton irradiation: a preliminary study," *Phys. Med. Biol.*, vol. 54, pp. 731–743, Feb. 2009.

[123] M. Zarifi, S. Guatelli, D. Bolst, B. Hutton, A. Rosenfeld, and Y. Qi, "Characterization of prompt gamma-ray emission with respect to the Bragg peak for proton beam range verification: A Monte Carlo study," *Physica Medica*, vol. 33, pp. 197–206, Jan. 2017.

[124] C. Golnik, F. Hueso-González, A. Müller, P. Dendooven, W. Enghardt, F. Fiedler, T. Kormoll, K. Roemer, J. Petzoldt, A. Wagner, and G. Pausch, "Range assessment in particle therapy based on prompt gamma-ray timing measurements," *Phys. Med. Biol.*, vol. 59, pp. 5399–5422, Sept. 2014.

[125] F. Hueso-Gonz, F. Fiedler, C. Golnik, G. Janssens, J. Petzoldt, D. Prieels, and M. Priegnitz, "First test of the prompt gamma ray timing method with heterogeneous targets at a clinical proton therapy facility," *Phys. Med. Biol.*, p. 27, 2015.

[126] 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," *Phys. Med. Biol.*, vol. 57, pp. 3371–3405, June 2012.

[127] J. Krimmer, D. Dauvergne, J. Letang, and E. Testa, "Prompt-gamma monitoring in hadrontherapy: A review," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 878, pp. 58–73, Jan. 2018.

[128] G. Pausch, J. Berthold, W. Enghardt, K. Römer, A. Straessner, A. Wagner, T. Werner, and T. Kögler, "Detection systems for range monitoring in proton therapy: Needs and challenges," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 954, p. 161227, Feb. 2020.

[129] F. Roellinghoff, A. Benilov, D. Dauvergne, G. Dedes, N. Freud, G. Janssens, and J. Krimmer, "Real-time proton beam range monitoring by means of prompt-gamma detection with a collimated camera," *Phys. Med. Biol.*, p. 13, 2014.

[130] M. Priegnitz, S. Helmbrecht, G. Janssens, I. Perali, J. Smeets, F. Vander Stappen, E. Sterpin, and F. Fiedler, "Measurement of prompt gamma profiles in inhomogeneous targets with a knife-edge slit camera during proton irradiation," *Phys. Med. Biol.*, vol. 60, pp. 4849–4871, June 2015.

[131] E. Sterpin, "Analytical computation of prompt gamma ray emission and detection for proton range verification," *Phys. Med. Biol.*, p. 33, 2015.

[132] C. H. Min, H. R. Lee, C. H. Kim, and S. B. Lee, "Development of array-type prompt gamma measurement system for *in vivo* range verification in proton therapy: Development of array-type prompt gamma measurement system," *Med. Phys.*, vol. 39, pp. 2100–2107, Mar. 2012.

[133] 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," *Phys. Med. Biol.*, vol. 59, pp. 7653–7674, Dec. 2014.

[134] S. W. Peterson, D. Robertson, and J. D. Polf, "Optimization of a Compton camera for measuring gamma rays during proton treatment," vol. 55, p. 17, 2010.

[135] C. Richter, "First clinical application of a prompt gamma based in vivo proton range verification system," *Radiotherapy and Oncology*, p. 6, 2016.

[136] Y. Xie, "Prompt gamma imaging for in vivo range verification of pencil beam scanning proton therapy," p. 18, 2017.

[137] J. C. Polf, S. Avery, D. S. Mackin, and S. 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.*, p. 16, 2015.

[138] E. Draeger, D. Mackin, S. Peterson, H. Chen, S. Avery, S. Beddar, and J. C. Polf, "3D prompt gamma imaging for proton beam range verification," *Phys. Med. Biol.*, vol. 63, p. 035019, Jan. 2018.

[139] M. Testa, M. Bajard, M. Chevallier, D. Dauvergne, N. Freud, P. Henriquet, S. Karkar, F. Le Foulher, J. M. Létang, R. Plescak, C. Ray, M.-H. Richard, D. Schardt, and E. Testa, "Real-time monitoring of the Bragg-peak position in ion therapy by means of single photon detection," *Radiat Environ Biophys*, vol. 49, pp. 337–343, Aug. 2010.

[140] A. K. Biegun, E. Seravalli, P. C. Lopes, I. Rinaldi, M. Pinto, D. C. Oxley, P. Dendooven, F. Verhaegen, K. Parodi, P. Crespo, and D. R. Schaart, "Time-of-flight neutron rejection to improve prompt gamma imaging for proton range verification: a simulation study," *Phys. Med. Biol.*, vol. 57, pp. 6429–6444, Sept. 2012. Publisher: IOP Publishing.

[141] M. Priegnitz, "Measurement of prompt gamma profiles in inhomogeneous targets with a knife-edge slit camera during proton irradiation," *Phys. Med. Biol.*, p. 24, 2015.

[142] L. Tian, G. Landry, G. Dedes, F. Kamp, M. Pinto, K. Niepel, C. Belka, and K. Parodi, "Toward a new treatment planning approach accounting for *in vivo* proton range verification," *Phys. Med. Biol.*, vol. 63, p. 215025, Oct. 2018.

[143] F. 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," *Phys. Med. Biol.*, vol. 59, pp. 4427–4441, Aug. 2014.

[144] J. Constanzo, M. Vanstalle, C. Finck, D. Brasse, and M. Rousseau, "Dosimetry and characterization of a 25-MeV proton beam line for preclinical radiobiology research," *Medical Physics*, vol. 46, no. 5, pp. 2356–2362, 2019. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/mp.13512.

[145] D. Breton, E. Delagnes, J. Maalmi, and P. Rusquart, "The WaveCatcher family of SCA-based 12-bit 3.2-GS/s fast digitizers," in *2014 19th IEEE-NPSS Real Time Conference*, pp. 1–8, May 2014.

[146] R. Brun and F. Rademakers, "ROOT — An object oriented data analysis framework," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 389, pp. 81–86, Apr. 1997.

[147] A. Boudard, J. Cugnon, S. Leray, and C. Volant, "Intranuclear cascade model for a comprehensive description of spallation reaction data," *Phys. Rev. C*, vol. 66, p. 044615, Oct. 2002.

[148] A. Boudard, J. Cugnon, J.-C. David, S. Leray, and D. Mancusi, "New potentialities of the Liège intranuclear cascade model for reactions induced by nucleons and light charged particles," *Phys. Rev. C*, vol. 87, p. 014606, Jan. 2013.

[149] D. Mancusi, A. Boudard, J. Cugnon, J.-C. David, P. Kaitaniemi, and S. Leray, "Extension of the Liège intranuclear-cascade model to reactions induced by light nuclei," *Phys. Rev. C*, vol. 90, p. 054602, Nov. 2014.

[150] G. Folger, V. N. Ivanchenko, and J. P. Wellisch, "The Binary Cascade: Nucleon nuclear reactions," *Eur. Phys. J. A*, vol. 21, pp. 407–417, Sept. 2004.

[151] A. Koning and D. Rochman, "Modern Nuclear Data Evaluation with the TALYS Code System," *Nuclear Data Sheets*, vol. 113, pp. 2841–2934, Dec. 2012.

[152] K. Gudima, S. Mashnik, and V. Toneev, "Cascade-exciton model of nuclear reactions," *Nuclear Physics A*, vol. 401, pp. 329–361, June 1983.

[153] V. F. Weisskopf and D. H. Ewing, "On the Yield of Nuclear Reactions with Heavy Elements," *Phys. Rev.*, vol. 57, pp. 472–485, Mar. 1940. Publisher: American Physical Society.

[154] S. Furihata, "Statistical analysis of light fragment production from medium energy proton-induced reactions," *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, vol. 171, pp. 251–258, Nov. 2000.

[155] M. Guttormsen, T. S. Tveter, L. Bergholt, F. Ingebretsen, and J. Rekstad, "The unfolding of continuum y-ray spectra," p. 6, 1996.

[156] D. Radford, I. Ahmad, R. Holzmann, R. Janssens, and T. Khoo, "A prescription for the removal of Compton-scattered gamma rays from gamma-ray spectra," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 258, pp. 111–118, July 1987.

[157] M. Morháč, "Deconvolution methods and their applications in the analysis of -ray spectra," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 559, pp. 119–123, Apr. 2006.

[158] M. Morháč and V. Matoušek, "Complete positive deconvolution of spectrometric data," *Digital Signal Processing*, vol. 19, pp. 372–392, May 2009.

[159] R. Gold, "An iterative unfolding method for response matrices," Tech. Rep. ANL6984, Argonne National Laboratory, 1964.

[160] W. H. Richardson, "Bayesian-Based Iterative Method of Image Restoration*," *J. Opt. Soc. Am., JOSA*, vol. 62, pp. 55–59, Jan. 1972. Publisher: Optica Publishing Group.

[161] L. Lucy, "An iterative technique for the rectification of observed distributions," 1974.

[162] M. Morháč and V. Matoušek, "High-resolution boosted deconvolution of spectroscopic data," *Journal of Computational and Applied Mathematics*, vol. 235, pp. 1629–1640, Jan. 2011.

[163] B. Kozlovsky, R. J. Murphy, and R. Ramaty, "Nuclear Deexcitation Gamma-Ray Lines from Accelerated Particle Interactions," *ASTROPHYS J SUPPL S*, vol. 141, pp. 523–541, Aug. 2002.

Alexandre SÉCHER

Outils logiciels et algorithmiques pour la reconstruction de données nucléaires en hadronthérapie

Résumé

La hadronthérapie est une modalité de traitement du cancer utilisant des ions pour irradier les tumeurs. L'utilisation de ces ions apporte des propriétés balistiques supérieures par rapport à la radiothérapie conventionnelle, qui utilise des rayons X. Cependant, elle n'est pas sans contrepartie : les interactions entre les ions du faisceau et les noyaux du milieu ciblé peuvent produire des fragments qui vont altérer le dépôt de dose envisagé pour le traitement.
Ainsi, il est crucial d'être capable de prendre en compte l'impact de ce processus de fragmentation lors d'un traitement, afin de réaliser le potentiel latent de cette technologie. Pour cela, la détermination des sections efficaces de ces réactions est nécessaire. C'est en effet la tâche entreprise par la collaboration FOOT. Le travail effectué lors de cette thèse a constitué en l'élaboration d'un algorithme de reconstruction globale nécessaire à la détermination de ces données.
Le contrôle du traitement en hadronthérapie est difficile : malgré de nombreuses équipes de recherche travaillant sur le sujet, les méthodes envisagées doivent faire face à la faible statistique disponible pour établir un contrôle fiable. Une expérience autour du développement d'une nouvelle méthode de contrôle, basée sur le dopage de tumeurs avec des éléments caractérisés par leur profil d'émission gamma, a permis le développement d'algorithmes de déconvolution nécessaire à l'étude des spectres ainsi produits.

Abstract

Hadrontherapy is a cancer treatment modality using ions to irradiate tumors. The use of those ions brings superior balistical properties over conventional radiotherapy, which uses x-rays. However, this modality also has a few downsides: the interactions between the ions and the target nuclei can produce fragments which will alter the dose deposition.
Thus, it is crucial to be able to take into account the impact the fragmentation process during a treatment, in order to fully exploit the advantages of the modality. To do so, the determination of the fragmentation cross sections is mandatory. It is indeed the task undertaken by the FOOT collaboration. The work done during this thesis lead to the elaboration of a global reconstruction algorithm, necessary in order to extract this information.
Treatment monitoring in hadrontherapy is difficult: despite extensive work done by several research teams around the world, the methods developed have to face intrinsically low available statistics in order to establish range verification. An experiment focused on the development of a novel method, based on the enhancement of tumoral volume with elements selected according to their gamma emission profile, allowed for the development of unfolding algorithms mandatory in order to study the produced spectra.