Versnelling van GATE Monte Carlo simulaties

Acceleration of GATE Monte Carlo Simulations

Jan De Beenhouwer

Promotoren: prof. dr. I. Lemahieu, prof. dr. ir. S. Staelens
Proefschrift ingediend tot het behalen van de graad van
Doctor in de Ingenieurswetenschappen: Biomedische Ingenieurstechnieken

Vakgroep Elektronica en Informatiesystemen
Voorzitter: prof. dr. ir. J. Van Campenhout
Faculteit Ingenieurswetenschappen
Academiejaar 2008 - 2009

UNIVERSITEIT
GENT

ISBN 978-90-8578-231-5
NUR 954
Wettelijk depot: D/2008/10.500/51

Acknowledgments

This dissertation can provide at most one doctor's degree, while it is in fact the work of many who have guided and supported me through the course of this PhD. There is no better place for me to thank them all than at the very start of this dissertation.

In the first place I would like to thank my promotors prof. dr. Ignace Lemahieu and prof. dr. ir. Steven Staelens. I had already been interested in multiple disciplines as a computer science engineering student, so when I met Ignace I knew I was offered the opportunity of a lifetime. He opened the door to the fascinating world of nuclear medical imaging. He always supported my requests to join international workshops and conferences where I could present my work to the experts in our field. Thank you very much for this. Steven has contributed enormously to the genesis of this PhD both in terms of intellectual input and of support and encouragement. Not a word of mine has appeared in print or in this dissertation which he has not relentlessly read and reflected and commented on. This dissertation would simply not have reached its current level without you Steven, thank you very much.

Right from the start of my Phd I was introduced into the worldwide OpenGATE collaboration where I could join some of the foremost experts in Monte Carlo simulations for nuclear medicine. I thank every member of OpenGATE for all those nice GATE meetings and the useful comments on my work. I would especially like to thank dr. ir. Dennis Schaart for his genuine interest during my GATE presentations and prof. dr. Irène Buvat for her critical comments. Knowing

that my contributions to GATE are available to researchers all over the world gave me a great motivation to complete this dissertation.

At times a PhD is viewed upon as a solitary experience. This was however never the case for me and I have all my colleagues at MEDISIP and IBITECH to thank for it. In particular, my thanks go to Jeroen Verhaeghe with whom I shared an office for two years. We experienced many sometimes slightly absurd but always memorable situations together, especially on conferences and after hours during rock climbing. I wish you all the best in Canada. My appreciation also goes to Roel Van Holen for his down to earth view on and interest in my research. The short course in Archamps will be remembered for years to come. I wish you good luck with the writing of your dissertation. My thanks also goes to prof. dr. Stefaan Vandenberghe for his many useful suggestions and comments on how to improve my work and for always providing yet another viewpoint. I also wish my other colleagues and former colleagues: Erwann Rault, Long Zhang, Vincent Keereman, Jelle De Smet, Steven Deleye, Tom Ghekiere, Yves D'Asseler and Stijn De Clercq good luck with their research and jobs these next years. Furthermore I would like to thank my MR and EEG colleagues for their interest and support: Hans Hallez, Sara Assecondi, Peter Van Hese, Els Fieremans, Steven Delputte, Ronald Phlypo and Mahir Ozdemir. My gratitude also goes to the administrative staff of our department for taking care of all the administrative details during these years. In particular I would like to thank Rita Breems, Ayfer Ugurlu and Anne-Marie Van der Biest.

Aside from all these colleagues I would like to thank my family for their love and support. I especially want to thank my parents. My mother for her beloved concern and my father for never stopping to believe in my possibilities and for showing me what is important in life.

Finally, I dedicate this work to my wife Lieve. Countless were the nights during which I yet again had to work late into the night. She has continuously supported and encouraged me to complete this work.

Jan De Beenhouwer, June 2008

Table of Contents

List of Figures

List of Tables

List of Acronyms

A

AE	Acceleration Estimate
AF	Acceleration Factor
aGATE	Analog GATE
AM	Amdahl's law

B

Bexxar	Tositumomab, a therapeutic monoclonal antibody
Bq	Becquerel

C

CFD	Convolution-based Forced Detection
cps	counts per second

D

DOI	Depth Of Interaction
DOTATOC	DOTA-d-Phe(1)-Tyr(3)-octreotide

E

EGS4	Electron Gamma Shower 4
eV	electron Volt

F

FD	Forced Detection
FF	Form Factor
FOM	Figure of Merit
FOV	Field of View
FWHM	Full Width at Half Maximum
FWTM	Full Width at Tenth Maximum

G

GATE	Geant4 Application for Tomographic Emission
Geant4	GEometry ANd Tracking 4
GIS	Geometrical Importance Sampling

H

HEGP	High Energy General Purpose

L

LEHR	Low Energy High Resolution

M

MCAT	Mathematical Cardiac Torso
MCNP	Monte Carlo N-Particle
MEGP	Medium Energy General Purpose
MLCRNG	Multiplicative linear congruential random number generators
MPI	Message Passing Interface
MPS	Multiple Projection Sampling

P

PDF	Probability Density Function
PET	Positron Emission Tomography
PMT	Photo Multiplier Tube
PSC	Pure Scattered Coincidences
PSF	Point Spread Funciton
PVE	Partial Volume Effect

P

QF	Quality Factor

R

RFOM	Relative Figure of Merit
RNG	Random Number Generator
ROC	Receiver Operating Characteristic
ROI	Region of Interest
RR	Russian Roulette

S

SC	Spurious Contamination
SCC	Scattered Spurious Contamination
SF	Scattering Function
sGATE	Simplified GATE
SIMIND	Simulating Medical Imaging Nuclear Detectors
SimSET	Simulation System for Emission Tomography
SORTEO	Simulation Of Realistic Threedimensional Emitting Objects
SPECT	Single Photon Emission Computed Tomography

T

T	True Coincidences
TOF	Time of Flight
TRT	Targeted Radionuclide Therapy

Nederlandstalige samenvatting –Summary in Dutch–

Positron Emission Tomography (PET) en Single Photon Emission Computed Tomography (SPECT) zijn medische beeldvormingstechnieken die functionele beelden opleveren die biologische processen weerspiegelen. Beide zijn gebaseerd op het speurstof principe. Een biologisch actieve stof, een farmaceutisch product, wordt geselecteerd zodat de ruimtelijke en tijdsafhankelijke verspreiding doorheen het lichaam gerelateerd is aan een bepaalde lichaamsfunctie of stofwisseling. De speurstof wordt radioactief gemerkt tot een radiofarmaceutisch product of tracer. Na het toedienen van het radiofarmaceutisch product kan men de gegenereerde fotonen detecteren. De verkregen data worden achteraf gereconstrueerd om een driedimensionale verdeling van de radioactieve speurstof in de patiënt te verkrijgen.

Monte-Carlomethoden zijn numerieke methoden die gebruik maken van willekeurige getallen om een resultaat te berekenen. Men maakt gebruik van een willekeurige variabele waarvan de verwachtingswaarde gelijk is aan het beoogde resultaat. Door de willekeurige variabele te bemonsteren kan men het gemiddelde en de variantie gebruiken als schatting van het beoogde resultaat. Het is een poging om de natuur te modelleren door middel van directe simulatie van de essentiële dynamiek van het systeem in kwestie. Monte-Carlomodellen genieten de voorkeur voor toepassingen waarbij experimentele metingen niet mogelijk zijn of waarbij analytische modellen niet beschikbaar zijn door de complexe aard van het probleem. Deze methoden zijn heel geschikt voor de medische fysica

gezien de stochastische aard van de stralingsemissie, het deeltjestransport en de detectieprocessen. Ze vormen belangrijke bijdragen in toepassingsgebieden zoals detectorontwerp en correctietechnieken voor attenuatie en verstrooiing. Er zijn verscheidene gespecialiseerde Monte-Carlosimulatoren voor PET en/of SPECT beschikbaar. Deze zijn echter vaak niet gedetailleerd of flexibel genoeg voor realistische simulaties van complexe detectorgeometrieën en tijdsafhankelijke processen zoals het radioactief verval, speurstofverspreiding, patiëntbeweging en detectorbeweging. De Monte-Carlosimulator die gebruikt wordt in deze doctoraatsthesis, GEANT4 Application for Tomographic Emission (GATE), werd speciaal ontworpen om een oplossing te bieden voor deze problemen.

De flexibiliteit die eigen is aan GATE heeft echter een schaduwzijde. De simulatie van een eenvoudig SPECT detector prototype is misschien mogelijk binnen een tijdsduur van enkele uren, maar een acquisitie van een realistisch fantoom met een realistische detector kan evenwel jaren duren met een enkelvoudige CPU. Het kernwoord van deze doctoraatsthesis is daarom: efficiëntie. De versnelling van GATE simulaties kan enkel worden bereikt door middel van een combinatie van efficiënte data analyse, gespecialiseerde variantie reductie, snelle navigatie algoritmen en parallellisatie.

In het eerste deel van deze doctoraatsthesis beschouwen we de verbetering van de analyse module van GATE. De statische analyse module is niet flexibel en het is niet mogelijk om meer detail te bewaren zonder in te boeten aan efficiëntie. Een flexibele, gedetailleerde en efficiënte analyse module is echter onontbeerlijk voor ontwerp en de validatie van de versnellingstechnieken in deze thesis. Om dit doel te bereiken ontwerpen we een nieuw analyse raamwerk dat de mogelijkheid biedt om elk proces, gaande van het radioactief verval van isotopen tot de interactie en detectie van deeltjes in eender welk detector element voor eender welk type fantoom. De evaluatie van ons raamwerk bestaat enerzijds uit een analyse van valse detecties bij onzuivere PET isotopen zoals ^{124}I-Bexxar en anderzijds uit een analyse van contaminatie bij SPECT isotopen zoals ^{131}I-Bexxar. In het geval van PET beschrijven we hoe ons raamwerk zelfs voor realistische fantomen valse detecties kan opsporen bij onzuivere isotopen.

We tonen aan dat het mogelijk is om geoptimaliseerde energievensters te bepalen, welke onmiddellijk toepasbaar zijn in de kliniek, om de contaminatie door valse detecties te minimaliseren. We tonen ook aan dat de valse detecties zelf spatiaal niet uniform zijn verdeeld. Om deze reden zijn de standaard reconstructie- en correctiemethoden niet afdoend. In het geval van SPECT beschrijven we hoe het mogelijk is om detecties in te delen naargelang hun interactiegeschiedenis: geometrische detecties, verstrooiing in het fantoom, penetratie doorheen de collimator, verstrooiing in de collimator en vertrooiing in het eindgedeelte van de detector. We tonen aan dat standaard correctiemethoden gebaseerd op energievensters niet kunnen corrigeren voor septale penetratie. We tonen ook aan dat ^{124}I PET met geoptimaliseerde energievensters een betere beeldkwaliteit oplevert dan ^{131}I SPECT wanneer van standaard reconstructietechnieken gebruik wordt gemaakt.

In het tweede deel van deze doctoraatsthesis besteden we bijzondere aandacht aan het verbeteren van de efficiëntie van GATE door middel van een variantie reductie techniek: Geometrical Importance Sampling (GIS). We beschrijven hoe slechts 0.02% van alle fotonen het kristaloppervlak van een lage energie hoge resolutie SPECT detector kan bereiken. Heel wat rekenkracht gaat dus verloren bij het simuleren van fotonen die niet zullen bijdragen tot het beoogde resultaat. We maken gebruik van een tweevoudige strategie om dit probleem op te lossen: enerzijds maakt GIS gebruik van Russische Roulette om die fotonen te verwijderen uit de simulatie waarvan wordt vermoed dat ze waarschijnlijk niet zullen bijdragen tot het resultaat. Anderzijds worden fotonen in belangrijke regios gesplitst in meedere fotonen met een kleiner gewicht om de overlevingskans te vergroten. We tonen aan dat deze techniek vertakkingen introduceert in de interactiegeschiedenis van een deeltje. We beschrijven hoe we dit in rekening brengen door middel van een boomstructuur die de interactiegeschiedenis weerspiegelt en waarmee we de resultaten kunnen analyseren. De evaluatiestudie bestaat uit een validatie van energiespectra, spatiale resolutie en sensitiviteit voor isotopen van lage en middelhoge energie. We tonen aan dat GIS versnellingsfactoren tussen 5 en 13 kan bereiken tegenover analoge GATE simulaties voor

de isotopen in deze studie. GIS is een algemene versnellingstechniek die gebruikt kan worden voor eender welke combinatie van isotoop, fantoom en detector.

Hoewel GIS nuttig is als een veilige en nauwkeurige versnellingstechniek, is de tijdsduur van een simulatie niet klinisch aanvaardbaar. De belangrijkste reden hiervoor is dat het niet in staat is om fotonen uit te sturen in een voorafbepaalde, gedwongen richting. In het derde deel van deze doctoraatsthesis lossen we dit probleem op voor ^{99m}Tc SPECT simulaties. Onze aanpak is tweeledig.

Eerst introduceren we twee variantie reductie technieken: gedwongen detectie (eng.:Forced Detection, FD) en op convolutie gebaseerde gedwongen detectie (eng.: Convolution-based Forced Detection, CFD) waarbij meerdere projecties tegelijkertijd worden gegenereerd (eng.: Multiple Projection Sampling, MPS). FD en CFD zorgen ervoor dat bij elke emissie en interactie copieën worden gemaakt van een foton. Deze worden vervolgens getransporteerd doorheen het fantoom in een richting bemonsterd in een ruimtehoek die gericht is naar de SPECT detector en dit voor alle projectiehoeken tegelijkertijd. We beschrijven hoe we een gewicht toekennen aan iedere copie om te compenseren voor de gedwongen richting en het feit dat het foton niet wordt geabsorbeerd in het fantoom. We geven aan dat de gewichten worden berekend door middel van de totale en differentiële Compton en Rayleigh werkzame doorsneden per electron, met inbegrip van Hubbell's atomische factoren. In het geval van FD worden alle detectorinteracties gemodelleerd door Monte-Carlobemonstering, terwijl deze in het geval van CFD analytisch worden gemodelleerd.

Vervolgens beschrijven we het ontwerp van een voor FD en CFD gespecialiseerde navigator om de trage algoritmen voor deeltjestransport in GEANT4 te versnellen. De validatiestudie toont aan dat zowel FD als CFD de analoge GATE simulaties zeer goed benaderen en dat we een versnellingsfactor tussen 3 en 6 grootteordes kunnen bereiken in vergelijking met analoge GATE simulaties. Dit laat een simulatie toe van een realistische acquisitie met een complex fantoom binnen 130 seconden op een hedendaagse desktop computer.

In het vierde en laatste deel van deze doctoraatsthesis maken we

gebruik van een intrinsieke eigenschap van Monte-Carlosimulaties: parallelle verwerking. We tonen aan hoe Monte-Carlosimulaties theoretisch gezien lineair moeten schalen als functie van het aantal processoren, maar dat dit meestal niet mogelijk is door de insteltijd van simulatiejobs en dataverwerking. Onze aanpak is gebaseerd op twee stappen: jobdistributie en dataverwerking. De jobdistributie is gebaseerd op splitsing in het tijdsdomein waarbij alle experimentele parameters worden weerhouden en waarbij de statistische onafhankelijkheid van iedere subsimulatie wordt gegarandeerd. We reduceren ook de insteltijd van simulatiejobs door het ontwerp van een geparameteriseerd collimator model voor SPECT simulaties. We verminderen ook de tijd die nodig is voor de dataverwerking door middel van een ketting-gebaseerde dataverwerkingseenheid. Een schaalbaarheidsstudie gebaseerd op een cluster van 70 CPUs toont aan dat een versnellingsfactor van ongeveer 66 op 70 CPUs kan worden bereikt voor zowel PET als SPECT simulaties. We tonen ook aan dat onze methode van parallellisatie geen benaderingen introduceert en dat onze methode kan worden gecombineerd met alle voorgaande versnellingstechnieken van deze doctoraatsthesis.

English summary

Positron Emission Tomography (PET) and Single Photon Emission Computed Tomography are forms of medical imaging that produce functional images that reflect biological processes. They are based on the tracer principle. A biologically active substance, a pharmaceutical, is selected so that its spatial and temporal distribution in the body reflects a certain body function or metabolism. In order to form images of the distribution, the pharmaceutical is labeled with gamma-ray-emitting or positron-emitting radionuclides (radiopharmaceuticals or tracers). After administration of the tracer to a patient, an external position-sensitive gamma-ray camera can detect the emitted radiation to form a stack of images of the radionuclide distribution after a reconstruction process.

Monte Carlo methods are numerical methods that use random numbers to compute quantities of interest. This is normally done by creating a random variable whose expected value is the desired quantity. One then simulates and tabulates the random variable and uses its sample mean and variance to construct probabilistic estimates. It represents an attempt to model nature through direct simulation of the essential dynamics of the system in question. Monte Carlo modeling is the method of choice for all applications where measurements are not feasible or where analytic models are not available due to the complex nature of the problem. In addition, such modeling is a practical approach in nuclear medical imaging in several important application fields: detector design, quantification, correction methods for image degradations, detection tasks etc. Several powerful dedicated Monte Carlo simulators for PET and/or SPECT are available. However, they are often not detailed nor flexible enough to

enable realistic simulations of emission tomography detector geometries while also modeling time dependent processes such as decay, tracer kinetics, patient and bed motion, dead time or detector orbits. Our Monte Carlo simulator of choice, GEANT4 Application for Tomographic Emission (GATE), was specifically designed to address all these issues.

The flexibility of GATE comes at a price however. The simulation of a simple prototype SPECT detector may be feasible within hours in GATE but an acquisition with a realistic phantom may take years to complete on a single CPU. In this dissertation we therefore focus on the Achilles' heel of GATE: efficiency. Acceleration of GATE simulations can only be achieved through a combination of efficient data analysis, dedicated variance reduction techniques, fast navigation algorithms and parallelization.

In the first part of this dissertation we consider the improvement of the analysis capabilities of GATE. The static analysis module in GATE is both inflexible and incapable of storing more detail without introducing a large computational overhead. However, the design and validation of the acceleration techniques in this dissertation requires a flexible, detailed and computationally efficient analysis module. To this end, we develop a new analysis framework capable of analyzing any process, from the decay of isotopes to particle interactions and detections in any detector element for any type of phantom. The evaluation of our framework consists of the assessment of spurious activity in ^{124}I-Bexxar PET and of contamination in ^{131}I-Bexxar SPECT. In the case of PET we describe how our framework can detect spurious coincidences generated by non-pure isotopes, even with realistic phantoms. We show that optimized energy thresholds, which can readily be applied in the clinic, can now be derived in order to minimize the contamination. We also show that the spurious activity itself is not spatially uniform. Therefore standard reconstruction and correction techniques are not adequate. In the case of SPECT we describe how it is now possible to classify detections into geometric detections, phantom scatter, penetration through the collimator, collimator scatter and backscatter in the end parts. We show that standard correction algorithms such as triple energy

window correction cannot correct for septal penetration. We demonstrate that ^{124}I PET with optimized energy thresholds offer better image quality than ^{131}I SPECT when using standard reconstruction techniques.

In the second part of this dissertation we focus on improving the efficiency of GATE with a variance reduction technique called Geometrical Importance Sampling (GIS). We describe how only 0.02% of all emitted photons can reach the crystal surface of a SPECT detector head with a low energy high resolution collimator. A lot of computing power is therefore wasted by tracking photons that will not contribute to the result. A twofold strategy is used to solve this problem: GIS employs Russian Roulette to discard those photons that will not likely contribute to the result. Photons in more important regions on the other hand are split into several photons with reduced weight to increase their survival chance. We show that this technique introduces branches into the particle history. We describe how this can be taken into account by a particle history tree that is used for the analysis of the results. The evaluation of GIS consists of energy spectra validation, spatial resolution and sensitivity for low and medium energy isotopes. We show that GIS reaches acceleration factors between 5 and 13 over analog GATE simulations for the isotopes in the study. It is a general acceleration technique that can be used for any isotope, phantom and detector combination.

Although GIS is useful as a safe and accurate acceleration technique, it cannot deliver clinically acceptable simulation times. The main reason lies in its inability to force photons in a specific direction. In the third part of this dissertation we solve this problem for ^{99m}Tc SPECT simulations. Our approach is twofold. Firstly, we introduce two variance reduction techniques: forced detection (FD) and convolution-based forced detection (CFD) with multiple projection sampling (MPS). FD and CFD force copies of photons at decay and at every interaction point to be transported through the phantom in a direction sampled within a solid angle toward the SPECT detector head at all SPECT angles simultaneously. We describe how a weight must be assigned to each photon in order to compensate for the forced direction and non-absorption at emission and scatter. We

show how the weights are calculated from the total and differential Compton and Rayleigh cross sections per electron with incorporation of Hubbell's atomic form factor. In the case of FD all detector interactions are modeled by Monte Carlo, while in the case of CFD the detector is modeled analytically. Secondly, we describe the design of an FD and CFD specialized navigator to accelerate the slow tracking algorithms in GEANT4. The validation study shows that both FD and CFD closely match the analog GATE simulations and that we can obtain an acceleration factor between 3 (FD) and 6 (CFD) orders of magnitude over analog simulations. This allows for the simulation of a realistic acquisition with a torso phantom within 130 seconds.

In the fourth part of this dissertation we exploit the intrinsic parallel nature of Monte Carlo simulations. We show how Monte Carlo simulations should scale linearly as a function of the number of processing nodes but that this is usually not achieved due to job setup time, output handling and cluster overhead. We describe how our approach is based on two steps: job distribution and output data handling. The job distribution is based on a time-domain partitioning scheme that retains all experimental parameters and that guarantees the statistical independence of each subsimulation. We also reduce the job setup time by the introduction of a parameterized collimator model for SPECT simulations. We reduce the data output handling time by a chain-based output merger. The scalability study is based on a set of simulations on a 70 CPU cluster and shows an acceleration factor of approximately 66 on 70 CPUs for both PET and SPECT. We also show that our method of parallelization does not introduce any approximations and that it can be readily combined with any of the previous acceleration techniques described above.

1
General Introduction

In the 1940's crude spatial information about radioactive source distributions within the brain were produced using a single detector positioned at various locations around the head. During the 1950's this technology was improved upon and it became possible to obtain images of the distribution of radionuclides in the body. 1951 saw the development of the rectilinear scanner by Benedict Cassen (1). It produced planar images but required a very long imaging time. In 1953 Hal Anger used a lead pin hole to project a gamma ray image of a source distribution onto a scintillating screen with a photographic film behind it. Due to losses in the film, the technique required extremely long imaging times. This problem remained until 1958. In that year Hal Anger replaced the film by a NaI crystal and photo multiplier tubes (PMT) and presented the Anger camera (2). Its basic principles are still in use today. The final revolution that marked the start of a new age in nuclear medical imaging, was the development of code to reconstruct tomographic images from a set of projections around the patient by Cormack (3). This enabled the development of Positron Emission Tomography (PET) by Phelps (4) and Single Photon Emission Computed Tomography (SPECT) by Kuhl (5) during the 1970's. Both methods provide functional information by showing the 3D distribution of an injected radiopharma-

ceutical in the body. Today the functional information of PET and SPECT is combined with anatomical imaging modalities such as CT or MRI. It can therefore be used as a diagnostic tool for the detection of small tumours in an early stage which is of increasing interest given the aging population in Western Europe and North America.

Nowadays, the design of new detector prototypes, quantification analysis, correction methods to improve reconstructed images, etc. are done with Monte Carlo simulations. Monte Carlo methods have the advantage of being able to model all physics related to PET and SPECT. It is a numerical solution to a problem that models objects, such as photons and electrons, interacting with other objects or their environment based upon object-object or object-environment relationships defined by particle physics theory and experimental data. A stochastic solution is determined by random sampling of the relationships until the result converges (6). Since the history of every simulated particle can be analyzed, Monte Carlo simulations provide a solution to a wide range of problems that could not be solved by experimental or analytical approaches.

The breakthrough of Monte Carlo methods in medical physics came with the paper of Raeside in 1976 (7). Since then the influence of Monte Carlo simulations on PET and SPECT has only increased. Over a dozen Monte Carlo simulators for PET and/or SPECT have been developed, each with their own advantages, disadvantages and limitations. However, none of these was flexible enough to allow the realistic simulation of the detector geometry in PET and SPECT. This became possible with the public release of the GEANT4 Application for Tomographic Emission (GATE) (8) on may 7th 2004. The OpenGATE collaboration was founded by research groups at Lausanne University (UNIL), Ghent University (ELIS/Medisip) and Clermont-Ferrand University (LPC). The collaboration has since then grown to over 20 research labs worldwide and GATE has manifested itself as the golden standard for PET and SPECT Monte Carlo simulations.

The Achilles heel of GATE is however its computational complexity. A realistic simulation in GATE may take years to complete on a single processor. Therefore, the major aim of this dissertation was to contribute to the acceleration of GATE simulations. A fast GATE simulator will not only provide faster simulations which enhance the practical usability of GATE. It will also allow GATE to be incorporated into reconstruction algorithms which in turn will lead

to better image quality.

Chapter 2 gives an overview of nuclear medical imaging and more specifically PET and SPECT from acquisition to reconstruction. The general principles of Monte Carlo simulations are explained as well as their relevance to nuclear medical imaging. In addition, an overview is given of the different Monte Carlo simulators for PET and SPECT. The general high energy physics library GEANT4 is introduced and the most important physics implementations related to PET and SPECT are explained. Subsequently GATE, built upon GEANT4, is introduced as the most flexible and accurate simulator to date. The final part of chapter 2 will focus entirely on the computational complexity of analog GATE (aGATE) simulations (i.e. without acceleration methods). The different problems resulting in inefficiency and their relation to each other are presented and a number of solutions are proposed that will be discussed in the remaining chapters.

Chapter 3 is focused completely on the enhancement of analysis capabilities of PET and SPECT simulations with GATE. Firstly, a much more detailed analysis module called ProcessGATE is developed for aGATE, which will be used as a validation tool for the methods developed in all remaining chapters. Secondly, the efficiency of the analysis module is greatly enhanced for aGATE. This enhancement is used directly by all remaining chapters.

Chapters 4, 5 and 6 focus entirely on the acceleration of GATE. The use of geometrical importance sampling (GIS), a variance reduction technique, is discussed in Chapter 4. It provides a general efficiency improvement for any combination of SPECT isotope, phantom and detector. The efficiency of GATE SPECT simulations is further enhanced with the development of forced detection and convolution-based forced detection in chapter 5. Fast particle navigation algorithms were developed specifically for these methods. The application of parallelization to PET and SPECT GATE simulations is discussed subsequently in chapter 6.

Chapter 7 summarizes the main contributions of this dissertation and highlights some aspects for future study.

Key to the efficiency improvements in GATE is the integration of the different techniques described in chapters 3 to 6. The different chapters are related to each other through these techniques as shown in figure 1.1. Moreover, each of the techniques was incorporated into a single version of GATE in order to retain the general applicability

of GATE as a PET and SPECT Monte Carlo simulator. A suitable acceleration technique can be chosen, depending on the speed and accuracy that is required. A technical description of the integration can be found in the appendix.

The research performed within this PhD thesis resulted in 2 published articles and one under review as a first author in peer reviewed high-impact international journals and in 7 publications as first author in international conference proceedings. Another 5 coauthored peer review journal articles and 9 coauthored international proceedings resulted. The references to these publications are given in a separate section at the end of this dissertation. In addition to these publications, the software developed within this dissertation is currently actively used by most researchers in the Medical Imaging and Signal Processing group (Medisip) for their current and future research. Moreover, part of this software has been incorporated into the public release of GATE and is now successfully used by research labs worldwide.

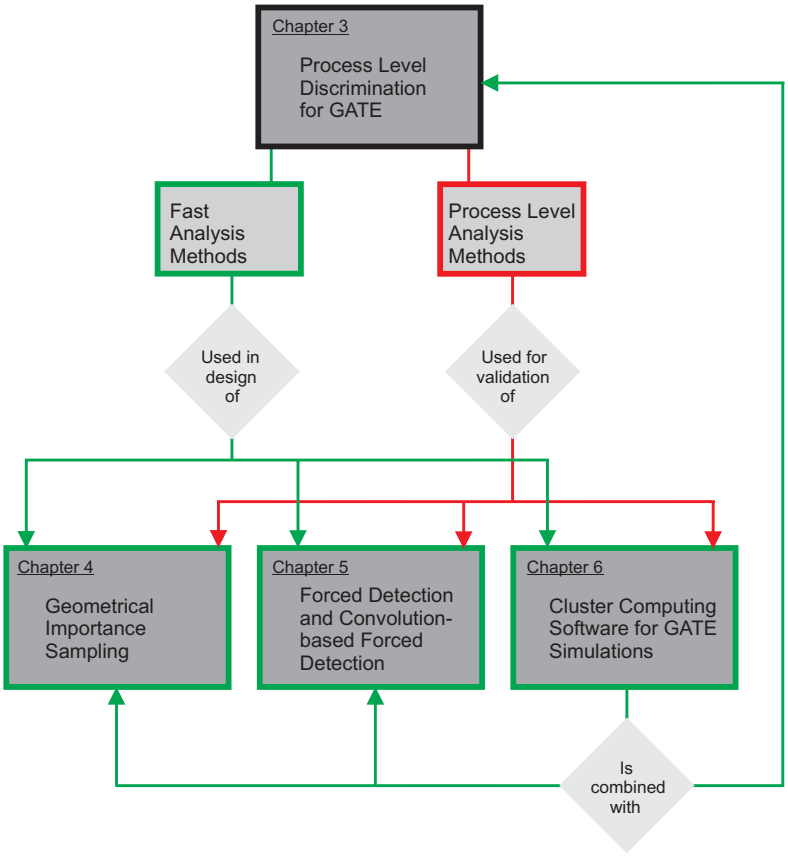


Figure 1.1: Overview of the different chapters and their relation to each other. Red indicates the validation framework. Green indicates the efficiency improvement framework.

Bibliography

[1] B. Cassen, L. Curtis, C. Reed, and R. Libby. *Instrumentation for 131I use in medical studies.* Nucleonics, 9:46–50, 1951.

[2] H.O. Anger. *Scintillation camera*. Rev. Sci. Instr., 29:27–33, 1958.

[3] A.M. Cormack. *Representation of a function by its line integrals, with some radiological applications.* J. Appl. Phys., 34:2722–2727, 1963.

[4] M.E. Phelps, E.J. Hoffman, N.A. Mullani, and M.M. Ter Pogossian. *Application of annihilation coincidence detection of transaxial reconstruction tomography.* J. Nucl. Med., 16:210–215, 1975.

[5] D.E. Kuhl, R.Q. Edwards, and A.R. Ricci et al. *The Mark IV system for radionuclide computed tomography of the brain.* Radiology, 121:405–413, 1976.

[6] A.F. Bielajew. *Fundamentals of the Monte Carlo method*. The University of Michigan Press, 2001.

[7] D. Raeside. *Monte Carlo principles and applications.* Phys. Med. Biol., 21:181–197, 1976.

[8] S. Jan et al. *GATE, a simulation toolkit for PET and SPECT*. Phys. Med. Biol., 49(19):4543–4561, 2004.

2

Monte Carlo Simulations in Nuclear Medicine

2.1 Introduction

This chapter serves as a brief introduction into the research field within which this dissertation on the acceleration of Monte Carlo simulations was started. Firstly, nuclear medical imaging and more specifically Single Photon Emission Computed Tomography (SPECT) and Positron Emission Tomography (PET) are introduced from acquisition to reconstruction. Secondly, the Monte Carlo method and its relevance to nuclear medicine is explained. Thirdly, we investigate the different Monte Carlo simulators for PET and SPECT and clarify our choice for the GEANT4 Application for Tomographic Emission (GATE) (1). Subsequently, a summary of the photon transport model as used in Geometry and Tracking 4 (GEANT4) (2) is presented. Thereafter we focus on Monte Carlo simulations with GATE. The last part of this chapter introduces the drawback of Monte Carlo simulations: the computational cost. More extensive information on this subject can be found in the excellent books and documentation on which this introduction is based (3; 4; 5; 6; 7; 8).

2.2 Nuclear medical imaging

Nuclear medical imaging is a form of medical imaging that produces functional images that reflect biological processes taking place at the cellular and subcellular level. It is based on the tracer principle. A biologically active substance, a pharmaceutical, is selected so that its spatial and temporal distribution throughout the body reflects a certain body function or metabolism. In presence of disease, the pharmaceutical will be distributed around the body differently. In order to form images of the distribution, the pharmaceutical is labeled with gamma-ray-emitting or positron-emitting radionuclides (radiopharmaceuticals or tracers). After administration of the tracer to a patient, an external position-sensitive gamma-ray camera can detect the emitted radiation to form an image of the radionuclide distribution (8). In many cases an increased physiological function will lead to an local increased concentration of the tracer, resulting in a hot-spot on the distribution image. In other cases, a disease may exclude a tracer, resulting in a cold-spot on the distribution image.

Functional imaging is also achievable using other modalities such as magnetic resonance imaging and ultrasound. However, nuclear medicine imaging has the advantage of sensitivity which makes it possible to target molecular processes. In addition, it has the advantage of an ever increasing amount of new radiopharmaceuticals. Examples of the use of nuclear medicine imaging include tissue perfusion, glucose metabolism, the somatostatin receptor status of tumours (see also chapter 3) and gene expression. There are two main classes of nuclear medical imaging: positron annihilation photon imaging, in which two annihilation photons are detected simultaneously and single photon imaging (9; 8; 4).

2.2.1 Positron emission tomography

Pharmaceuticals can also be labeled with neutron-deficient isotopes which are positron emitters. A number of elements that are fundamentally used by the human body can be positron emitters, allowing more scope for radiopharmaceutical design than is possible with single photon emitters. However, production costs are higher since a cyclotron is needed to generate these instable radionuclides. Fluorine 18 (^{18}F) is mostly used to label the fluorodeoxyglucose compound, which is a glucose analog and thus suitable for the visualization of

the glucose metabolism. The emitted positron annihilates with an electron giving rise to two virtually anti-parallel 511 keV photons, corresponding to the conversion of the rest mass of the two particles into energy. The primary advantage of detecting both annihilation photons in time coincidence is the electronic collimation. Detection of two photons is sufficient to determine the line on which the annihilation took place and so no physical collimation is required, allowing higher sensitivity than in single photon imaging. Tomographs appropriate for PET (figure 2.2(b)) typically consist of a series of ring detectors, having a large number of separate blocks coupled to photomultiplier tubes (PMTs) containing a photocathode which converts the optical photons into electrons that are transported over dynodes towards an Anger logic readout for position and energy signal calculation. Depending on the required sensitivity, the individual detectors can be in coincidence only with other opposing detectors in the same ring using lead septa, which is called PET in 2D mode, or also with detectors in other rings (3D mode). The 3D mode increases the sensitivity but also increases the image degradation resulting from the detection of scattered photons and random coincidences. PET application fields in order of importance are oncology, neurology, and cardiology (4).

2.2.2 Single photon imaging

Single photon imaging requires at minimum just one detector fixed in one position to obtain a two-dimensional (2D) projection of a three-dimensional (3D) radiopharmaceutical distribution. A typical imaging system suitable for this task is the gamma camera (figure 2.1) (4). A typical gamma camera consists of a collimator that limits the angle of incidence of the detected photons on the detector to a specific direction, for instance around 90° for a parallel hole collimator. The deposited energy is converted by a scintillating crystal to visible light which travels through the crystal and the light guide toward a set of PMTs (figure 2.1(c)). Single projection imaging can be improved upon by simply rotating the gamma camera around the patient, thus obtaining a series of 2D projections. These can be used to retrieve depth information, i.e. the 3D distribution of the radiopharmaceutical, which is known as Single Photon Emission Computed Tomography (SPECT). The quality of the information (i.e the time dependent reconstructed radiopharmaceutical images) obtained from

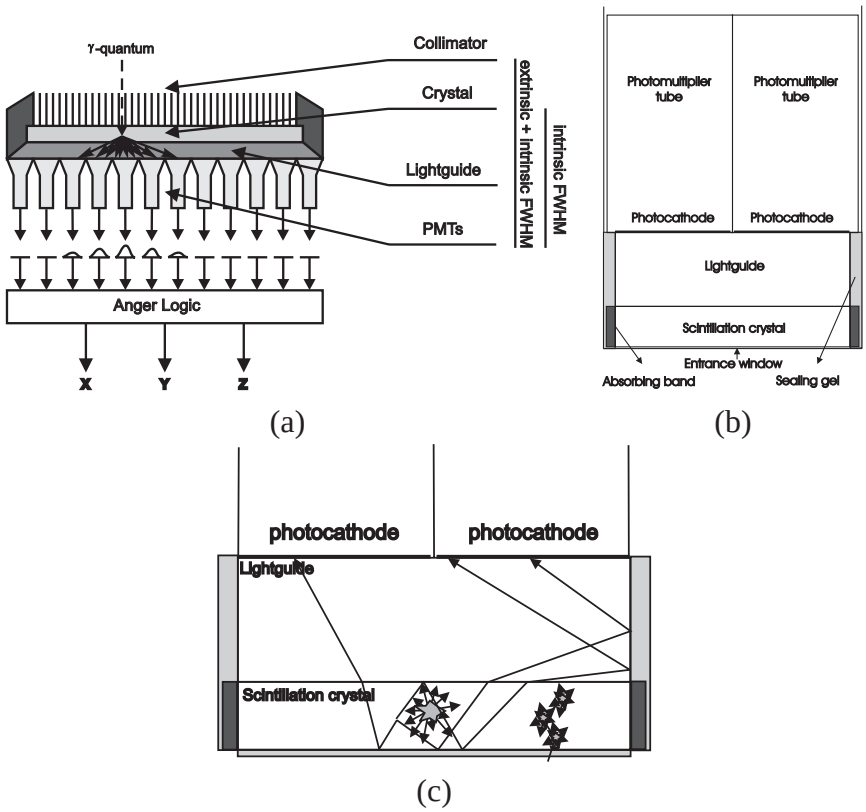


Figure 2.1: Gamma Camera : (a) technical scheme, (b) detailed view, (c) optical photon transport.

SPECT depends considerably on the imaging time, camera sensitivity and position resolution (4). Consequently the temporal and spatial resolutions of SPECT can be improved through the simultaneous use of more than one detector, and most systems currently consist of two or three heads mounted on a single gantry (figure 2.2(a)).

The radionuclide mostly used in SPECT is technetium (^{99m}Tc) which has a half life of 6.03h and which emits photons of 140.5 keV. One of the most important application fields for SPECT is functional cardiac imaging. More specifically, myocardial perfusion imaging is used to diagnose ischemic heart disease (10). Brain imaging is also performed with SPECT. In this case ^{99m}Tc is taken up by brain tis-

(a) (b)

Figure 2.2: (a) SPECT: Philips Irix, (b) PET camera: Philips Allegro.

sue proportionally to the brain blood flow. Blood flow in the brain is related to the brain metabolism, which in turn allows to diagnose the different types of dementia (11). A third important application of SPECT imaging is bone imaging. Many cancers that originally began in the lungs, kidneys, prostate or other organs, may metastasize to the bones (12). Hot-spots on SPECT bone scans may indicate fractures, tumours or infections while cold-spots may indicate poor blood flow to the bone or bone breakdown due to a tumour.

In addition, SPECT with indium (^{111}In) and iodine (^{131}I) is used for dosimetry in targeted radionuclide therapy such as 90Y-DOTATOC (13; 14), for the treatment of primary and secondary neuroendocrine tumours, and 90Y-Zevalin (13) for the treatment of B-cell lymphomas.

2.3 Image reconstruction

A key ingredient for SPECT and PET is image reconstruction, which consists of retrieving the 3D spatial distribution of the radiopharmaceutical from the projection data acquired under different angles. Tomographic reconstruction can be performed either analytically of iteratively:

- analytical reconstruction

  - Filtered BackProjection: analytical inversion of the X-Ray Transform based on the Central Slice Theorem

- deterministic iterative reconstruction

  - ART: Algebraic Reconstruction Technique
  - SIRT: Simultaneous Iterative Reconstruction Technique

- statistical iterative reconstruction

  - MLEM: Maximum Likelihood Expectation Maximization
  - MAP: Maximum a Posteriori

A deterministic model is obtained when the system characteristics are incorporated. Accordingly, the reconstruction problem can be solved analytically. If we assume the acquisition process to be statistical, then a statistical reconstruction technique is mandatory. Often used statistics are the Poisson model for the radioactive decay process and the Gaussian model for the intrinsic features (detector response) of the acquisition. The main advantage of a statistical model is that a smaller variance results in the reconstructed image if the appropriate models are used. Iterative reconstruction techniques are capable of incorporating image degrading effects in the reconstruction algorithm in order to improve the reconstructed image quality (4). The reconstructions in this dissertation are based on MLEM (15), which is a basic statistical iterative reconstruction algorithm that incorporates the Poisson statistics of the radioactive decay in the reconstruction.

2.4 Monte Carlo principle

The Monte Carlo method is a numerical solution to a problem that models objects interacting with other objects or their environment based upon simple object-object or object-environment relationships. It represents an attempt to model nature through direct simulation of the essential dynamics of the system in question. In this sense the

Monte Carlo method is essentially simple in its approach: a solution to a macroscopic system through simulation of its microscopic interactions. A stochastic solution is determined by random sampling of the relationships, or the microscopic interactions, until the result converges (3). We will first discuss random number generation and different sampling methods. Subsequently the Monte Carlo principle is explained with a simple example. This section is then closed with a discussion of scoring for Monte Carlo simulations.

2.4.1 Random numbers

Random number generators (RNG), or rather pseudo-random number generators are at the heart of Monte Carlo simulations. RNGs should provide a stream of seemingly uncorrelated random numbers in order to imitate the stochastic nature of particle interactions. Random number generation is an area of active research however and it is imperative for Monte Carlo simulations requiring long periods that an unbiased RNG is used. RNGs play an important role in chapter 6 of this dissertation where independent streams with long periods are required for parallelization.

2.4.1.1 Requirements

The requirements of RNGs are (16):

- Randomness: the RNG should pass relevant statistical tests.
- Long period: current Monte Carlo simulations regularly exceed the specifications of classical RNGs.
- Efficiency: Monte Carlo simulations are computationally expensive. Random number generation amounts to a significant portion of the simulation time.
- Repeatability and restarting: under the same initial conditions, a sequence should repeat itself. Additionally, it should be possible to restart a sequence from a certain point. Without this functionality it would be impossible to debug a failed Monte Carlo simulation.

- Independent sequences: the same generator should be able to provide independent streams of random numbers simultaneously.

2.4.1.2 Multiplicative linear congruential random number generators (MLCRNG)

MLCRNGs are the classical way to generate pseudo-random numbers. Subsequent seeds are obtained as in equation 2.1. Equation 2.2 describes an example of such a RNG, that is based on the 2-complement integer operation for a binary computer with 32-bit word length. It generates a 32-bit string of random bits from the previous step. At multiplication, the higher order bits are lost, leaving pseudo-random lower-order bits. In the equation, a and b should be chosen carefully (17). When b is chosen as an odd number, a period of 2^{32} can be obtained. Current Monte Carlo simulations can easily exhaust this period however and the generation of independent streams is not straightforward. In addition, this type of RNG may lead to biased results as shown in the classic paper of Marsaglia (18). The GEANT4 RNDM module implements a MLCRNG. One way to extend the period for these RNGs is to use longer integers. Using a 64-bit integer, a period of 2^{64} is possible as in the GEANT4 module RANECU.

$$r_{n+1} = a_0 r_n + a_1 r_{n-1} + ... + a_j r_{n-j} + b \ mod \ P \tag{2.1}$$

$$r_{n+1} = a r_n + b \ \ mod \ 2^{32} \tag{2.2}$$

2.4.1.3 Lagged Fibonacci generators

The need for longer periods has led to the introduction of lagged Fibonacci generators. These are initialized with t integers r_1, r_2,...,r_t as shown in equation 2.3, where $t < s$ and $\otimes$ can be either $+, -, \times$ or the exclusive-or operation $\otimes$. With P chosen as 2^u, a period of around 2^{t+u-1} can be obtained. However, not all lagged Fibonacci generators pass statistical randomness tests.

$$r_i = (r_{i-t} \otimes r_{i-s}) \ mod \ P \tag{2.3}$$

2.4.1.4 Universal generators

Universal generators as proposed by Marsaglia in (16) are a combination of two different approaches in order to obtain a very long period and excellent randomness behaviour. The RANMAR generator which is the standard RNG for GEANT4 employs this technique. The first part of RANMAR is a lagged Fibonacci generator as in equation 2.4. This is then combined with a simple arithmetic sequence for the prime modulus $2^{24}-3 = 16777213$ as in equation 2.5. Here c is chosen as $7654321/16777216$ and d as $16777213/16777216$. The combination of these two parts then defines the random number r_i as shown in equation 2.6. RANMAR has a period of around 2^{144}, excellent statistical behaviour and the possibility to generate independent streams. Two simulations starting from the same random seed would generate exactly the same results. Independent streams are therefore a prerequisite for parallelized Monte Carlo. These abilities made RANMAR widely used for long Monte Carlo simulations. It is also the generator used for the parallelization of GATE in chapter 6.

$$\begin{aligned} k_i &= k_{i-97} - k_{i-33}, \quad if \quad k_{i-97} \geq k_{i-33} \\ k_i &= k_{i-97} - k_{i-33} + 1, \quad otherwise \end{aligned} \tag{2.4}$$

$$\begin{aligned} l_i &= l_i - c, \quad if \quad l_i \geq c \\ l_i &= l_i - c + d, \quad otherwise \end{aligned} \tag{2.5}$$

$$\begin{aligned} r_i &= k_i - l_i, \quad if \quad k_i \geq l_i \\ r_i &= k_i - l_i + 1, \quad otherwise \end{aligned} \tag{2.6}$$

2.4.1.5 Very long period generators

Very long period generators will be required for future Monte Carlo simulations. These include RNGs based on the ”add-with-carry” and ”subtract-with-borrow” methods as proposed by Marsaglia in (19) and MLCRNGs based on 64-bit integers. The RNGs originally proposed by Marsaglia were subsequently improved in (20) and (21).

However, at their highest setting these RNGs are relatively slow. An entirely new class of RNGs was proposed in (22). The Mersenne twister provides fast generation of very high quality pseudo-random numbers, and rectified many flaws found in older RNGs. With a period of around 2^{19937} it suitable to become the RNG of choice for future Monte Carlo simulators.

2.4.2 Sampling methods

To obtain a stochastic variable that follows a particular pdf, $pdf(x)$, three different sampling methods can be used (4).

2.4.2.1 The direct method:

A cumulative distribution function $F(x) = P(X \leq x)$ is constructed from the integral of $pdf(x)$ over the interval $[a,x]$ according to:

$$F(x) = \int_a^x pdf(x')dx'. \tag{2.7}$$

The variable x is then sampled by replacing $F(x)$ in (2.7) with a uniformly distributed random number in the range of $[0,1]$ and solving for x. This sampling method is used if the inverse of $F(x)$ can be easily calculated.

2.4.2.2 The rejection method:

In principle, the invertible cumulative probability distribution function method can always be used. However,the inversion may be impractical to calculate. In this case, the rejection technique may offer a solution as follows (3).

1. Scale the probability distribution function, shown in figure 2.3a, by its maximum value obtaining a new distribution function, $pdf(x) = p(x)/p(x_{max})$. The new distribution function has a maximum value of 1 which occurs at $x = x_{max}$ as shown in figure 2.3b. This method works only if the probability distribution function is not infinite anywhere and if it is not difficult to determine the location of the maximum value.

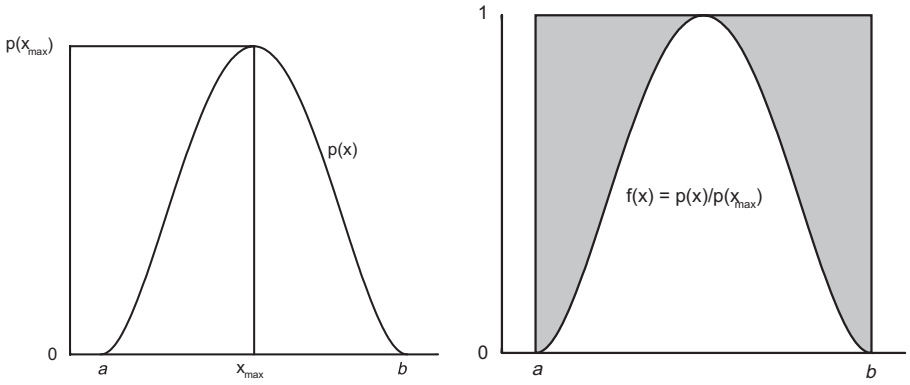


Figure 2.3: (a) a simple probability distribution, (b) The probability distribution scaled for the rejection technique. Images courtesy of (3)

2. Choose a random number, r1, uniform in the range $[0,1]$ and use it to obtain an x which is uniform in the probability distribution function's range $[a,b]$ as follows: $x = a + (b-a)r1$

3. Choose a second random number r2. If $r2 < p(x)/p(x_{max})$ then accept x, else, reject it and go back to step 2. In the case of figure 2.3b, a rejected x belongs to the shaded area.

The efficiency of the rejection technique is defined as in equation 2.8. This is the ratio of the expected number of random numbers pairs that are accepted to the total number of pairs employed. Whether or not to use the rejection method depends on the probability distribution at hand. If the inverse cumulative distribution is computationally expensive, then it might be worthwhile to "waste" a few random numbers with the rejection method instead.

$$efficiency = \frac{1}{p(x_{max})} \int_a^b dx p(x) \tag{2.8}$$

2.4.2.3 Mixed methods:

When the previous two methods are not applicable, a mixture of the two methods above is used. The $f(x)$ is written as the product of two probability distribution functions $m(x).r(x)$ with $r(x)$ acting as the rejection function defined in the previous section. One determines an x value using the distribution function method on $m(x)$ and applies the rejection method with that x to $r(x)$.

2.4.3 Scoring

Monte Carlo results represent an average of the contributions from many histories by accumulating relevant physical quantities into tallies or scores. It is important to determine the statistical error associated with the results (3). To this end, we need an estimate of the expected value, an estimate of the variance of a sample and of the standard deviation of the mean.

2.4.3.1 Estimate of the expected value

Suppose that $pdf(x)$ is a continuous probability density function over a range (a,b) which can be used to select a random walk to score x to the tally being estimated. The true mean, $\bar{x}$, is then the expected value of x:

$$\bar{x} = \int_a^b x \; pdf(x)dx \tag{2.9}$$

with

$$pdf(x) \geq 0 \; over \; (a,b) \; and \; \int_a^b pdf(x)dx = 1 \tag{2.10}$$

In the case of a discrete distribution with K possible values with probability p_i, this becomes:

$$\bar{x} = \sum_{i=1}^{K} p_i x_i \tag{2.11}$$

with

$$all \; p_i \geq 0 \; and \; \sum_{i=1}^{K} p_i = 1 \tag{2.12}$$

Unfortunately, $pdf(x)$ is mostly not known explicitly. It must therefore be implicitly sampled by a Monte Carlo random walk. The expected value is thus estimated by the sample mean $\hat{x}$, as in equation 2.13. Here, N is the number of simulated histories and x_i the value for the score associated with the i'th history (23). The relationship between the expected value and the sample mean is given by the Strong Law of Large Numbers (24), which states that if the expected value is finite, $\hat{x}$ will approach $\bar{x}$ in the limit as N approaches infinity.

$$\hat{x} = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{2.13}$$

2.4.3.2 Estimate of the variance of the sample

The variance, σ^2, is the expected value of the squared error. In the case of the continuous distribution we get:

$$\sigma^2 = \int_a^b pdf(x)(x-\bar{x})^2 dx \tag{2.14}$$

In the case of the discrete distribution this becomes:

$$\sigma^2 = \sum_{i=1}^{K} p_i (x_i - \bar{x})^2 \tag{2.15}$$

When sampling with Monte Carlo, the variance S_x^2 associated with the distribution of the x_i can be estimated with equation 2.16:

$$S_x^2 = \frac{1}{N-1}\sum_{i=1}^{N} (x_i - \hat{x})^2 \tag{2.16}$$

Equation 2.16 can be written in a more useful form as:

$$S_x^2 = \frac{N}{N-1}\left(\frac{\sum_{i=1}^{N} x_i^2}{N} - \left(\frac{\sum_{i=1}^{N} x_i}{N}\right)^2\right) \approx \hat{x^2} - \hat{x}^2 \tag{2.17}$$

Finally, the error in $\hat{x}$ can be determined by estimating the standard variance of the mean as in equation 2.18. The final result can then be reported as $x = \hat{x} \pm S_{\hat{x}}$.

$$S_{\hat{x}}^2 = \frac{S_x^2}{N} \tag{2.18}$$

2.4.4 Monte Carlo simulation example

A general approach to estimate a physically measurable variable therefore breaks down into two steps:

1. Set up an numerical experiment that would have an expected value corresponding to the desired physically measurable value, $\overline{x}$

2. Sample the numerical experiment in order to obtain an estimate of the expected value, $\hat{x}$

The numerical experiment reflects the physical situation which in itself is stochastic. In other words, it is a perfect analog of the physical situation. This type of Monte Carlo simulation is therefore called an analog simulation.

A simple example of an analog simulation is the numerical estimation of π. Figure 2.4 shows that the ratio of the area of the circle to the area of the square is $\pi r^2/(2r)^2 = \pi/4$. If the radius is chosen as 1 with the origin at the centre of the circle, then the area of the circle is exactly π. The simulation is then as follows:

1. A random point in the square is chosen by choosing a random u and v between -1 and 1

2. The resulting point is scored: a hit scores 4 when the point is inside the circle ($u^2 + v^2 \leq 1$) or 0 if it is outside

3. The first two steps are repeated N times.

The final estimate of π is then given by $\hat{\pi} = \sum_{i=1}^{N} x_i/N$, with x_i the score of trial i.

2.5 Relevance of Monte Carlo to nuclear medicine

Monte Carlo modeling is the method of choice for all applications where measurements are not feasible or where analytic models are not available due to the complex nature of the problem (25; 26). In addition, such modeling is a practical approach in nuclear medical imaging in four important application fields (27; 28; 4).

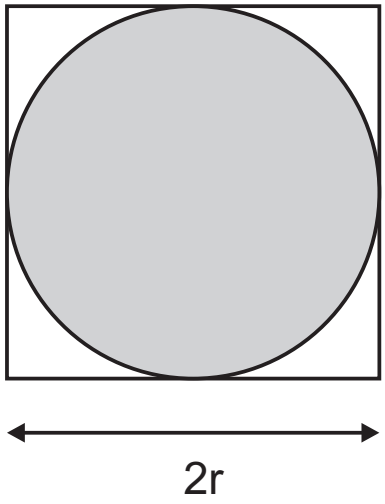


Figure 2.4: Estimation of π by the estimation of the area of a circle enclosed by a rectangle.

2.5.1 Studying detector design

Monte Carlo methods can assist in the development of new collimator and detector designs, evaluation of new electronics, aso (4). Monte Carlo simulations have been extensively used to analyze the performance of new long bore parallel collimators as well as rotating slit collimators (29), fan beam (30), cone beam and pinhole collimators (31). For the latter, simulations have been performed with various aperture span angles, different hole sizes and various materials to evaluate the penetration (32). Monte Carlo methods also play an important role in new system design, for instance in the research field of solid state detectors with improved energy resolution and low-noise electronics (33; 34; 35). Similarly, in the field of PET, Monte Carlo techniques have been used to study the performance of interplane septa with varying constitution, thickness and geometry, to compare single to true coincidence event ratios in single-slice, multi-slice and open collimator 3D configurations and to assess the effect of collimation on scatter fraction (e.g. (36; 37; 38)). Monte Carlo simulations of detector responses and efficiencies are of key importance since the scintillation crystal is the critical component in emission tomography (39). Simulations also prove their useful-

ness in Time-of-Flight PET (TOF) design since they can be used to test several detector crystals and to simulate the influence of timing resolution on variance reduction (e.g. (40; 41; 42)). Monte Carlo methods can also assist in the design of new detectors with depth-of-interaction (DOI) information by simulating multilayer crystals and by estimating the gain in reconstructed resolution through incorporation of DOI information (43; 44).

2.5.2 Analysing quantification issues

The presence of the scatter, attenuation and partial volume effects in nuclear medicine images limits the accuracy of activity estimates (i.e. what is called quantification) (4). Each of these degradations has a particular impact (45; 46; 47). Scatter does not produce major artefacts comparable to those caused by attenuation but reduces contrast by introducing a low-frequency blur in the image. Moreover, the impact of scatter and attenuation generally depends on the photon energy, camera energy resolution and energy window settings, as well as the object size and shape, detector geometry and the source distribution. Many of these parameters are non-stationary which implies a potential difficulty when developing proper scatter and attenuation correction techniques. Monte Carlo calculations have been found to be powerful tools for an in depth analysis since the user has the ability to separate the detected photons according to their interaction history: primary (i.e. unscattered) events, scattered events, contribution of downscattered events, scatter in patient or detector, etc. (48). Monte Carlo modeling thus enables a detailed investigation of the spatial and energy distribution of Compton scatter: energy spectra, point-spread functions and scatter fractions can be simulated (49). This is impossible to perform using present experimental techniques, even with very good energy resolution detectors. Another important application field for Monte Carlo simulations in quantification is resolution recovery. For instance in SPECT, distance dependent collimator response can be subdivided in a geometric component, collimator scatter and septal penetration. Monte Carlo simulations can be used to determine the relative contribution of each of these components in specific imaging configurations (50).

In PET, the penetration of annihilation photons into the detector material before interaction is a statistical process which leads to significant displacement and anisotropy of the point spread func-

tion (51; 52). Compensation for crystal penetration is thus one of the most important issue in order to recover the spatial resolution in PET. Theoretical models for that anisotrope spatial resolution can only be verified by Monte Carlo simulations since those DOI-dependent experiments are very hard to perform with current devices.

Finally, Monte Carlo analysis has been proven useful in evaluating motion artefacts. If the simulator is capable of using explicit timing information to synchronize decay with detector and patient movement then degradations such as cardiac or respiratory motions can be studied (53).

2.5.3 Correction methods for image degradations

The image degrading factors (as described in the previous section) can have a large impact on quantitative accuracy and on clinical diagnosis and ideally all have to be corrected for to achieve optimal clinical imaging (4). Traditionally this was done by preprocessing the projection data or by postprocessing the reconstructed images. Monte Carlo simulated datasets are often used to evaluate the accuracy of these correction methods. For such evaluation studies, the simulated datasets must have the appropriate statistical properties and include all detector specific imperfections. For instance when evaluating a scatter correction method, Monte Carlo simulations can be used to generate scatter-free projections. Images reconstructed from scatter-free projections constitute the gold standard for any scatter correction method. Nowadays, Monte Carlo simulations often form the basis of the aforementioned correction techniques since iterative reconstruction algorithms are able to correct for image degradations by simulating the effect during the reconstruction. The simulations can be based on Monte Carlo methods or on a combination of Monte Carlo and analytical methods. These SPECT simulators face a considerable challenge; they have to balance accuracy against efficiency. A large variety of simulators have been proposed (54; 55; 56; 57; 58). Other groups achieve similar goals by simulating and store the system matrix (59; 60). Reconstruction methods based on Monte Carlo simulators have been shown to improve contrast-to-noise (61; 62; 63) and lesion detectability (64; 65; 66). Finally Monte Carlo generated datasets are also used in evaluation or comparison of different reconstruction schemes, for instance to evaluate the robustness of an algorithm to parameter variations.

2.5.4 Detection tasks using Monte Carlo simulations

Receiver Operating Characteristic (ROC) analysis is considered as the most reliable method for evaluating the diagnostic ability of medical imaging techniques (4). It provides a measure of the diagnostic performance of an imaging modality by plotting the sensitivity versus the specificity for a wide and continuous range of decision criteria. These studies require many images (typically hundreds) so that statistical analysis of the detection performance can be performed. Observer studies based on Monte Carlo based projection data can answer a large range of problems from the simple comparison between analytical and statistical reconstruction (67) to complex studies on the benefits of anatomical a priori information in the reconstruction (68). Using Monte Carlo simulations to generate the datasets is therefore only feasible if efficient codes are available. Also, sharing these datasets encourages and facilitates evaluation studies. In addition, evaluating data processing tools developed by different groups worldwide using the same datasets removes the evaluation biases introduced by the use of different data. Therefore, some efforts have recently been accomplished to propose publicly available databases of Monte Carlo simulated data, dedicated to evaluation studies (69; 70).

2.6 Monte Carlo simulators for PET and SPECT

Some Monte Carlo simulation codes dedicated to the simulation of SPECT and PET have been developed, such as SimSET (71) enabling SPECT and PET modeling, SIMIND (72) and SimSPECT (73) for SPECT, PETSIM (74), Eidolon (75) and SORTEO (76) for PET (4).

Such dedicated packages are very powerful but are often not detailed and flexible enough to enable realistic simulations of emission tomography detector geometries. Moreover, most of them do not account for time explicitly, which limits their use for modeling time dependent processes such as decay, tracer kinetics, patient and bed motion, dead time or detector orbits. To ensure high flexibility in simulation design, general purpose nuclear physics codes such as EGS4 (77), GEANT4 (2; 6) or MCNP (23) have also been used for

SPECT and PET simulations. The main advantages of these general purpose codes are that they are widely used and extensively tested, they can be regarded as long-term existent as well as supported, and they are continuously evolving and therefore use the best of current hardware and software capabilities. There are fewer limitations on their possible applications than for dedicated codes, and fewer simplifying assumptions are made. For example, the processes taking place in the collimator can be thoroughly simulated whereas dedicated codes often use a parametric model. Also, time-dependent processes can be simulated and their non-specific design makes them appropriate to implement non conventional SPECT and PET cameras (28). End of the nineties, 14 codes were used for Monte Carlo simulations in SPECT and PET, among which 10 were not used by another group than those who developed the code, and 4 were publicly released or available from the authors. In early 2000, 15 codes were used among which 8 were home-made and 7 were publicly released or available from authors (78). This suggests that up to recently, there was no code that was considered as a standard for Monte Carlo simulations in emission tomography. Based on this observation, a new code, GATE (79; 80; 1), has recently been designed as an upper layer of the GEANT4 nuclear physics code (2; 6) tuned for simulating SPECT and PET acquisitions. GATE thus takes advantage of all GEANT4 features, including a great variety of physics models, basic event timing information, geometry modeling tools and visualization tools. It is written in an object-oriented language that ensures high modularity. In addition, GATE uses a scripting language making it easy for the user to design a SPECT or PET simulation. GATE has been developed as a collaboration effort of about 10 labs worldwide (the OpenGATE collaboration), and was publicly released in May 2004. It is now shared by a large research community, while long-term support and maintenance is still ensured by the OpenGATE collaboration. During 2006 already more than 650 individuals had subscribed to the GATE community and at the 2005 IEEE Medical Imaging Conference (the most relevant conference in this domain), GATE and/or GEANT4 were used as tool of excellence in 33% of all papers involving Monte Carlo simulations while this number was 13% for SimSET and 7% for SIMIND respectively (78). It is therefore expected that GATE will become the code most frequently used for SPECT and PET simulations. Most of the simulations and all of the implementation in this dissertation is based

on GATE. The next sections will therefore describe the GEANT4 photon transport model used by GATE and subsequently the basic features of GATE.

2.7 GEANT4 Monte Carlo model

In the case of Monte Carlo simulations in nuclear medicine, some a priori information about the occurring physics processes is needed. This information is expressed in the form of probability density functions (pdfs). When simulating photon interactions, the partial and total cross section data (based on the material constitution) represent such information used to calculate the path length and the type of interaction. During a simulation these pdfs are sampled by predefined rules using randomly generated numbers. The energy of a photon can be dissipated along its path or the photon can penetrate all scattering and attenuating media to reach the detector where a new pdf-sampling decides whether it should be accounted for in the scoring region or whether it should be discarded (81). Firstly, photon interactions in the GEANT4 low energy physics model are described. Subsequently, the transportation of particles in GEANT4 is presented.

2.7.1 Photon interaction in GEANT4

The electromagnetic interactions in GATE are based on GEANT4. The physics package manages electrons, positrons, γ-rays, X-rays, optical photons, muons, hadrons, and ions. When using the low energy extension, the modeling of photons and electrons is extended and verified down to 250 eV, includes Rayleigh scattering and covers elements with atomic number between 1 and 99 (6). Because atomic shell structure is more important in most cases at low energies than it is at higher energies, the low energy processes make direct use of shell cross section data instead of parameterizations (6). Photon transport and interaction mechanisms are generally based on four processes: pair production, Compton scatter, Rayleigh scatter and the photoelectric effect. In the case of nuclear medicine only Compton scatter, the photoelectric effect and, to a lesser extent, Rayleigh scatter are of importance. All processes involve two distinct phases: the calculation and use of the total cross sections and the generation

of the final state. All of the following is based on GEANT4 low energy physics (6).

2.7.1.1 Calculation of the total cross sections

The energy dependence of the total cross section is derived for each process from data libraries. Since the libraries provide cross sections for a set of discrete incident energies, the total cross section at a given energy, E, is obtained by interpolation according to the formula:

$$\log(\sigma(E)) = \frac{\log(\sigma_1)\log(E_2/E) + \log(\sigma_2)log(E/E_1)}{\log(E_2/E_1)} \qquad (2.19)$$

with E_1 and E_2 respectively the closest lower and higher energy for which data (σ_1 and σ_2) are available. For a particle of energy E, the mean free path for a given process is then calculated as:

$$\lambda = \frac{1}{\sum_i \sigma_i(E) n_i} \qquad (2.20)$$

with σ_i the microscopic integrated cross section of the processes considered at energy E, n_i the atomic density of the ith element contributing to the particular material composition and $\sum_i$ the sum over all the elements of which the material is composed (6).

2.7.1.2 The photoelectric effect

The total photoelectric cross section for a given energy, E, is calculated using discretized libraries by means of interpolation. The incident photon is absorbed and an electron is emitted in the same direction as the incident photon. The kinetic energy of the electron is defined as the difference between the energy of the original photon and the binding energy of the electron. The subshell of emission is again randomly sampled, thereby using cross section data for all subshells.

2.7.1.3 Compton scatter

In the case of unpolarized radiation and when neglecting the binding energy of an electron, the quantum mechanical Klein-Nishina differential cross section per electron for scattering into a unit solid angle

at scatter angle θ is given as (82):

$$\frac{d\sigma_c}{d\Omega} = \frac{r_e^2}{2}\left(\frac{E_1}{E_0}\right)^2 \left[\frac{E_1}{E_0} + \frac{E_0}{E_1} - sin^2\theta\right] \tag{2.21}$$

with

r_e = classical electron radius
E_0 = energy of the incident photon
E_1 = energy of the scattered photon
θ = scatter angle

This can be rewritten as the differential cross section per atom (6):

$$\frac{d\sigma_c}{d\varepsilon} = \pi r_e^2 \frac{m_e c^2}{E_0} Z\left[\frac{1}{\varepsilon} + \varepsilon\right]\left[\frac{1-\varepsilon\sin^2\theta}{1+\varepsilon^2}\right] \tag{2.22}$$

where

$m_e c^2$ = electron mass
$\varepsilon = \frac{E_1}{E_0}$
Z = atomic number

Assuming an elastic collision, the scattering angle θ is defined by the Compton formula:

$$E_1 = E_0 \frac{m_e c^2}{m_e c^2 + E_0(1-\cos\theta)}. \tag{2.23}$$

The value of ε corresponding to the minimum photon energy (backscatter) is given by:

$$\varepsilon_0 = \frac{m_e c^2}{m_e c^2 + 2E_0}, \tag{2.24}$$

hence $\varepsilon \in [\varepsilon_0, 1]$. For low energy incident photons the atomic form factor or scattering function (SF) is taken into account. The angular and energy distribution of the incoherently scattered photon is then given by the product of the Klein-Nishina formula $\frac{d\sigma_c}{d\varepsilon}$ and the scattering function, $SF(q)$:

$$\frac{d\sigma_{c,SF}}{d\varepsilon} = \frac{d\sigma_c}{d\varepsilon} \times SF(q) \tag{2.25}$$

with the momentum transfer given by $q = E_0\, sin^2(\theta/2)$. The effect of the scattering function is to multiply the Klein-Nishina distribution

with the square of the energy of the scattered photon at very low energies and to suppress forward angles. At high energies the effect can be ignored because the scattering function becomes isotropic. One may state that:

$$\begin{aligned}\frac{d\sigma_{c,SF}}{d\varepsilon} &\approx [\frac{1}{\varepsilon}+\varepsilon][\frac{1-\varepsilon\sin^2\theta}{1+\varepsilon^2}]\ SF(q)\\ &= m(\varepsilon).r(\varepsilon)\\ &= [\alpha_1 m_1(\varepsilon)+\alpha_2 m_2(\varepsilon)].r(\varepsilon)\end{aligned} \tag{2.26}$$

where

$$\begin{aligned}\alpha_1 &= ln(\frac{1}{\varepsilon_0})\\ m_1(\varepsilon) &= \frac{1}{\alpha_1\varepsilon}\\ \alpha_2 &= \frac{(1-\varepsilon_0^2)}{2}\\ m_2(\varepsilon) &= \frac{\varepsilon}{\alpha_2}.\end{aligned} \tag{2.27}$$

m_1 and m_2 are probability density functions defined on the interval $[\varepsilon_0, 1]$, and $r(\varepsilon)$ is set to:

$$r(\varepsilon) = [1-\frac{\varepsilon\sin^2\theta}{1+\varepsilon^2}]\ SF(q) \tag{2.28}$$

being the rejection function $\forall\ \varepsilon \in [\varepsilon_0, 1] \Rightarrow 0 < r(\varepsilon) \leq Z$. Given a set of 3 random numbers r_0,r_1,r_2 uniformly distributed on the interval $[0, 1]$, the sampling procedure for ε is the following (83):

1. decide whether to sample from $m_1(\varepsilon)$ or $m_2(\varepsilon)$: if $r_1 < \frac{\alpha_1}{(\alpha_1+\alpha_2)}$ select $m_1(\varepsilon)$ otherwise select $m_2(\varepsilon)$,

2. sample ε from the distributions corresponding to m_1 or m_2:
   for m_1 : ε=$\varepsilon_0{}^{r_1}$
   for m_2 : ε^2=ε_0^2+$(1-\varepsilon_0^2)r_1$,

3. calculate $\sin^2\theta$=$t(2-t)$ where $t \equiv (1-\cos\theta)$=$\frac{m_e c^2(1-\varepsilon)}{E_0\varepsilon}$

4. test the rejection function: if $r(\varepsilon) \geq r_2\, Z$ accept ε, otherwise go to step 1.

The polar angle θ is deduced from the sampled ε value and in the azimuthal direction the angular distribution of both the scattered photon and the recoil electron are considered to be isotropic (6). The kinetic energy, E_{el}, and momentum vector, $\vec{V}_{el}$, of the recoil electron are then defined as:

$$E_{el} = E_0 - E_1 \text{ and } \vec{V}_{el} = \vec{V}_{\gamma 0} - \vec{V}_{\gamma 1} \tag{2.29}$$

with $\vec{V}_{\gamma 0}$ and $\vec{V}_{\gamma 1}$ the momentum vectors of the original and scattered photon respectively.

2.7.1.4 Rayleigh scatter

The coherent scattered photon angle θ is sampled according to the distribution obtained from the product of the energy independent Thomson formula and the square of Hubbel's form factor $FF^2(q)$ (84). This is shown in equation 2.30, where $q = 2E\sin(\theta/2)$ is the momentum transfer.

$$\frac{d\sigma_{r,FF^2}}{d\Omega} = \frac{{r_e}^2}{2}[1+\cos^2\theta] \times FF^2(q) \tag{2.30}$$

Form factors introduce a dependency on the initial energy E of the photon that is not taken into account in the Thomson formula. At low energies form factors are isotropic and do not affect the angular distribution, at higher energies they are forward peaked. The sampling procedure is then as follows (6):

1. $\cos\theta$ is chosen from a uniform distribution between -1 and 1.
2. The form factor FF is extracted from the data table for the considered element, using logarithmic data interpolation, for $q = 2E\sin(\theta/2)$.
3. If the value obtained for $\frac{d\sigma_{r,FF^2}}{d\Omega}$ is larger than a random number uniformly distributed between 0 and Z^2, the procedure is repeated from step 1, otherwise θ is taken as the photon scattering angle with respect to its incident direction.
4. The azimuthal direction of the scattered photon is chosen at random.

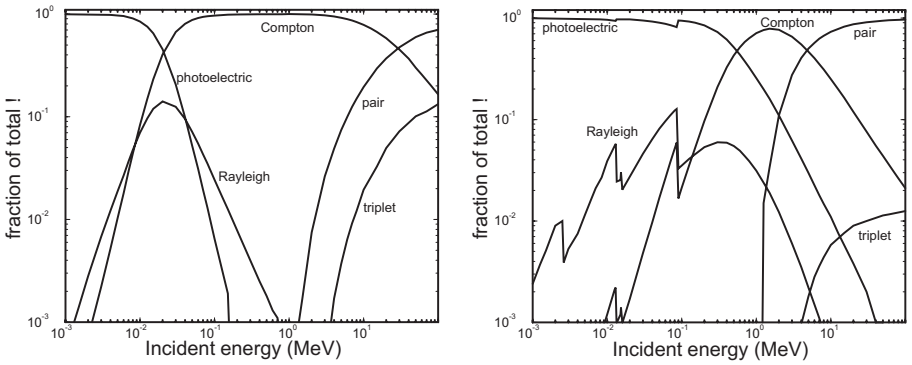


Figure 2.5: (a) Photon cross section in carbon, (b) Photon cross section in lead. Images courtesy of (3)

2.7.1.5 Relative importance

The contributions of the processes to the total photon cross section for carbon (low-Z material, present in human tissue) and lead (high-Z material, present in collimator material) are shown in figure 2.5a and b respectively. In the case of carbon the photoelectric effect is dominant below 20 keV, while pair production is dominant above 30 MeV which is out of the range of nuclear medicine. In between those regions Compton scatter is the dominant process. In the case of lead, there dominant region for Compton scatter is much smaller, ranging from 700 keV to 4 MeV (3).

2.7.2 GEANT4 Particle transportation

In GEANT4, the world geometry defined by the simulation setup basically consists of analytical or parameterized volumes with material properties. The source definition determines the number of decays, or events in GEANT4, that will be simulated over time. Every particle generated during a single event is represented by a track and is sequentially numbered by a trackID and placed on a stack. A track holds all the transient information concerning that particle as it is transported from one point to the next. During tracking, a particle is transported one step at a time through the geometry from interac-

tion point to interaction point until it either ceases to exist or until it leaves the ”world”. The following describes the differential approach to particle transport, which is used in most simulation codes. In this approach the continuous energy loss can limit the step size since the cross sections are energy dependent. It is generally assumed that the step is chosen small enough so that the particle cross sections remains approximately constant during each step. In principle a very small step size should be used in order to obtain an accurate simulation, but the computing time increases as the step size decreases (6). We first describe two critical points in particle tracking: the sampling of the number of interaction lengths and the evaluation of the distance to the next interaction point. Using these principles we then describe the tracking algorithm used by GEANT4.

2.7.2.1 Sampling the number of interaction lengths

The number of interaction lengths that the particle is going to travel, before undergoing each of the possible processes must be sampled in order to determine the next interaction point. The mean free path of a particle for a given process, λ, depends on the material and cannot be used directly to sample the probability of an interaction in a heterogeneous detector. The probability an interaction in a certain distance dx is given as $P(interaction\ in\ distance\ dx) = dx/\lambda$. Therefore, the number of mean free paths which a particle travels is given by equation 2.31.

$$N_\lambda = \int dx/\lambda(x) \tag{2.31}$$

The latter equation is independent of the material traversed. If N_R is a random variable denoting the number of mean free paths from a given point until the point of interaction, it can be shown that N_R has the distribution function:

$$P(N_R < N_\lambda) = 1 - e^{-N_\lambda} \tag{2.32}$$

The number of mean free paths or interaction lengths can then be sampled, using the direct sampling method from equation 2.33, with η a random number uniformly distributed in the range $(0,1)$.

$$N_\lambda = -log(\eta) \tag{2.33}$$

2.7.2.2 Evaluation of the distance to the interaction point

The number of interaction lengths remaining to travel before each of the possible processes is multiplied by the inverse of the macroscopic cross section for that process in the current material. This provides the distances that the particle has to travel before each of the processes occurs in the current medium. The minimum among these numbers is the step over which the particle will be transported and determines which process is selected, unless the step is limited by geometric boundaries or thresholds such as energy cuts and maximum step limits. In the case of boundary crossing into a new medium, the number of mean free paths must be adjusted as in equation 2.34, with Δx the step length, and tracking will be continued in the new medium until an interaction point is reached.

$$N_{\lambda}^{'} = N_{\lambda} - \Delta x/\lambda(x) \tag{2.34}$$

2.7.2.3 Tracking algorithm

The following describes the tracking algorithm used by GEANT4:

1. Sample the number of interaction lengths that the particle is going to travel for each of the possible processes.

2. Evaluate the distance to the interaction point and determine the step and the selected process.

3. Transport the particle along a straight line in case there is no magnetic field or in case of a neutral particle. Otherwise, transport the particle over a helicoidal path in case of charged particles in a magnetic field.

4. The energy of the particle is updated in case of continuous energy loss for charged particles in a medium.

5. The final state of the interaction is generated for the selected process. This includes the generation of secondary particles.

6. If the original particle survived the interaction, return to point (1).

2.8 Monte Carlo simulations with GATE

2.8.1 Basic features

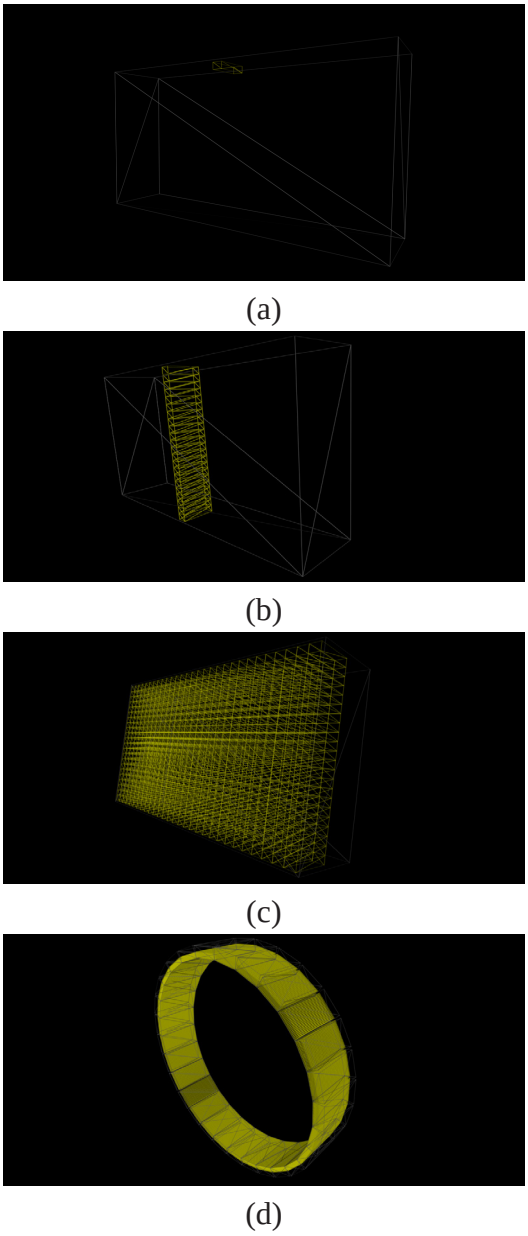
2.8.1.1 Software architecture

GATE is a C++ design, enabling a modular structure built on three fundamental layers: 1) the core layer defines the basic mechanisms available in GATE for geometry definition, time management, source definition, digitization, and data output, 2) the application layer is composed of classes derived from the core layer classes to model specific objects or processes and finally 3) the user layer. The latter allows the user to simulate a setup through the use of scripting. Indeed, the functionality provided by each class is available through script commands, so that the end-user of GATE does not have to manipulate any C++ language. Therefore, a complete nuclear medicine experiment can be defined using the GATE script language, including the object geometry, the radioactive sources, the camera geometry, the detector electronics, the physics processes, the passing of time, the kinetic parameters, and the output format. Some important modeling instructions will be demonstrated hereafter while discussing the implemented functionality as in (4).

2.8.1.2 Defining the geometric features of an acquisition

All geometric features involved in modeling a SPECT or PET acquisition are defined using basics elements known as “logical volumes”. A logical volume is defined by its name, shape, size, and material composition. When logical volumes are placed at specific positions, they form “physical volumes”. Repeaters can be used to replicate and place logical volumes at multiple positions and orientations to form physical volumes. Repeaters are elementary geometrical transformations such as rotations and translations applied in succession.

More complex structures can be created by combining various types of repeaters (e.g., ring, linear, quadrant and cubic array), as shown in figure 2.6 which demonstrates how the ring detector geometry is constructed for a PET scanner. A single crystal is repeated with a cubic array repeater into a block of 22×29 crystals. In turn these blocks are repeated are repeated over a ring. When modeling a scanner using combinations of volumes with material properties,



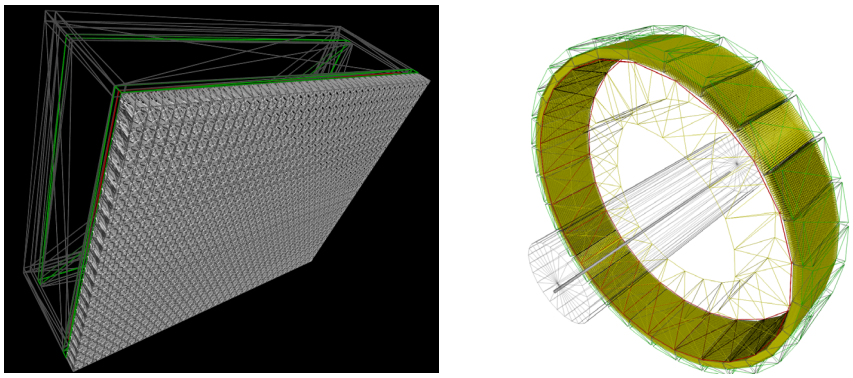
(a)

(b)

(c)

(d)

Figure 2.6: (a) Single crystal, (b) Single crystal repeated into an array of 22, (c) 22×29 crystals forming a block, (d) Blocks repeated over a ring.



(a) (b)

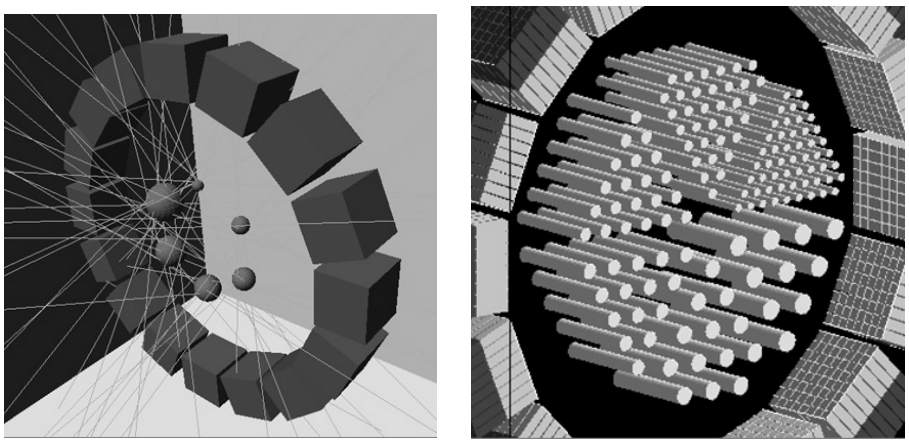
Figure 2.7: Example scanner types modeled in GATE: (a) SPECT system , (b) PET system.

specific guidelines with respect to the geometrical hierarchy must be followed. Most PET scanners are built following comparable concepts: one or more rings, each ring consisting of several blocks of scintillating material, each block being subdivided in crystal parts, etc. For SPECT, the detector mostly consists in a gamma camera with a continuous or a pixelated crystal and a collimator. Most of these geometrical concepts are common to many different imaging systems. To facilitate the modelling of detectors, predefined global *systems* are used. From the user's point of view, the main property of a *system* is that its geometric hierarchy is automatically accounted for by the corresponding data output formats. ASCII and ROOT (85) output files are available for all *systems* and can easily be transformed to a list mode file for reconstruction purposes.

Examples of a modeled SPECT and PET scanner using these predefined systems are shown in figure 2.7.

2.8.1.3 Defining radioactive sources

The spatial distribution of a radioactive source can be specified using five two-dimensional shapes: circle, annulus, ellipse, square, and rectangle, and four three-dimensional shapes: sphere, ellipsoid, cylinder, and parallelepiped. Examples of simple analytical sources



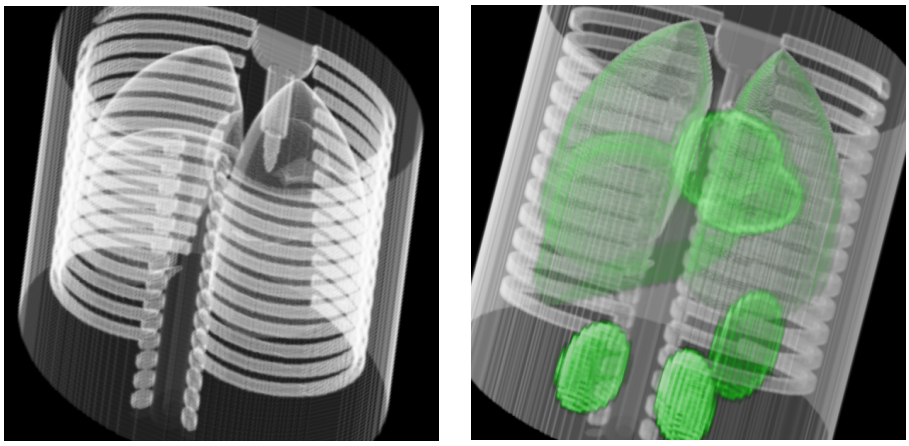
(a) (b)

Figure 2.8: Example of analytical sources: (a) jaszczak, (b) Derenzo phantom.

are shown in figure 2.8. Moreover, the angular distribution of the decay particles can be tuned by using azimuthal and polar angles, making it possible to emit particles in different angular spans. In PET, the user can force the annihilation photons to be emitted back-to-back to significantly decrease the simulation time.

Radioactive sources can also be described using images, what is called a voxelized description. Using such voxelized descriptions, any inhomogeneous, anthropomorphic source or patient image can be used to model realistic acquisitions. As an example, a rendered image of the MCAT (86) voxelized attenuation geometry is shown in figure 2.9a. A voxelized source distribution with activity in the heart, liver, lungs, kidneys and spleen fused with the attenuation geometry is shown in figure 2.9b.

Although quite realistic anthropomorphic phantoms are available for human and for small animal studies, the impact of Monte Carlo simulations in SPECT and PET would be increased by the development of a wider variety of numerical anthropomorphic phantoms. These should include humans with a wide variety of body habitus and motions, and small animal models with breathing and cardiac motions. Several GATE development efforts are currently ongoing



(a) (b)

Figure 2.9: Rendered images of the MCAT phantom: (a) attenuation geometry; (b) source distribution fused with the attenuation geometry.

in this direction (87).

2.8.2 GATE: Time management

A major and recent advance in Monte Carlo simulations is the current possibility to model time-dependent phenomena. This makes it possible to realistically model radioactive decay, dynamic biodistributions, physiological motions such as respiratory and cardiac motions, rotation of the scanner, TOF-PET, dead time, count rate, random coincidences and event pile-up. In GATE, time is explicitly kept track of during the simulation by a virtual clock that synchronizes all time-dependent processes. The start and stop times of the acquisition must be provided by the user, as well as the sampling interval. In such a “time slice”, the geometry is kept at rest and the update is performed at the time slice transition. However, the sources are allowed to decay within the time slice and the particle transport proceeds.

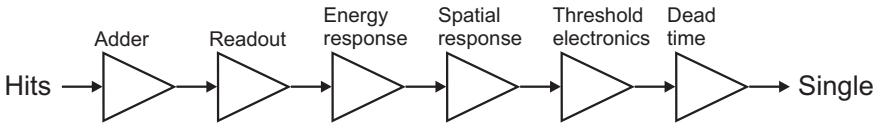


Figure 2.10: The digitizer is organized as a chain of several modules that processes the *hits* to yield a single, which represents a physical observable.

2.8.3 GATE: Digitization

Parts of the scanner geometry are designated as sensitive detectors, and particle interactions are recorded and scored within these regions. Firstly, *hits* are generated from interactions that occur inside the detector parts of the scanner (e.g., the crystal). The data contained in these *hits* include the energy that has been deposited, the interaction positions, the origin of the particle, the type of interaction, the volume name in which the interaction took place, and the time information. Secondly, similar sensitive regions are used to detect and count the Compton and Rayleigh interactions occurring within the scanner's field-of-view (FOV). For example the number of scattering interactions that occurred in the light guide, the isolation, the shielding and the phantom or patient can be recorded.

The digitizer consists of a chain of processing modules (shown in figure 2.10) that takes a list of *hits* from the sensitive detectors and transforms them into *pulses* referred to as singles. The processing modules of the digitizer can be controlled using the GATE scripting language. The key elements of this chain are now briefly described. Firstly the adder sums the deposited energy of all interactions in a sensitive detector (*hits*) to yield a *pulse*. The position of the *pulse* is calculated from the energy-weighted centroid of the *hit* positions, and the time of the *pulse* is set to that of the first *hit* within the volume. If a particle interacts in several sensitive detectors, for instance after crystal scatter or after crystal penetration, the hit adder will generate a list of *pulses*, one for each sensitive detector. The second instruction of the digitizer (readout) implements the readout segmentation of the scanner. The position of the *pulse* is set to that of the *pulse* from the adder that has the most energy (winner-takes-all paradigm).

The energy response instruction applies a blurring to the energy

of the *pulse*. The spatial response instruction applies a Gaussian blur of the position for SPECT. For PET, spatial resolution is calculated by the pulse reader, which simulates the intrinsic spatial resolution of the detector. Furthermore, an energy discrimination is performed to set an acceptance window (thresholder electronics). Also, both paralyzable (the detector is dead for a period t after every event, even if the event arrives during the dead period of the previous event) and non-paralyzable dead times (the detector is dead for a period t after every event) can be modeled on an event-by-event basis.

Other user-defined modules can be added individually to model more specific properties such as e.g. the time resolution. Moreover for PET, at the end of a digitizer chain, a sorter can be added to find pairs of singles that are in coincidence (whenever the time interval between the singles is less than a user-defined coincidence window). Using the event number and the Compton flag, randoms and scatter coincidences can be differentiated from true coincidences. Multiple coincidences and auto-coincidences are also taken into account.

The optimization of the digitizer chain parameters to reproduce the behaviour of a specific scanner is very time-consuming. This is best done by comparing the results from different sets of digitizer parameters using the same series of *hits* with those of the real tomography when available. To perform this optimization, GATE offers an operating mode named digiGATE. In this mode, *hits* are read from a data file generated by GATE and fed directly into the digitizer chain. All conditions are kept identical in the simulations including time dependencies.

2.9 Efficiency

As pointed out in this chapter, Monte Carlo simulations are an indispensable research tool in the field of nuclear medicine. It has its applications in studying detector design, quantification analysis, correction methods for image degradations, detection tasks and other fields. In addition, we have shown that GATE is quickly becoming the tool of excellence for Monte Carlo simulations in PET and SPECT. As GATE is built on GEANT4, it provides almost limitless possibilities.

However, although the design of a new prototype SPECT detector may be feasible within hours in GATE, a realistic acquisition with

a simple phantom for that detector may take several weeks to complete. Similarly, Monte Carlo simulators hold the key to improved reconstructions when used as a forward projector, but the reconstruction may take a few weeks instead of the 10 minutes expected by the physician. Matters become even more dramatic when acquisitions with realistic phantoms are considered. A realistic SPECT simulation of the MCAT phantom takes approximately 7 years and 111 days to complete with GATE on one CPU (88).

2.9.1 The computational cost of GATE simulations

There are several reasons why GATE simulations are computationally expensive, or rather inefficient. They can mainly be grouped in two categories that are discussed next.

2.9.1.1 Generality and complexity of GATE

GEANT4 nor GATE were built with efficiency as a primary objective. The flexibility and generality of Monte Carlo simulations with GEANT4, including the photon transport model, navigation algorithms and probability density functions are inherited by GATE and mainly extended with functionality specifically for PET and SPECT simulations. The GATE software modules therefore display a complete lack of algorithms dedicated to the efficient simulation of PET and SPECT.

The effects are most pronounced in two areas. Firstly, the general particle tracking algorithms employed by GEANT4 are accurate and detailed, but no implementations optimized for the geometry used in PET and SPECT are incorporated into GATE. This results in slow navigation through voxelized and parameterized geometries. Both of these are used frequently in Monte Carlo simulations in nuclear medicine for realistic phantoms, real patient-based attenuation maps and collimators. Therefore, a simulation with a realistic voxelized phantom and a LEHR collimator with 161120 individual air holes is extremely time-consuming. Moreover, a simulation with a parameterized fan beam collimator is nearly impossible with GATE. Secondly, the current GATE analysis module is based on a computationally expensive recursive search through the event history. This results in a slow analysis of each event. In addition, it poses a limitation, as

more detailed information cannot be recorded without introducing an even larger computational overhead.

2.9.1.2 Absence of acceleration methods

Aside from the efficiency problems described above, there is also a lack of dedicated acceleration methods in GATE. A Monte Carlo simulation involving the exact pdfs may be impractical from a computational point of view, requiring unrealistic computation time to achieve reliable results. This computational issue is most cumbersome for SPECT since the total number of detected counts in SPECT is less than 0.02 % of the generated events because the collimator in front of the crystal stops most of the incoming photons. The effect is most pronounced in high resolution collimator variants.

2.9.1.3 Strategies to improve efficiency

Several strategies can be employed separately or in combination with each other to improve the computational efficiency of Monte Carlo simulations in nuclear medical imaging in general, and GATE in particular. The goal is always to calculate a result with a given variance in a shorter amount of time without introducing (significant) bias in the end results.

In some cases these techniques do not require any approximations to be made to the physics. Exploitation of the CPU versus memory trade-off by using pre-calculation or lookup tables and improved tracking algorithms are typical examples. A more direct approach consists of parallelizing the code so that the computational load can be spread over a number of CPUs (89; 90; 91).

In other cases, the improved efficiency is paid for in terms of accuracy by introducing approximations. Hybrid methods replace a computationally inefficient part of the simulation, such as the detector, by an analytical or tabulated model (92; 93). It is also possible to bias the sampling so that particle histories that are likely to contribute to the result are sampled more often than others. This class of techniques must be applied with caution in order to avoid a bias of the result. A few well known techniques are:

- Stratification: Some regions of a phantom may be of higher interest than others. With stratification those outcomes which

are more likely to result in a detected photon are sampled more frequently (94). A production table must be built during a short simulation in order to set the frequencies and to associate weights to the particle histories.

- Forced detection: Photons are forced to scatter into a solid angle (depending on the acceptance angle of the collimator) so that they are forced through the collimator air holes. Weight corrections must again be applied to avoid a bias of the result (94). This method drastically improves the efficiency for SPECT. However, it also limits its applicability as collimator and detector interactions and septal penetration depend on the size of the acceptance angle that is used.

- Splitting and Russian Roulette: Photons may be be split into a number of photons with a lower weight that reflects the number of photons the original photon was split into. This is one of the most safe methods, although care must be taken to avoid excessive splitting (94).

- Forced non-interaction: Efficiency may improve if some interactions, which would otherwise not contribute to the result, are not allowed. A commonly used method is not to allow photoelectric absorption. This way, photons live longer and may have a higher chance to be detected (3).

- Forced interaction: In some cases there is a loss in efficiency because the photons leave a part of the geometry, such as the detector, without interacting. Time is spent in tracking these photons through the geometry, without any contribution to the result. These photons can be forced to interact within the geometry. The interaction probability distribution $p(N_\lambda)dN_\lambda = e^{-N_\lambda}dN_\lambda / \int_0^{N_\Lambda} e^{-N_\lambda'}dN_\lambda'$ is then used to calculate weight corrections to avoid a bias in the result. Here N_λ is the distance measured in mean free paths and N_Λ is the total number of mean free paths along the direction of motion of the photon to the end of the geometry (3).

The inefficiency of GATE simulations in general and SPECT simulations with GATE in particular results from a combination of the problems described above. Any single acceleration method may

not be sufficient to reach our goal of clinically acceptable calculation times. The acceleration of Monte Carlo simulations with GATE can only be achieved through a combination of fast navigation algorithms, code parallelization, dedicated variance reduction techniques and efficient data analysis.

2.10 Summary

This introductory chapter gave an overview of the research field within which this dissertation was conducted. Nuclear medical imaging was introduced from acquisition to reconstruction. The basic principles of Monte Carlo were introduced and the relevance of Monte Carlo simulations to nuclear medicine was discussed. The main Monte Carlo simulators for PET and SPECT were introduced and GATE, based on the GEANT4 nuclear physics code, was presented as the Monte Carlo tool of excellence that is increasingly used in nuclear medical imaging. One of the largest shortcomings of Monte Carlo simulations with GATE was pointed out: the large computational cost. Chapter 3 will significantly extend the flexibility of the analysis module of GATE to allow for a detailed and efficient validation of the acceleration methods in the subsequent chapters. Chapters 4, 5 and 6 will tackle the efficiency problem through particle splitting, forced detection and parallelization respectively.

References

[1] S. Jan et al. *GATE, a simulation toolkit for PET and SPECT*. Phys. Med. Biol., 49(19):4543–4561, 2004.

[2] GEANT4 Collaboration. *GEANT4 - A Simulation Toolkit*. Nucl. Inst. & Meth. in Phys. Res. A, 506:250–303, 2003.

[3] A.F. Bielajew. *Fundamentals of the Monte Carlo method*. The University of Michigan Press, 2001.

[4] S. Staelens and I. Buvat. *Monte Carlo Simulations in Nuclear Medicine Imaging*. Elsevier Science, 2008.

[5] S. Staelens and I. Lemahieu. *Monte Carlo based image reconstruction in emission tomography*. CRC Press, 2008.

[6] Geant4. http://geant4.web.cern.ch/geant4/UserDocumentation/, 2008.

[7] OG Collaboration. http://opengatecollaboration.healthgrid.org, 2008.

[8] S. Cherry, J. Sorensen, and M. Phelps. *Physics in Nuclear Medicine*. Saunders, 3 edition, 2003.

[9] Z. Cho, J. Jones, and M. Singh. *Foundations of medical imaging*. John Wiley and sons Inc., 1993.

[10] G.W. Sloof, F.C. Visser, J.J. Bax, A. van Lingen, J. Eersels, F.F. Knapp, and G.J.J Teule. *Increased uptake of iodine-123-BMIPP in Chronic Ischemic Heart Disease: Comparison with Fluorine-18-FDG*. J. Nucl. Med., 39(2):255–260, 1998.

[11] B.L. Holman and M.D. Devous. *Functional Brain SPECT: The Emergence of a Powerful Clinical Method*. J. Nucl. Med., 20:342–352, 1992.

[12] E. Even-Sapir, U. Metser, E. Mishani, G. Lievshitz, H. Lerman, and I. Leibovitch. *The Detection of Bone Metastases in Patients with High-Risk Prostate Cancer: 99mTc-MDP Planar Bone Scintigraphy, Single- and Multi-Field-of-View SPECT, 18F-Fluoride PET, and 18F-Fluoride PET/CT*. J. Nucl. Med., 47:287–297, 2006.

[13] B. Brans, O. Linden, F. Giammarile, J. Tennvall, and C. Punt. *Clinical applications of newer radionuclide therapies*. Eur. J. Cancer, 42:994–1003, 2006.

[14] C. Hindorf, S. Chittenden, L. Causer, V.J. Lewington, H.R. Macke, and G.D. Flux. *Dosimetry for 90Y-DOTATOC therapies in patients with neuroendocrine tumors*. Cancer Biother. Radio., 22(1):130–135, 2007.

[15] L. Shepp and Y. Vardi. *Maximum likelihood reconstruction for emission tomography*. IEEE Trans. Med. Imaging, 1(2):113–122, 1982.

[16] G. Marsaglia, A. Zaman, and W.W. Tsang. *Toward a Universal Random Number Generator*. Stat. Prob. Let., 9:35–39, 1990.

[17] D.E. Knuth. *Seminumerical algorithms*. Addison Wesley, 1997.

[18] G. Marsaglia, A. Zaman, and W.W. Tsang. *Random numbers fall mainly in the plains*. Nat. Acad. Sci., 61:25–28, 1968.

[19] G. Marsaglia, B. Narasimhan, and A. Zaman. *A Random Number Generator for PCs*. Comp. Phys. Comm., 60:345–349, 1990.

[20] F. James. *A Review of Pseudorandom Number Generators*. Comp. Phys. Comm., 60:329–344, 1990.

[21] M. Luscher. *A Portable High-Quality Random Number Generator for Lattice Field Theory Simulations*. Comp. Phys. Comm., 79:100–110, 1994.

[22] M. Matsumoto and T. Nishimura. *Mersenne Twister: a 623-dimensionally equidistributed uniform pseudorandom number generator*. ACM Transactions on Modeling and Computer Simulations: Special Issue on Uniform Random Number Generation, 8:3–30, 1998.

[23] F.B. Brown et al. *MCNP Version 5*. Trans. Am. Nucl. Soc., 87:273–381, 2002.

[24] R. Durrett. *Probability: Theory and Examples*. 1995.

[25] M. Ljungber, S. Strand, and M. King. Institute of Physics, 1998.

[26] P. Andreo. *Monte Carlo techniques in medical radiation physics*. Phys. Med. Biol., 36:861–920, 1991.

[27] H. Zaidi. *Relevance of accurate Monte Carlo modeling in nuclear medical imaging*. Med. Phys., 26:574–608, 1999.

[28] I. Buvat and I. Castiglioni. *Monte Carlo simulations in SPECT and PET*. J. Nucl. Med., 46:48–61, 2002.

[29] S. Webb, D. Binnie, M. Flower, and R. Ott. *Monte Carlo modeling of the performance of a rotating slit collimator for improved planar gamma camera imaging*. Phys. Med. Biol., 37:1095–1108, 1992.

[30] A. Cot, J. Sempau, D. Pareto, S. Bullich, J. Pavia, F. Calvino, and D. Ros. *Evaluation of the geometric, scatter, and septal penetration components in fan-beam collimators using Monte Carlo simulation*. IEEE Trans. Nucl. Sci, 49:12–16, 2002.

[31] T. Song, Y. Choi, Y. Chung, J. Jung, Y. Choe, K. Lee, S. Kim, and B. Kim. *Optimization of pinhole collimator for small animal SPECT using Monte Carlo simulation*. IEEE Trans. Nucl. Sci, 50:327–332, 2003.

[32] F. van der Have and F. Beekman. *Photon penetration and scatter in micro-pinhole imaging: a Monte Carlo investigation*. Phys. Med. Biol., 49:1369–1386, 2004.

[33] D. Gagnon, G. Zeng, M. Links, J. Griesmer, and C. Valentino. *Design considerations for a new solid state gammacamera: SOLSTICE*. In Proceedings of the 2001 IEEE Medical Imaging Conference, pages 1156–1160, 2001.

[34] J. Griesmer, B. Kline, J. Grosholz, K. Parnham, and D. Gagnon. *Performance evaluation of a new CZT detector for nuclear medicine: SOLSTICE*. In Proceedings of the 1996 IEEE Medical Imaging Conference, pages 1050–1054, 2001.

[35] S. Staelens, M. Koole, S. Vandenberghe, Y. D'Asseler, I. Lemahieu, and R. Van de Walle. *The geometric transfer function for a slat collimator mounted on a strip detector*. IEEE Trans. Nucl. Sci, 52:708–714, 2005.

[36] C. Groiselle, Y. D'Asseler, J. Kolthammer, C. Matthews, and S. Glick. *A Monte Carlo simulation study to evaluate septal spacing using triple-head hybrid PET imaging*. IEEE Trans. Nucl. Sci, 50:1339–1346, 2003.

[37] C. Thompson. *The effect of collimation on singles rates in multi-slice PET*. IEEE Trans. Nucl. Sci, 35:598–602, 1988.

[38] S. Surti, J. Karp, and G. Muehllehner. *Image quality assessment of LaBr3-based whole-body 3D PET scanners: a Monte Carlo evaluation*. Phys. Med. Biol., 49:4593–4610, 2004.

[39] S. Surti and J. Karp. *A count-rate model for PET scanners using pixelated Anger-logic detectors with different scintillators*. Phys. Med. Biol., 50:5697–5715, 2005.

[40] A. Bice, T. Lewellen, R. Miyaoka, R. Harrison, D. Haynor, K. Pollard, C. Hanson, and S. Gillespie. *Monte Carlo simulations of BaF2 detectors used in time of flight positron emission tomography*. IEEE Trans. Nucl. Sci, 37:696–701, 1990.

[41] S. Surti, J. Karp, G. Muehllehner, and P. Raby. *Investigation of lanthanum scintillators for 3-D PET*. IEEE Trans. Nucl. Sci, 50:348–354, 2003.

[42] W. Moses. *Time of fligh in PET revisited*. IEEE Trans. Nucl. Sci, 50:1325–1330, 2003.

[43] W. Moses and S. Derenzo. *Design studies for a PET detector module using a PIN photodiode to measure depth of interaction*. IEEE Trans. Nucl. Sci, 41:1441–1445, 1994.

[44] C. Lartizien, A. Reilhac, N. Costes, M. Janier, and D. Sappey-Marinier. *Monte Carlo simulation-based design study of a LSO-LuAP small animal PET system*. IEEE Trans. Nucl. Sci, 50:1433–1448, 2003.

[45] G. El Fakhri, I. Buvat, M. Pelegrini, H. Benali, P. Almeida, A. Bendriem, A. Todd-Pokropek, and R. Di Paolo. *Respective roles of scatter, attenuation, depth-dependent collimator response and finite spatial resolution in cardiac SPECT quantitation: a Monte Carlo study*. Eur. J. Nucl. Med., 26:437–446, 1999.

[46] G. El Fakhri, I. Buvat, H. Benali, A. Todd-Pokropek, and R. Di Paolo. *Relative impact of scatter, collimator response, attenuation, and finite spatial resolution corrections in cardiac SPECT*. J. Nucl. Med., 41:1400–1408, 2000.

[47] M. Soret, P. Koulibaly, J. Darcourt, S. Hapdey, and I. Buvat. *Quantitative accuracy of dopaminergic neurotransmission imaging using 123I SPECT*. J. Nucl. Med., 44:1184–1193, 2003.

[48] H. Zaidi and K. Koral. *Scatter modeling and compensation in emission tomography*. Eur. J. Nucl. Med., 31:761–782, 2004.

[49] S. Staelens, D. Strul, G. Santin, S. Vandenberghe, M. Koole, Y. D'Asseler, I. Lemahieu, and R. Van de Walle. *Monte Carlo simulations of a scintillation camera using GATE: validation and application modelling*. Phys. Med. Biol., 48(18):3021–3042, 2003.

[50] A. Cot, J. Sempau, D. Pareto, S. Bullich, J. Pavia, F. Calvino, and D. Ros. *Study of the point spread function (PSF) for 123I SPECT imaging using Monte Carlo simulation*. Phys. Med. Biol., 49:3125–3136, 2004.

[51] H. Huesman, E. Salmeron, and J. Baker. *Compensation for crystal penetration in high resolution positron tomography*. IEEE Trans. Nucl. Sci, 36:1100–1107, 1989.

[52] S. Staelens, Y. D'Asseler, S. Vandenberghe, M. Koole, I. Lemahieu, and R. Van de Walle. *A threedimensional theoretical model incorporating spatial detection uncertainty in continuous detector PET*. Phys. Med. Biol., 49:2337–2351, 2004.

[53] F. Lamare, M. Ledesma Carbayo, G. Kontaxakis, A. Santos, A. Turzo, Y. Bizais, C. Cheze-Le Rest, and D. Visvikis. *Incorporation of elastic transformations in list-mode based reconstruction for respiratory motion correction in PET*. In Proceedings of the 2004 IEEE Medical Imaging Conference, pages M03–247, 2004.

[54] F.J. Beekman and M.A. Viergever. *Fast SPECT simulation including object shape dependent scatter*. IEEE Trans. Med. Imaging, 14:271–282, 1995.

[55] F.J. Beekman, J.M. den Harder, M.A. Viergever, and P.P. van Rijk. *SPECT scatter modeling in non-uniform attenuating objects*. Phys. Med. Biol., 42:1133–1142, 1997.

[56] E.C. Frey and B.M.W. Tsui. *A new method for modeling the spatially variant object dependent scatter response function in SPECT*. In Proceedings of the 1996 IEEE Medical Imaging Conference, pages 1082–1086, 1996.

[57] S.R. Meikle, B.F. Hutton, and D.L. Bailey. *A transmission dependent method for scatter correction in SPECT*. J. Nucl. Med., 23:360–367, 1994.

[58] S.H.M. Walrand, L.R. van Elmbt, and S. Pauwels. *Quantification in SPECT using an effective model of the scattering*. Phys. Med. Biol., 39:719–734, 1994.

[59] C.E. Floyd, R.J. Jaszczak, K.L. Greer, and R.E. Coleman. *Inverse Monte Carlo as a unified reconstruciton algorithm for ECT*. J. Nucl. Med., 27:1577–1585, 1986.

[60] D. Lazaro, Z. El Bitar, V. Breton, D. Hill, and I. Buvat. *Fully 3D Monte Carlo reconstruction in SPECT: a feasibility study*. Phys. Med. Biol., 50:3739–3754, 2005.

[61] F.J. Beekman, C. Kamphuis, and M.A. Viergever. *Improved SPECT quantitation using fully three-dimensional iterative spatially variant scatter response compensation*. IEEE Trans. Med. Imaging, 15:491–499, 1996.

[62] F.J. Beekman, C. Kamphuis, and E.C. Frey. *Scatter compensation methods in 3D iterative SPECT reconstruction: A simulation study*. Phys. Med. Biol., 42:1619–1632, 1997.

[63] D. Kadrmas, E. Frey, and B. Tsui. *Application of reconstruction-based scatter compensation to Thallium-201 SPECT:implementations for reduced reconstructed noise*. IEEE Trans. Med. Imaging, 17:325–333, 1998.

[64] T. Farncombe, G.H.M. Narayanan, H. Pretorius, E. Frey, and M. King. *Assessment of scatter compensation strategies for 67Ga SPECT using numerical observer and human LROC studies*. J. Nucl. Med., 45:802–812, 2004.

[65] E.C. Frey, K. Gilland, and B.M.W. Tsui. *Application of task-based measures of image quality to optimization and evaluation of three-dimensional reconstruction-based scatter compensation in myocordial perfusion SPECT*. IEEE Trans. Med. Imaging, 21:1040–1050, 2002.

[66] M. Narayanan, H. Pretorius, S. Dahlberg, J. Leppo, N. Botkin, J. Krasnow, W. Berndt, E. Frey, and M. King. *Evaluation of scatter compensation strategies and their impact on human detection performance Tc-99m myocardial perfusion imaging*. IEEE Trans. Nucl. Sci, 50:1522–1527, 2003.

[67] G. Tourassi, E. Carey, and M. Munley. *Improved lesion detection in SPECT using MLEM reconstruction*. IEEE Trans. Nucl. Sci, 38:780–783, 1991.

[68] P. Bruyant, H. Gifford, G. Gindi, and M. King. *Numerical observer study of MAP-OSEM regularization methods with anatimical priors for lesion detection in Ga-67 images*. IEEE Trans. Nucl. Sci, 51:193–197, 2004.

[69] I. Castiglioni, I. Buvat, G. Rizzo, M. Gilardi, J. Feuardent, and F. Fazio. *A publicly accessible Monte Carlo database for validation purposes in emission tomography*. Eur. J. Nucl. Med., 32:1234–1239, 2005.

[70] A. Reilhac, G. Batan, C. Michel, C. Grova, J. Tohka, D. Collins, N. Costes, and A. Evans. *PET-SORTEO: valida-*

tion and development of database of simulated PET volumes. IEEE Trans. Nucl. Sci, 52:1231–1328, 2005.

[71] R.L. Harrison, S. Dhavala, P.N. Kumar, R. Yiping Shao Manjersshwar, T.K. Lewellen, and F.P. Jansen. *Acceleration of SimSet photon history generation.* In Nuclear Science Symposium Conference Record, volume 3, pages 1835–1838, 2002.

[72] M. Ljungberg. *The SIMIND Monte Carlo program*. In M. Ljungberg, S.E. Strand, and M. A. King, editors, Monte Carlo calculations in nuclear medicine: Applications in diagnostic imaging, pages 145–163, Bristol, 1998. Institute of Physics Publishing.

[73] J. Yanch, C. Dobrzeniecki, C. Ramanathan, and R. Behrman. *Physically realistic Monte Carlo simulaiton of source, collimator and tomographic data acquisition for emission computed tomography*. Phys. Med. Biol., 37:853–870, 1992.

[74] C. Thomson, J. Cantu, and Y. Picard. *PETSIM: Monte Carlo program simulation of all sensitivity and resolution parameters of cylindrical positron imaging systems*. Phys. Med. Biol., 37:731–749, 1992.

[75] H. Zaidi, A. Hermann Scheurer, and C. Morel. *An Object-Oriented Monte Carlo simulator for 3D cylindrical positron tomographs*. Comp. Meth. Prog. Biomed., 58:133–145, 1999.

[76] A. Reilhac, C. Lartizien, N. Costes, S. Sans, C. Comtat, R. Gunn, and A. Evans. *PET-SORTEO: A Monte Carlo-based simulator with high count rate capabilities*. IEEE Trans. Nucl. Sci, 51:46–52, 2004.

[77] A. Bielajew, H. Hirayama, W. Nelson, and D. Rogers. *History, overview and recent improvements of EGS4*. Tech. Rep., 1994.

[78] I. Buvat and D. Lazaro. *Monte Carlo simulations in emission tomography and GATE: an overview*. Nucl. Inst. & Meth. in Phys. Res. A, in press, 2006.

[79] D. Strul, G. Santin, D. Lazaro, V. Breton, and C. Morel. *GATE (Geant4 Application for Tomographic Emission): a*

PET/SPECT general-purpose simulation platform. Nucl. Phys. B, 125:75–79, 2003.

[80] G. Santin, D. Strul, D. Lazaro, L. Simon, M. Krieguer, M. Vieira, V. Breton, and C. Morel. *GATE: A Geant4-based simulation platform for PET, SPECT integrating movement and time management*. IEEE Trans. Nucl. Sci, 50:1516–1521, 2003.

[81] H. Zaidi. *Relevance of accurate Monte Carlo modeling in nuclear medical imaging*. Med. Phys., 26:574–608, 1999.

[82] O. Klein and Z. Nishina. *Uber die Streuung von strahlung durch freie elektronen nach der neaen relativistischen Quatendynamik von Dirac*. Z. Phys., 52:853–868, 1929.

[83] H. Kahn. *Use of Different Monte Carlo Sampling Techniques*. In Symposium on Monte Carlo Methods, pages 146–190, 1956.

[84] J. H. Hubbel. *Summary of Existing Information on the Incoherent Scattering of Photons particularly on the Validity of the Use of the Incoherent Scattering Function*. Radiat. Phys. Chem., 50(1):113–124, 1997.

[85] R. Brun and F. Rademakers. *ROOT - An Object Oriented Data Analysis Framework*. In Nucl. Inst. & Meth. in Phys. Res. A, volume 389, pages 81–86, 1997.

[86] B.M.W. Tsui, J.A. Terry, and G.T Gullberg. *Evaluation of cardiac cone-beam SPECT using observer performance experiments and ROC analysis*. Investigative Radiology, 28(12):1101–1112, 1993.

[87] R. Taschereau, P. Chow, and A. Chatziioannou. *Monte Carlo simulations of dose from microCT imaging procedures in a realistic mouse phantom*. Med. Phys., 33:216–224, 2006.

[88] J. De Beenhouwer, S. Staelens, S. Vandenberghe, and I. Lemahieu. *Acceleration of GATE SPECT Simulations*. Med. Phys., 35(4):1476–1485, 2008.

[89] L. Deng and Z.S. Xie. *Parallelization of MCNP Monte Carlo neutron and photon transport code in parallel virtual machine and message passing interface*. J. Nucl. Sci. Tech., 36(7):626–629, 1999.

[90] Y.K. Dewaraja, M. Ljungberg, A. Majumdar, A. Bose, and K.F. Koral. *A parallel Monte Carlo code for planar and SPECT imaging: implementation, verification and applications in (131)I SPECT*. Comput. Methods Programs Biomed., 67(2):115–124, 2002.

[91] H. Zaidi, C. Labbé, and C. Morel. *Implementation of an environment for Monte Carlo simulation of fully 3D positron tomography on a high-performance parallel platform*. Parallel Comput., 24:1523–1536, 1998.

[92] H. W. A. M. de Jong, E. T. P. Slijpen, and F. J. Beekman. *Acceleration of Monte Carlo SPECT simulation using convolution-based forced detection*. IEEE Trans. Nucl. Sci, 48(1):58–64, 2001.

[93] X. Song, W. Segars, Y. Du, B. Tsui, and E. Frey. *Fast modeling of the collimator-detector response in Monte Carlo simulation of SPECT imaging using the angular response function*. Phys. Med. Biol., 50:1791–1804, 2005.

[94] D. Haynor, R. Harrison, and T. Lewellen. *The use of importance sampling techniques to improve the efficiency of photon tracking in emission tomography simulations*. Med. Phys., 18:990–1001, 1991.

3

Physics Process Level Discrimination of Detections for GATE: Assessment of Contamination in SPECT and PET

GATE is an extensively validated and one of the most accurate and versatile Monte Carlo codes currently available. It is capable of simulating a nuclear medical acquisition with a high degree of flexibility, detail and accuracy. It allows for complex decay schemes, time-based detector movement and realistic detector geometries. Detailed information with regard to particle interactions is stored during the propagation of particles through the geometry. However, only a predefined fraction of this information is available for post-processing. The analysis module currently used in GATE is inflexible and incapable of storing more detail without introducing a large computational overhead.

However, the validation and design of the acceleration techniques that are discussed in chapters 4, 5 and 6 requires a flexible, detailed and computationally efficient analysis module. To this end, a flexible analysis framework for GATE (version 3.1.2) called ProcessGATE was developed. Our aim is to allow the complete particle history information to be stored temporarily. Simple user-defined algorithms may then process this information during the simulation. At the end of a simulation, the additional data is available for post-processing. Our framework is applicable to any isotope, detector and phantom available in GATE and will become available to the general public as part of a future GATE release. The usefulness of our analysis framework is demonstrated through a case-study.

Radiopharmaceuticals such as ^{131}I-Bexxar are widely used in targeted radionuclide therapy (TRT) for the treatment of B-cell lymphomas. The imaging and therapy can be performed by the same isotope. An accurate estimation of the activity distribution is required for the absorbed dose estimation before and during the TRT. PET imaging with ^{124}I-Bexxar may offer better image quality when compared to SPECT. The long half life of 4.18 days compared to ^{18}F makes it practical to use. In addition, whole body tomographic images can be acquired within a reasonable amount of time due to the high sensitivity of PET.

Unfortunately, ^{124}I is a non-pure positron emitter. It does not only decay by positron emission (like ^{18}F) or by electron capture, but it also generates prompt gamma rays. These gamma rays are emitted simultaneously with the positron, or are in cascade with other gamma rays. These additional gamma rays can result in spurious coincidences. These are much like random coincidences, but are not corrected for with standard protocols for ^{18}F imaging. Several correction methods and spurious activity estimates have been investigated. In (1, 2) a tail fitted function was subtracted in order to correct for the spurious activity. Another correction was based on a modification of the Bergstrom scatter estimate which was then tail fitted (3). A convolution subtraction was performed in (4) using combined scatter and spurious coincidence kernels. However, a measured coincidence could not be labeled as scattered, random or spurious. Therefore, all these methods are based on crude estimates of the spurious activity. In (5) the energy window setting of a PET scanner was optimized for ^{86}Y and ^{124}I using GATE. However, it was not possible to distinguish contamination from scattered or true coincidences in realistic

objects. Therefore, the estimation of the spurious activity was done by tail fitting sinogram data of NEMA-like phantoms.

SPECT imaging may provide quantitative results if a number of image degrading factors are corrected for (6; 7; 8). These isotopes are not in the energy range of a low energy high resolution (LEHR) collimator. Therefore, medium or high energy collimators with a poor spatial resolution are often required. In addition, there may be contamination due to collimator penetration of photons from higher energy peaks. In turn that may lead to low image quality (9; 10).

The object of this case-study is to assess the image degrading factors for ^{124}I and ^{131}I with our analysis framework.

3.1 ProcessGATE analysis framework

The primary unit of a GEANT4 / GATE simulation is an event. It consists of a set of particles produced in interactions and a set of detector responses to these particles (11). Each particle is associated with a track. It holds all information necessary for tracking this particle, such as time, position and kinematic information. Particles are transported one step at a time from one interaction position to the next. The step size can either be limited by geometry or physics. During and at the end of every step, the continuous and discrete physics processes that apply to the particle are evaluated. Thereby, energy may be deposited and new particles may be created. If energy is deposited in a detector component marked as 'sensitive', such as a crystal, then hit objects are created. Hit objects can be processed with a digitizer to model the detector readout. After each step, the final status of the particle is updated through its associated track. The event is completed after all particles have been tracked through the geometry. The result is an event history tree that represents particles in a parent-child relationship.

Figure 3.1 shows a simple example of how an event is handled. A detector consisting of a collimator, a crystal and a back compartment is depicted. A ^{99m}Tc point source is located in a cylindrical water phantom. The event described by figure 3.1 consists of 4 particles, each associated with a track. At decay (D), a gamma photon is created which is associated with track 1. The photon is then transported to the first interaction position where it is Compton scattered (C). This marks the completion of step S1, at which point the track 1 in-

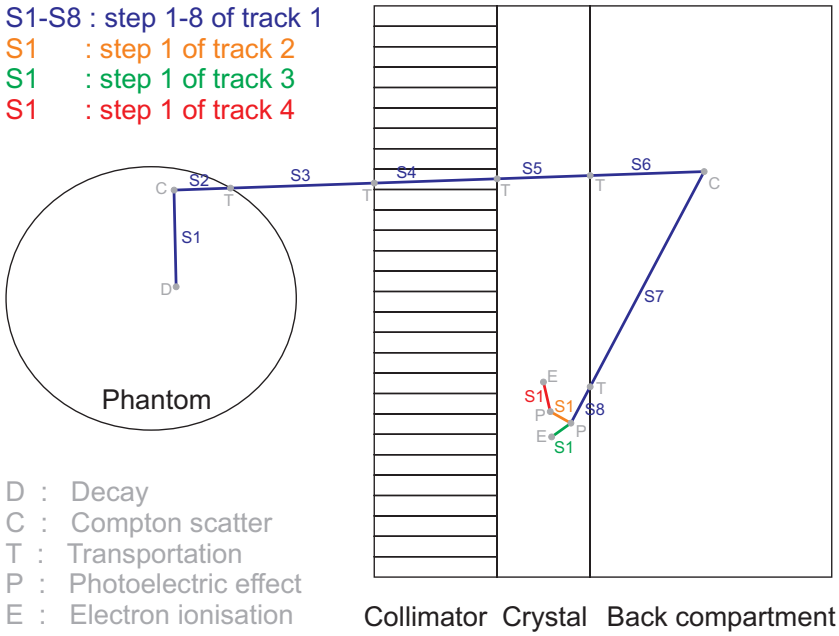


Figure 3.1: Tracking example for one event. A point source in a water phantom is shown together with a detector consisting of a collimator, a crystal and a back compartment. A single gamma photon is tracked through the geometry.

formation is updated. In step S2 the photon is transported (T) to the edge of the phantom because the step size is limited by the phantom geometry. Similarly, the photon is transported to the surface of the collimator (step S3), to the surface of the crystal (step S4) and to the front of the back compartment (step S5). At the end of step S6, the photon is Compton scattered and the track information is updated. Subsequently, the photon is transported to the surface of the crystal (step S7). Finally, the photon is subjected to the photoelectric effect (P) at the end of step S8. As a result of this, a gamma photon (track 2) and an electron (track 3) are created. The information in track 1 is updated and the photon is killed. The photon associated with track 2 is subjected to the photoelectric effect at the end of step 1. Due to this, an electron is created (track 4) and track 2 is killed. The electrons associated with track 3 and 4 both end in electron ionization

after the first step.

At this point the event is completed and it can be analyzed. The interaction information is filtered out for all sensitive volumes by the GATE analysis module and the hits are entered into a digitizer chain. The filter is however based on a computationally expensive recursive search through the event history. Due to this reason, only limited information is finally stored for detector parts other than the crystal: only the volume in which a particle interacted last will be stored in the simulation output. This relates back to figure 3.1 as follows: before detection two Compton interactions were observed and the last Compton interaction was in the back compartment. The analysis module cannot easily be modified and more information could not be added without introducing a large computational overhead and large data files. Ideally, it should be possible to store specific information for a particle along its trajectory through the detector geometry.

In order to meet these demands, we developed a flexible analysis framework for GATE called ProcessGATE. The core structure of our framework is shown in figure 3.2. During the simulation of an event, the particle type of each track and the track number of its parent (parentID) are stored. For each step (S1 to SN) of a track, the physics process (Compton scatter, photoelectric effect,...), and other properties such as the (kinetic) energy, the current time and the current interaction volume are stored in separate 2D interaction tables. Track 4 of the example in figure 3.1 is stored as follows: it is an electron, with parent track 2. The process table shows that it was created by a photoelectric effect and that it ended in electron ionization. At the end of an event, any custom algorithm can be used to filter out information. The tables are cleared again at the start of a new event. The simplicity of the framework, its low memory footprint and the amount of available information leads to limitless analysis possibilities for both SPECT and PET simulations with GATE. A few examples made available through ProcessGATE are:

- Compton and Rayleigh scatter information per volume for every event
- Collimator and Crystal fluorescence detection
- Collimator penetration detection
- User defined classification of detections in PET (see part 3.2)

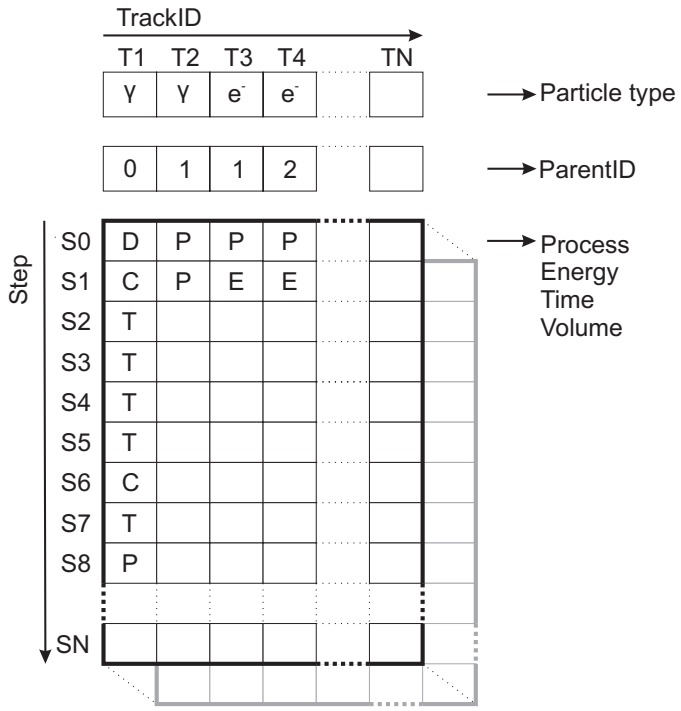


Figure 3.2: The ProcessGATE framework presented as interaction tables. Selected information of particle interactions such as the particle type (gamma, electron,..), physics process (D,C,T,P,E,..), energy, time and the volume is stored in tables during an event. The tables can be indexed by track number (trackID : T1...TN) and step number (S1...SN).

- User defined classification of detections in SPECT (see part 3.2)

3.2 Assessment of contamination

As a first example, we developed an algorithm for PET to identify spurious coincidences in the case of non-pure isotopes. A coincidence in GATE is composed of two single detections that in turn are composed of one or more crystal hits. If any such hit originates

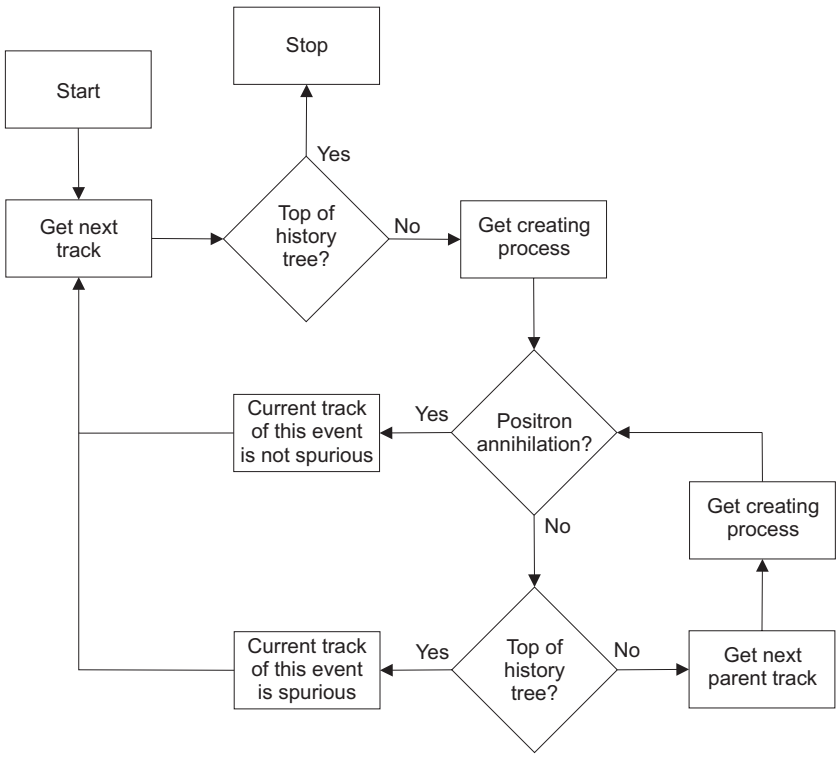


Figure 3.3: Detection scheme for spurious coincidences.

from a photon that was not generated by positron annihilation, then the whole coincidence is regarded as spurious. As such, the possible combinations are: two prompt gammas, one prompt gamma and one annihilation gamma. To retrieve this information, it is necessary to investigate the history of each particle that deposited energy in a detector crystal. The structure of the algorithm is shown in figure 3.3. The process table is indexed to retrieve the creating physics process of the current track. If the creating process was not positron annihilation, then the parentID table is used to identify the creating process of the direct parent of the current track. This continues until we are either at the top of the particle history or positron annihilation is found. The track is marked as spurious if positron annihilation cannot be found. The outcome for each track is then stored for the last 5 events since two separate events can lead to a coincidence de-

tection. At the digitization step of GATE this information can be used to determine the spurious nature of all hits forming a single and consequently of each two singles forming a coincidence. This way, the coincidences can be classified into 4 categories:

- **true coincidences (T):** Both photons are non-spurious (pure) and neither photon scattered before detection.
- **pure scattered coincidences (PSC):** Both photons are pure but at least one photon scattered once or more before detection.
- **spurious contamination (C):** At least one photon is spurious. Neither of the photons scattered before detection.
- **scattered spurious contamination (SCC):** At least one photon is spurious. At least one photon scattered once or more before detection.

As a second example, we facilitate the assessment of contamination in SPECT using ProcessGATE. The classification into five categories is done as follows (12):

- **geometric:** photons that are detected without any preliminary interaction.
- **phantom scatter:** photons that scatter in the phantom and are then directly detected in the detector crystal.
- **penetration through the collimator:** photons that penetrate the collimator directly before detection, independently of their previous interactions.
- **collimator scatter:** photons that undergo scatter in the collimator directly before detection, independently of their previous interactions.
- **backscatter in the end parts:** photons that backscatter in the end parts of the detector before detection, independently of their previous interactions. High energy photons that penetrate the collimator and that are detected through backscatter are classified as backscatter.

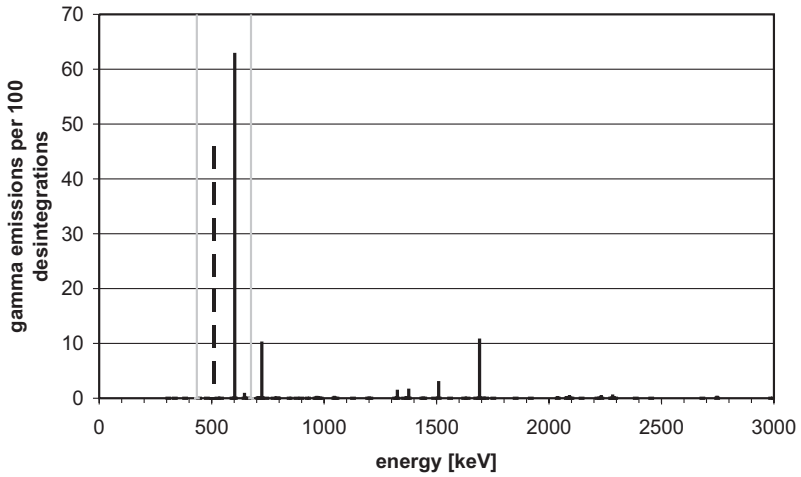


Figure 3.4: Emission spectra for ^{124}I. The dotted line represents the annihilation photons. The light grey lines show the 434-670 keV energy thresholds on the Allegro.

	λ	e^+	e^- capture
^{124}I	4.2d	23%	77%

Table 3.1: ^{124}I decay properties.

3.3 Radionuclides

Table 3.1 shows the decay properties of ^{124}I (13). Most of the positron annihilations and the transitions by electron capture are in cascade with a number of prompt gamma rays. Figure 3.4 shows the gamma

	λ	main γ	other γ
^{131}I	8d	364 keV	637 / 723 keV

Table 3.2: ^{131}I decay properties.

emission spectra per 100 decays. The gamma photons resulting from the positron annihilation are shown as well as the prompt gammas. The 434-670 keV energy thresholds used on the PET camera show that there is a number of prompt gammas directly in the coincidence energy window. Other high energy prompt gammas may also contribute to the spurious coincidences after or by Compton scatter.

Table 3.2 shows the decay properties of ^{131}I (14). The main gamma emission for ^{131}I is at 364 keV. However, the gamma emissions at 637 and 723 keV may also have a large influence.

3.4 Scanners

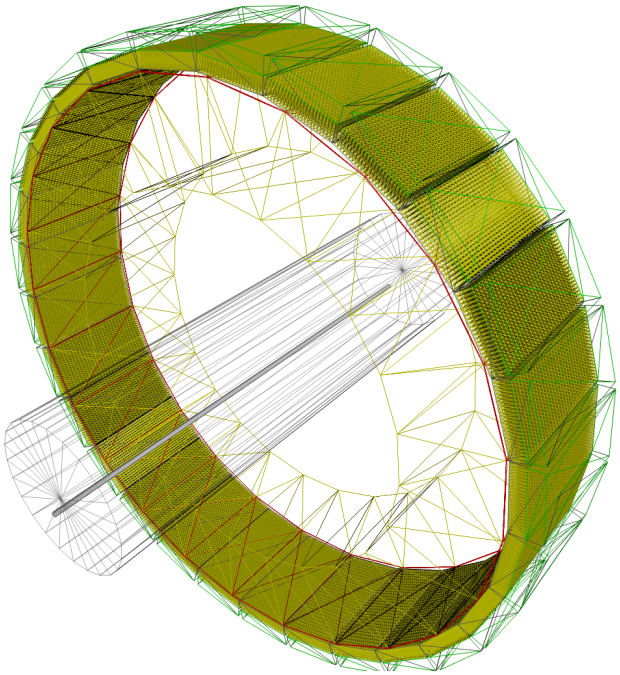


Figure 3.5: Allegro PET setup with a line source in a cylindrical water phantom.

The PET simulations in this chapter are based on the Philips Allegro (15), which is composed of 28 modules of a 22×29 array of

GSO crystals ($4 \times 6 \times 20\,mm^3$). The fully 3D mode was used with a paralysable singles dead time of 210 *ns*, a coincidence window of 7.5 *ns* and an energy resolution of 16.5%. The standard energy window is 434 to 670 keV. The SPECT detector used for the simulation in this chapter is based on the Philips AXIS camera (16) and was modeled as in (17). A high energy general purpose (HEGP) collimator was used. The hexagonal hole inner radius was 0.340 cm (hexagon with side length 0.3926 cm), with a septal thickness of 0.201 cm and a collimator height of 5.84 cm. The NaI crystal (1.905 cm thickness, 9.5% energy resolution, 0.33 cm intrinsic spatial resolution) was covered with a 1.27 cm plastic cover. The back compartment consisted of a glass lightguide (0.9525 cm thickness), plexiglass photomultiplier tubes (10.6 cm thickness), an aluminium compression plate (1.029 cm thickness), an air gap (2.017 cm thickness) and a Pb-Sb ending (1.59 cm thickness).

3.5 Simulations

3.5.1 Line source phantom

The standard energy thresholds used on the Allegro are acceptable for ^{18}F. However, it has been shown previously (5) that this is most likely not the case for ^{124}I. Therefore, we derive optimal thresholds from the simulation of a 1 MBq line source. The line source was suspended in a cylindrical water phantom (diameter 20 cm, length 70 cm) at a 4.5 cm radial offset from the center as shown in figure 3.5. The Noise-Equivalent Counting rate (NEC) (18) was calculated as in formula 3.1. The NEC describes the equivalent coincidence counting rate that would have the same noise properties as the actual counting rate, corrected for spurious coincidences and scatter. A maximum NEC can therefore be used to determine the optimal energy thresholds. The true coincidences (T), pure scattered coincidences (PSC), spurious contamination (C) and scattered spurious contamination (SCC) were determined with ProcessGATE. Random coincidences could be neglected due to the low activity. The energy thresholds derived from the optimal NEC were used subsequently for the thorax phantom simulations.

$$NEC = \frac{T \times T}{T + PSC + SCC + C} \tag{3.1}$$

3.5.2 Thorax phantom

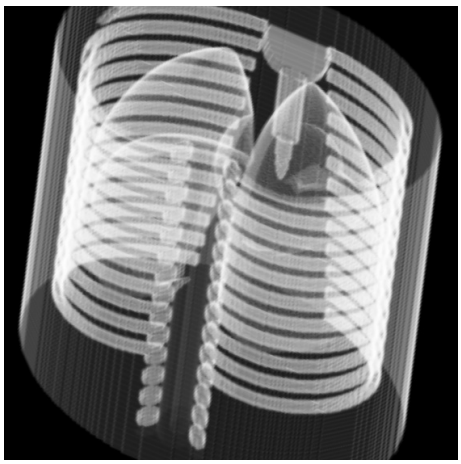


Figure 3.6: Rendered image of the MCAT phantom.

The MCAT thorax phantom (19), shown in figure 3.6, was simulated with the cluster version of GATE developed in chapter 6. An activity distribution for ^{124}I/^{131}I-Bexxar as in (20) was used. The distribution is shown in figure 3.7. The activity concentration ratios of the 80 MBq extended source distribution were chosen as 28:48:4:28::80:52:48:126 for lung, heart, background, liver, kidney, spleen, bloodpool and tumour respectively.

The PET acquisition time was 800 seconds. The sinogram data was reconstructed into images of 144×144 pixels with 4 mm per pixel using OSEM (8 subsets, 8 iterations) with sensitivity and attenuation correction. Single scatter estimation was used as in (21). A constant background was used by the single scatter algorithm during its tail fitting procedure in order to avoid a large overestimation of scatter. The same data were used for three different reconstruction cases: for a standard energy window, for an optimized energy

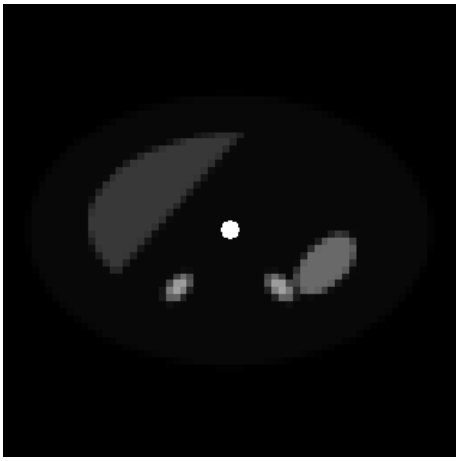


Figure 3.7: Bexxar activity distribution for a slice of the thorax phantom.

window based on the NEC curves for ^{86}Y and without spurious contamination and phantom scatter.

The SPECT acquisition time was 30 seconds per projection over 120 projections with a HEGP collimator. The projections were acquired into a 20% energy window centered at 364 keV and two 10% scatter windows at 318 keV and 413 keV. Sinograms were corrected for scatter using standard triple energy window (TEW) correction (22). After scatter rejection, the data were reconstructed into 128×128 pixel images (3.67 mm pixel size) using standard MLEM (64 iterations) with attenuation correction as in (23). In addition to the normal reconstruction, a contamination free case was also considered in which penetration, collimator scatter, backscatter and phantom scatter were filtered out with ProcessGATE prior to reconstruction.

All reconstructed images were post-smoothed with a 3D Gaussian filter with FWHM of 2 pixels.

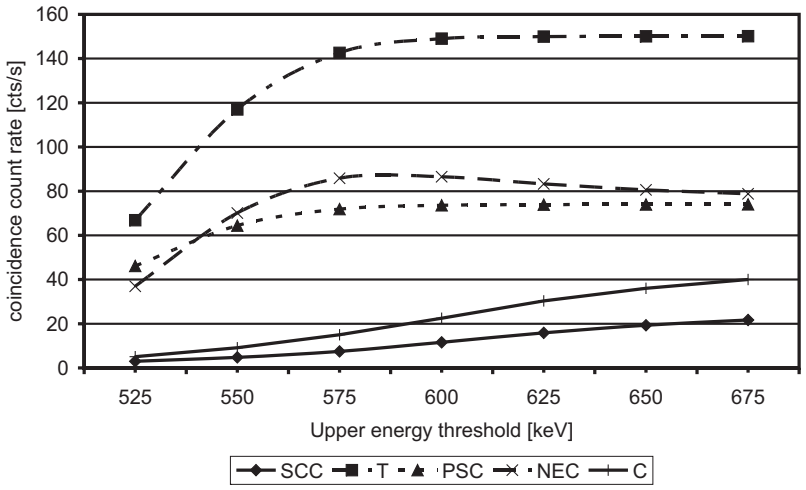


Figure 3.8: NEC curve with a fixed lower energy threshold of 434 keV in the case of the ^{124}I line source simulation. The different contributions are shown: true coincidences (T), pure scattered coincidences (PSC), spurious contamination (C) and scattered spurious contamination (SCC) for 7 upper energy cutoffs.

3.6 Results

3.6.1 Line source phantom

Due to the low abundance of prompt gamma's under 511 keV for ^{124}I, only a variation of the upper energy cutoff was required. The optimal upper energy cutoff was determined using the NEC. All combinations between 434 keV and 675 keV in steps of 25 keV were examined. Figure 3.8 shows the different contributions and the optimal NEC curve. The optimal threshold was found to be 585 keV. Figure 3.9 shows the total energy spectrum for the ^{124}I line source simulation and the different contributions: T, PSC, C and SCC.

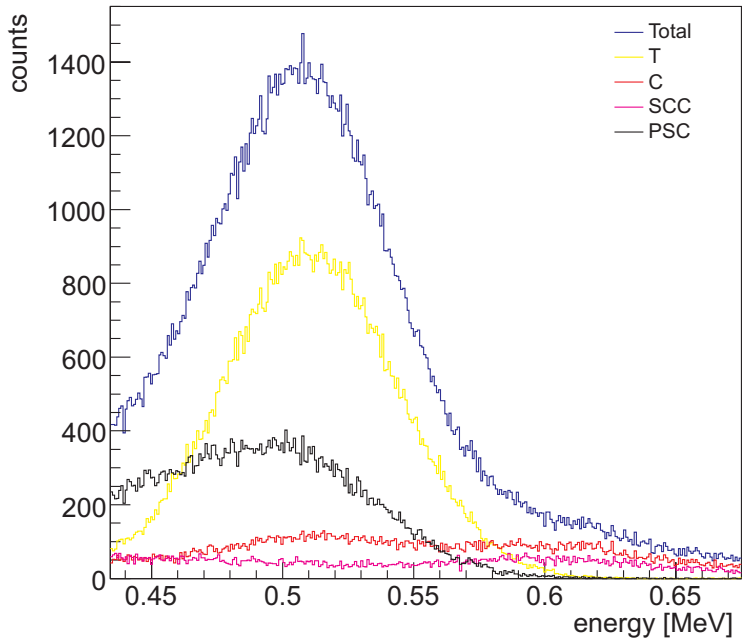


Figure 3.9: Energy spectrum of the ^{124}I line source simulation. The different contributions are shown: true coincidences (T), pure scattered coincidences (PSC), spurious contamination (C) and scattered spurious contamination (SCC).

3.6.2 Thorax phantom

The profile of a sinogram summed over all angles is shown in figure 3.10 for the ^{124}I-Bexxar simulation. The total profile is shown as well as the different contributions from T, PSC, C and SCC. Note the non-uniformity of C and to a lesser extent SCC. Reconstructed images are shown in figure 3.11. Figures 3.11a, b and c show a reconstructed slice of the kidneys and spleen for three energy window settings: the standard energy window (a), the optimized energy window (b) and without spurious contamination (c). Similarly, figures 3.11d, e and f show a reconstructed tumour slice for the different energy window settings. Profiles drawn through the kidneys/spleen and the

tumour are shown in figures 3.12 and 3.13 respectively. The three reconstruction cases are shown together with the simulated activity.

The energy spectrum of the ^{131}I-Bexxar simulation is shown in figure 3.14. The total energy spectrum is shown together with the different contributions: collimator scatter, phantom scatter, backscatter, penetration and geometric photons. The kidneys/spleen and tumour reconstructed slices are shown in figures 3.15a,b and c,d respectively for the two reconstruction cases. Similarly as before, profiles were drawn through the kidneys/spleen and the tumour images. The profiles are shown in figures 3.16 and 3.17 and are compared to both the optimized energy window ^{124}I-Bexxar PET profiles and the simulated activity.

3.7 Discussion

Our framework is essentially composed of two layers. In the first layer it is possible to select specific particle history information. This information can be stored directly (i.e. Compton scatter in a certain detector part) or it can be processed during the simulation (i.e. 3.3). The additional data is available at the end of the simulation together with regular GATE data output. This enables enhanced post-processing, which forms the second layer of our framework. To our knowledge, no Monte Carlo simulator was capable of simulating complex decay schemes and detecting spurious coincidences for realistic phantoms. However, with our framework this was accomplished by a simple algorithm. The user-defined classification of detections has clearly demonstrated the flexibility of ProcessGATE. This will lead to a better assessment of the different components of contamination. Its applications range from the optimization of correction methods to the investigation of the properties of current and future detectors.

A remaining difficulty for PET imaging with a non-pure isotope is that some spurious activity still remains, despite optimized energy thresholds and single scatter estimation. The emission spectra already suggest this with prompt gamma rays close to the 511 keV positron annihilation photons. Only now it could be assessed accurately with the classification of the singles energy spectra and the resulting NEC curves. Two additional problems arise as a result. Firstly, it is obvious that attenuation correction is not correctly

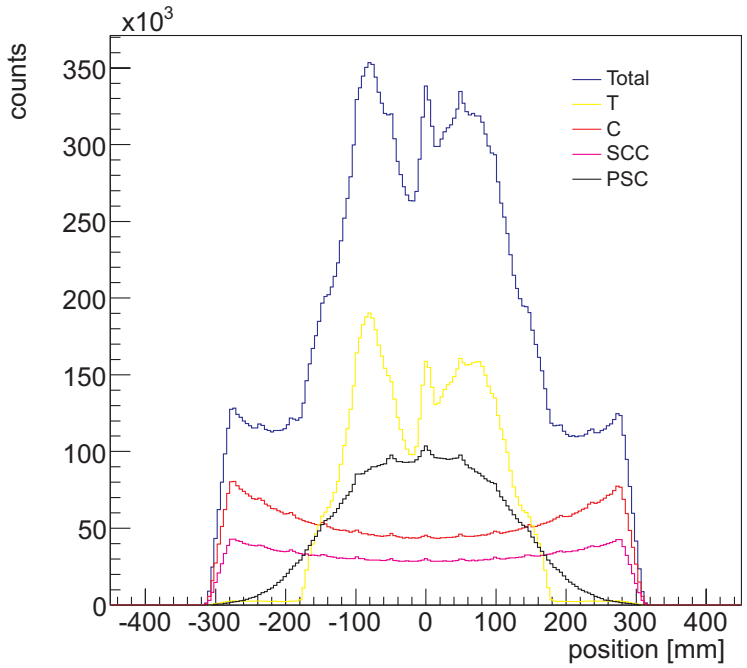
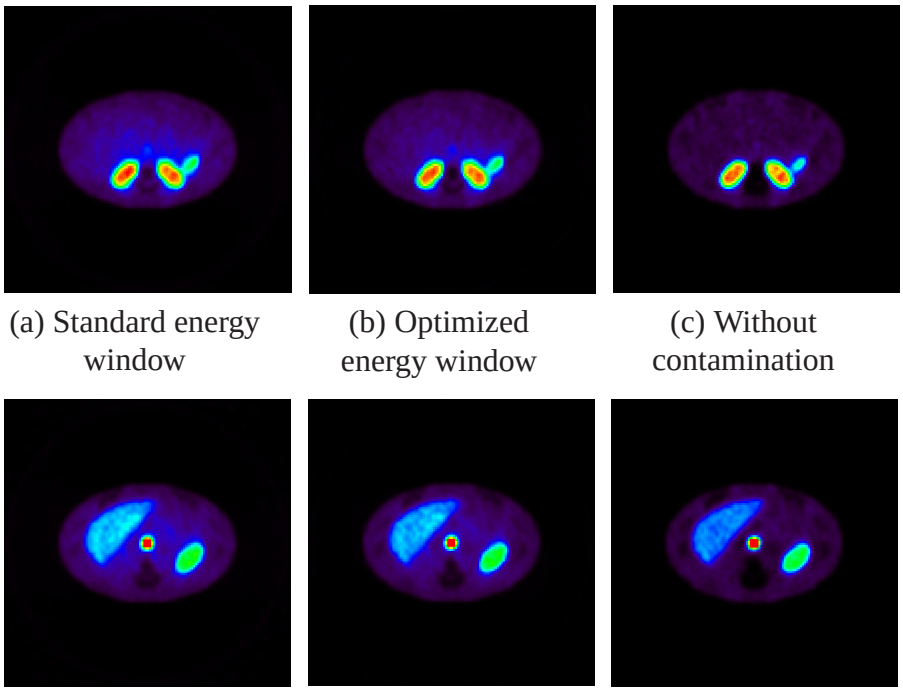


Figure 3.10: Sinogram profile for the ^{124}I-Bexxar simulation. The total profile is shown together with the different contributions: true coincidences (T), pure scattered coincidences (PSC), spurious contamination (C) and scattered spurious contamination (SCC).

modeled for spurious coincidences. Secondly, the single scatter algorithm must add a constant background for its tail fitting procedure. However, it is clear from our classification of the sinogram profiles that SCC and especially C cannot be assumed to be spatially uniform. The removal of a constant background is thus only a crude estimate. A dedicated algorithm to remove spurious coincidences in combination with single scatter estimation will likely provide better results.

Straightforward TEW correction was used for the SPECT isotope in this case study. The emission spectra already indicate that high energy photons are the main difficulty but it could now also be assessed accurately with our classification. TEW correction is able to partly



(a) Standard energy window (b) Optimized energy window (c) Without contamination

(d) Standard energy window (e) Optimized energy window (f) Without contamination

Figure 3.11: Reconstructed slices of the MCAT phantom for ^{124}I-Bexxar. Kidney slice: a,b,c. Tumour slice: d,e,f.

correct for phantom scatter and backscatter, but not for septal penetration since the scatter window does not take these detections into account. The influence of septal penetration is high in the case of ^{131}I due to the high energy emission peaks despite the HEGP collimator. The combination of septal penetration with the large spatial resolution results in poor reconstructed image quality.

The comparison of PET and SPECT in this case study shows that ^{124}I-Bexxar with simple optimized energy thresholds offers better image quality when standard reconstruction techniques are used. However, accurate modeling of collimator scatter and penetration in the reconstruction can improve image quality. In (8), a Monte Carlo projector with accurate physical modeling of the collimator was used, which resulted in superior image quality when compared

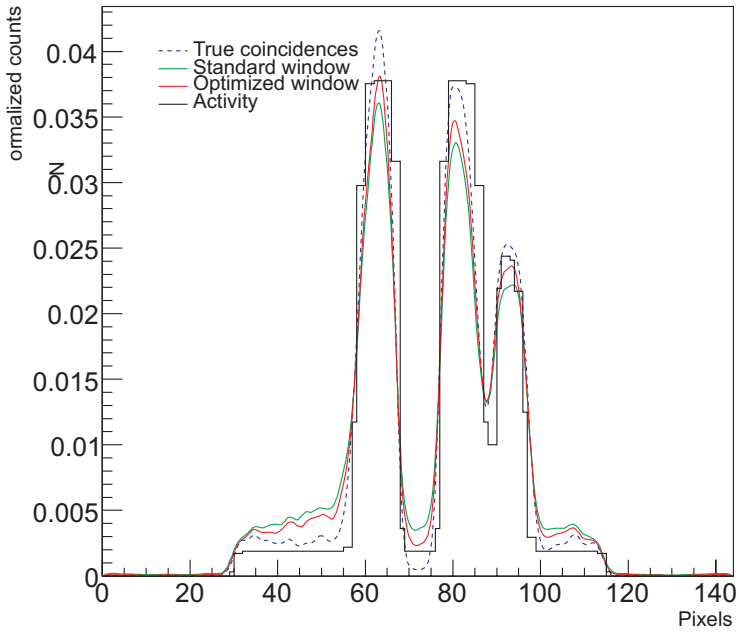


Figure 3.12: Kidney profile of the reconstructed slices shown in figure 3.11 for ^{124}I-Bexxar. The three reconstruction cases are shown: standard window, optimized window and without contamination. The simulated activity is also shown.

to TEW correction.

3.8 Summary and original contributions

GATE simulations allow for complex decay schemes, realistic phantoms and a large variety of detector geometries. However, only a fraction of the information in each particle history is available for post-processing. In this chapter we presented an analysis framework that extends the analysis capabilities of GATE. It allows for a user defined classification of detections based on complete particle histories. Our framework can be applied to any isotope, phantom and

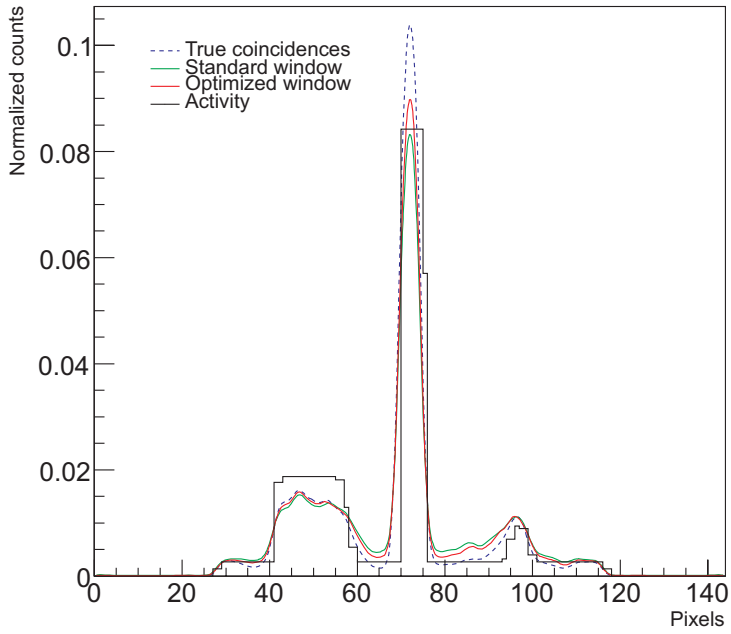


Figure 3.13: Tumour profile of the reconstructed slices shown in figure 3.11 for ^{124}I-Bexxar. The three reconstruction cases are shown: standard window, optimized window and without contamination. The simulated activity is also shown.

detector geometry available in GATE.

Our framework is applied to a case-study. Radiopharmaceuticals such as ^{131}I-Bexxar are widely used in targeted radionuclide therapy (TRT) for the treatment of B-cell lymphomas. It may provide quantitative results, provided that image degradation is corrected for. Quantitative PET imaging with ^{124}I on the other hand offers a number of advantages. Unfortunately, this is a non-pure positron emitter. It does not only decay by positron emission, but it also generates prompt gamma rays which can result in spurious coincidences.

In the case of PET an algorithm was designed to detect spurious coincidences with ProcessGATE in order to assess their distribution

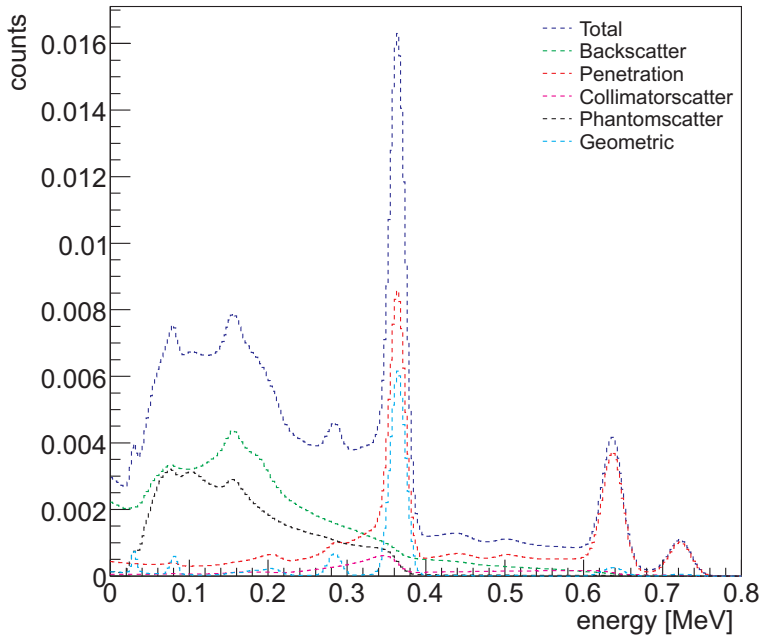
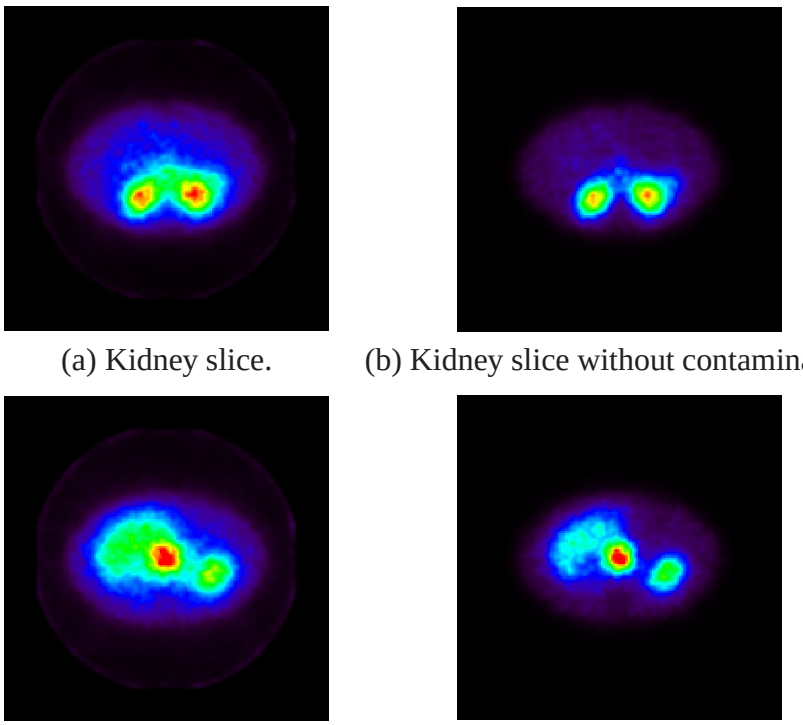


Figure 3.14: ^{131}I-Bexxar simulation: total energy spectrum, collimatorscatter, phantom scatter, backscatter, penetration and geometric photons.

and influence on image quality. In addition, optimal energy thresholds for ^{124}I were derived by calculating noise equivalent count rate curves. *In the case of SPECT* an accurate classification of contamination such as septal penetration, collimator scatter and backscatter was made. The assessment of contamination was then made for ^{124}I- and ^{131}I-Bexxar distributions in a complex torso phantom. The SPECT data was reconstructed with standard MLEM using TEW correction and attenuation correction. In addition to the normal reconstruction, a contamination free case was also considered in which penetration, collimator scatter, backscatter and phantom scatter were filtered out prior to reconstruction. The PET data was reconstructed using OSEM with sensitivity correction, attenuation correction and single scatter estimation. The same data were used for three recon-



(a) Kidney slice. (b) Kidney slice without contamination.

(c) Tumour slice. (d) Tumour slice without contamination.

Figure 3.15: Reconstructed slices of the MCAT phantom for ^{131}I-Bexxar.

struction cases: for a standard energy window, for an optimized energy window based on the NEC curves and without spurious contamination.

Our framework is essentially composed of two layers. In the first layer it is possible to select specific particle history information. This information can be stored directly or it can be processed during the simulation. The additional data is available at the end of the simulation together with regular GATE data output, which forms the second layer of our framework. The user-defined classification of detections has clearly demonstrated the flexibility of ProcessGATE. This will lead to a better assessment of the different components of contamination. Its applications range from the optimization of correction methods to the investigation of the properties of current and future

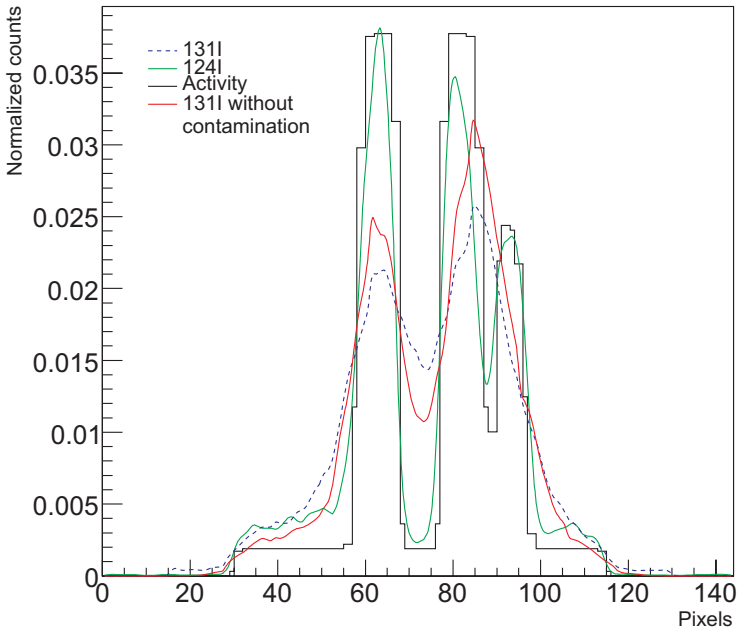


Figure 3.16: Kidney profile of the reconstructed slices shown in figure 3.15 for ^{131}I-Bexxar. The ^{131}I profiles with and without contamination are shown. In addition, the simulated activity is shown as well as the ^{124}I-Bexxar profile with optimized energy window.

detectors.

A remaining difficulty for PET imaging with a non-pure isotope is that some spurious activity still remains, despite optimized energy thresholds and single scatter estimation. Two additional problems arise as a result. Firstly, it is obvious that attenuation correction is not correctly modeled for spurious coincidences. Secondly, the single scatter algorithm adds a constant background for tail fitting while the spurious activity is not spatially uniform. A dedicated algorithm to remove spurious coincidences in combination with single scatter estimation will likely provide better results. In the case of SPECT our classification showed that TEW correction is able to partly cor-

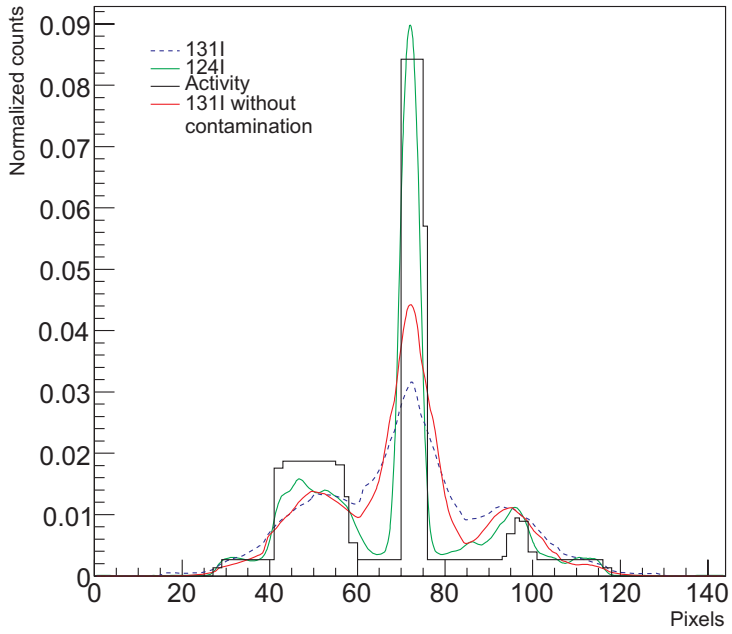


Figure 3.17: Tumour profile of the reconstructed slices shown in figure 3.15 for ^{131}I-Bexxar. The ^{131}I profiles with and without contamination are shown. In addition, the simulated activity is shown as well as the ^{124}I-Bexxar profile with optimized energy window.

rect for phantom scatter and backscatter, but not for septal penetration. The comparison of PET and SPECT in this case study shows that ^{124}I-Bexxar with simple optimized energy thresholds offer better image quality when standard reconstruction techniques are used. However, accurate modeling of collimator scatter and penetration in the reconstruction can improve image quality.

The work presented in this chapter resulted in an A1 journal paper (24) and primarily serves as the validation tool for the remaining chapters and as the basis of the efficient analysis required in chap-

ters 4 and 5. As illustrated in the case study, ProcessGATE allows a more detailed analysis of GATE results and can be used to design and evaluate contamination correction algorithms and to investigate the properties of current and future prototype detectors. As a result of this, the framework was used in 7 co-authored conference proceedings (25; 26; 27; 28; 29; 30; 31) and 4 co-authored A1 journal papers (32; 33; 34; 12).

References

[1] K.S. Pentlow, R.D. Finn, S.M. Larson, Y.E. Erdi, B.J. Beattie, and J.L. Humm. *Quantitative imaging of yttrium-86 with PET. The occurrence and correction of anomalous apparent activity in high density regions*. Clin. Positron Imaging, 3:85–90, 2000.

[2] M. Lubberink, H. Schneider, M. Bergstrom, and H. Lundqvist. *Quantitative imaging and correction for cascade gamma radiation of Br-76 with 2D and 3D PET*. Phys. Med. Biol., 47:3519–3534, 2002.

[3] S.G. Kohlmyer, R.S. Miyaoka, S.C. Shoner, T.K. Lewellen, and J.F. Eary. *Quantitative accuracy of PET imaging with yttrium-86*. J. Nucl. Med., 40:1237, 1999.

[4] S. Walrand, F. Jamar, I. Mathieu, J. De Camps, M. Lonneux, M. Sibomana, D. Labar, C. Michel, and S. Pauwels. *Quantitation in PET using isotopes emitting prompt single gammas: Application to yttrium-86*. Eur. J. Nucl. Med. Mol. Imaging, 30:354–361, 2003.

[5] S. Vandenberghe. *Three-dimensional positron emission tomography imaging with 124I and 86Y*. Nucl. Med. Comm., 27:237–245, 2006.

[6] B. He, Y. Du, W.P. Segars, and E.C. Frey. *A Monte Carlo and physical phantom evaluation of quantitative In-111 SPECT*. Phys. Med. Biol., 50:4169–4185, 2005.

[7] B. He and E.C. Frey. *Comparison of conventional, model-based quantitative planar, and quantitative SPECT image processing methods for organ activity estimation using In-111 agents*. Phys. Med. Biol., 51:3967–3981, 2006.

[8] Y.K. Dewaraja, M. Ljungberg, and J.A. Fessler. *3-D Monte Carlo-based scatter compensation in quantitative I-131 SPECT reconstruction*. IEEE Trans. Nucl. Sci., 53(1):181–188, 2006.

[9] D. Autret, A. Bitar, L. Ferrer, A. Lisbona, and M. Bardies. *Monte Carlo modeling of gamma cameras for 131I imaging in*

targeted radiotherapy. Cancer Biother. Radio., 20(1):77–84, 2005.

[10] A.N. Bica, J.F. Eary, and W.B. Nelp. *Quantification of iodine-131 distribution by gamma camera imaging*. Eur. J. Nucl. Med., 18(2):142–143, 1991.

[11] Geant4. http://geant4.web.cern.ch/geant4/UserDocumentation/, 2008.

[12] R. Van Holen, S. Vandenberghe, S. Staelens, J. De Beenhouwer, E. Rault, and I. Lemahieu. *Reduced Influence of High Energy Contamination in 123I Imaging by Using Rotating Slat Collimators*. Nucl. Instr. and Meth. A, submitted, 2008.

[13] H. Limura, K. Katakura, and T. Tamura. *Nuclear Data Sheets for A = 124*. Nuclear Data Sheets, 80(4):895–1068, 1997.

[14] Y. Khazov, I. Mitropolsky, and A. Rodionov. *Nuclear Data Sheets for A = 131*. Nuclear Data Sheets, 107(11):2715–2930, 2006.

[15] S. Surti and J.S. Karp. *Imaging characteristics of a 3-dimensional GSO whole-body PET camera*. J. Nucl. Med., 45:1040–1049, 2004.

[16] Philips Medical Systems, 595 Miner Road, Cleveland, OH 44143, USA.

[17] S. Staelens, D. Strul, G. Santin, S. Vandenberghe, M. Koole, Y. D'Asseler, I. Lemahieu, and R. Van de Walle. *Monte Carlo simulations of a scintillation camera using GATE: validation and application modelling*. Phys. Med. Biol., 48(18):3021–3042, 2003.

[18] NEMA Standards Publication NU 2-2001. *Performance measurements of positron emission tomographs*. Rosslyn, VA: National Electrical Manufacturers Association, 2001.

[19] B.M.W. Tsui, J.A. Terry, and G.T Gullberg. *Evaluation of cardiac cone-beam SPECT using observer performance experiments and ROC analysis*. Investigative Radiology, 28(12):1101–1112, 1993.

[20] Y.K. Dewaraja, S.J. Wilderman, M. Ljungberg, K.F. Koral, K. Zasadny, and M.S. Kaminiski. *Accurate dosimetry in 131-I radionuclide therapy using patient-specific, 3-dimensional methods for SPECT reconstruction and absorbed dose calculation*. J. Nucl. Med., 46(5):840–849, 2005.

[21] R. Accorsi, L. Adam, M.E. Werner, and J.S. Karp. *Optimization of a fully 3D single scatter simulation algorithm for 3D PET*. Phys. Med. Biol., 49:2577–2598, 2004.

[22] Y. Dewaraja, Jia. Li, and K. Koral. *Quantitative 131I SPECT with Triple Energy Window Compton Scatter Correction*. IEEE Trans. Nucl. Sci., 45(6):3109–3114, 1998.

[23] Y. Seo, K.H. Wong, M. Sun, B.L. Franc, R.A. Hawkins, and B.H. Hasegawa. *Correction of Photon Attenuation and Collimator Response for a Body-Contouring SPECT/CT Imaging System*. J. Nucl. Med., 46(5):868–876, 2005.

[24] J. De Beenhouwer, S. Staelens, S. Vandenberghe, J. Verhaeghe, and I. Lemahieu. *Process level discrimination for GATE: assessment of contamination in SPECT and spurious activity in PET*. Med. Phys., submitted, 2008.

[25] E. Rault, S. Vandenberghe, S. Staelens, J. De Beenhouwer, and I. Lemahieu. *Investigation of Yttrium-90 Bremsstrahlung with GATE*. In Journal of Nuclear Medicine, abstract book of the 54th SNM annual meeting, volume 48, page 422, 2007.

[26] E. Rault, S. Vandenberghe, R. Van Holen, J. De Beenhouwer, and I. Lemahieu. *Comparison of Quantification and Image Degrading Factors with Different Iodine Isotopes*. In European Journal of Nuclear Medicine, abstracts of the annual congress of the EANM, volume 33, page 374, 2006.

[27] E. Rault, S. Vandenberghe, S. Staelens, J. De Beenhouwer, and I. Lemahieu. *Investigation of Yttrium-90 Bremsstrahlung with GATE*. In Abstract book of the Workshop Quantitative Imaging and Dosimetry in Nuclear Medicine, page 25, 2007.

[28] R. Van Holen, S. Vandenberghe, E. Rault, J. De Beenhouwer, and I. Lemahieu. *Reducing High Energy Contamination in*

SPECT Using a Rotating Slat Collimator. In Book of Abstracts 4th International Conference on Imaging Technologies in Biomedical Sciences, page 65, 2007.

[29] E. Rault, S. Vandenberghe, R. Van Holen, J. De Beenhouwer, S. Staelens, and I. Lemahieu. *Comparison of Quantification and Image Quality for Different Iodine isotopes (I-123, I-124 and I-131)*. In Abstract book of the 22nd Annual BHPA Symposium on Physics in Medicine, page 20, 2007.

[30] S. Staelens, S. Vandenberghe, J. De Beenhouwer, S. De Clercq, Y. D'Asseler, I. Lemahieu, and R. Van de Walle. *Simulation study comparing the imaging performance of a solid date detector with a rotating slat collimator versus parallel beam collimator setups*. In Proceedings of SPIE Medical Imaging 2004, volume 5372, pages 301–310, 2004.

[31] E. Rault, S. Vandenberghe, R. Van Holen, J. De Beenhouwer, and I. Lemahieu. *Comparison of quantification and image degrading factors for different iodine isotopes (I-123, I-124 and I-131)*. In Abstracts 7e UGent-FirW doctoraatssymposium, page 38, 2006.

[32] S. Staelens, K. Vunckx, J. De Beenhouwer, F. Beekman, Y D'Asseler, J. Nuyts, and I. Lemhieu. *GATE simulations for optimization of pinhole imaging*. Nucl. Instr. and Meth. A, A569:359–363, 2006.

[33] E. Rault, S. Vandenberghe, R. Van Holen, J. De Beenhouwer, S. Staelens, and I. Lemahieu. *Comparison of image quality of different iodine isotopes (I-123, I-124 and I-131)*. Cancer Biotherapy and Radiopharmaceuticals, 22(3):423–430, 2007.

[34] E. Rault, S. Vandenberghe, S. Staelens, J. De Beenhouwer, and I. Lemahieu. *Fast Simulation of Yttrium-90 Bremsstrahlung with GATE*. Nucl. Instr. and Meth. A, submitted, 2008.

4

Geometrical Importance Sampling

4.1 Introduction

Importance sampling is a well-known and widely used variance reduction technique (1; 2) that attempts to sample in such a way that the number of particles sampled in a region is proportional to the importance of that region to the expected result. A region can thereby be defined in space, energy or both at the same time. In this chapter we focus on the use of geometrical importance sampling (GIS) based on Russian Roulette (RR) and particle splitting as an acceleration method for SPECT simulations with GATE.

As pointed out in chapter 2, the parallel hole collimator used in a typical SPECT system is made of highly attenuating material, such as lead, consisting of parallel holes to achieve perpendicular projection of the activity distribution onto the detector. Figures 4.1 and 4.2 show a SPECT detector head and a detail of the collimator surface respectively. Only about 0.02% of all emitted photons pass through a low energy high resolution (LEHR) collimator. Therefore it is obvious that a large quantity of photons are tracked in vain. With GIS

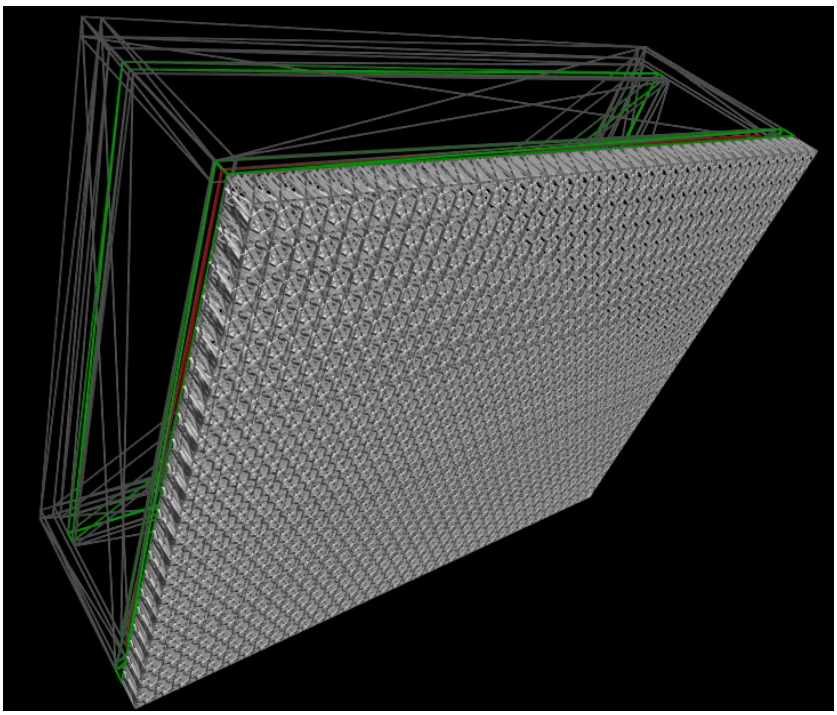


Figure 4.1: Wireframe model of a SPECT detector head with a LEHR collimator.

this problem can be ameliorated by only tracking photons with a high detection chance. The bias that is introduced in this way must then be compensated for during the scoring of those photons.

In the first part of this chapter the principle of GIS is explained and the incorporation of splitting and RR into GATE is discussed. Secondly, the introduction of branches into the particle history is discussed, along with the solution to this problem: history detangling. The third part of this chapter is focused on the validation of GIS for SPECT simulations with GATE. To this end an evaluation study is performed comparing GIS with analog GATE (aGATE): The efficiency and variance behaviour of GIS is investigated through the simulation of a set of 100 simulations with and without GIS and the introduction of a figure of merit (FOM). Subsequently, the particle history detangling approach is validated by examining the energy spectra of both a low and medium energy isotope in air and in a wa-

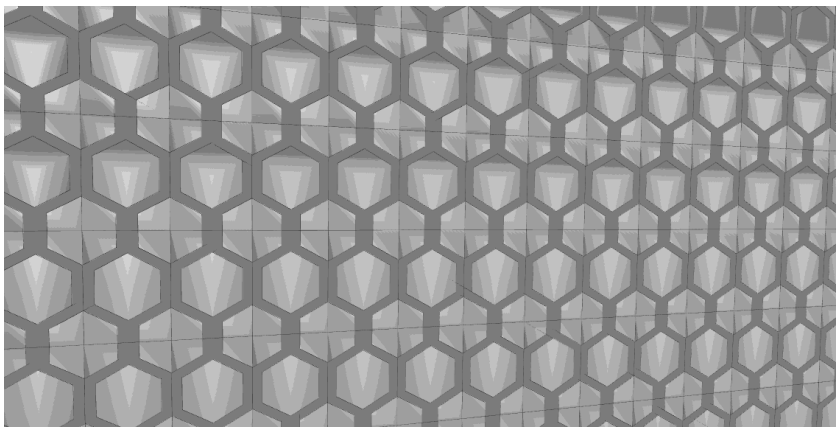


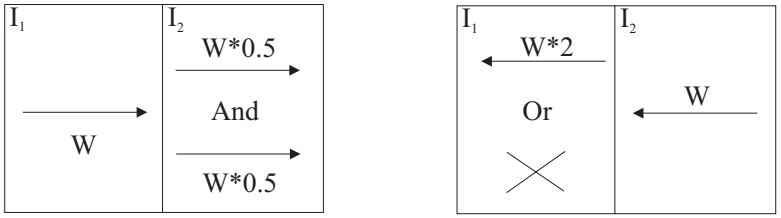
Figure 4.2: Detail of the collimator surface model.

ter phantom. Furthermore, the spatial resolution is examined and the sensitivity is checked for a medium energy point source at different depths in a water phantom.

4.2 Geometrical importance sampling

Photon paths leading towards the detector and through a collimator hole contribute to the expected result but occur infrequently. On the other hand, photon paths leading away from a detector are less likely to result in detection and occur frequently.

The solution to the first problem lies in particle splitting at certain intervals in the geometry of the problem. Photons are increasingly split into exact copies as the distance to a detector decreases. Each copy is then tracked separately through the geometry. Since no approximations have been made with regard to the physics models that apply, it is still possible to generate secondary particles. In order to avoid a bias in the result, a particle weight must be introduced that resembles the contribution of that particle to the result. The weight of a photon crossing a boundary between two regions is divided by a splitting factor that depends on the importance values of those regions. Figure 4.3a shows the simple case of 2 regions, where the region on the left has importance value I_1 while the region on the right has importance $I_2 = 2 * I_1$. A photon crossing the boundary with



(a) Splitting (b) Russian Roulette

Figure 4.3: The principle of GIS: The region on the left has half the importance of the region on the right. Figure (a) : The photon is split in two photons with half the weight. Figure (b) : The photon is either deleted or its weight is doubled after surviving RR.

weight W is therefore split into two photons with half the weight.

The second problem can be solved using RR. RR provides an alternative estimator $\langle Weight \rangle_{RR}$ of the weights of photons. Given a survival probability r and a uniform random number $\zeta \in [0,1]$, then the estimate is given as:

$$\langle Weight \rangle_{RR} = \begin{cases} \langle Weight \rangle_{aGATE} / r & \text{if } \zeta < r \\ 0 & \text{otherwise} \end{cases}$$

The photons can thus be deleted with probability $(1-r)$ in order to increase the simulation efficiency or increase their weight with probability r in order to avoid a bias in the result. An example is shown in figure 4.3b, where the photon weight is either doubled with probability $r = I_1/I_2$ or the photon is deleted.

A map consisting of all the regions with their respective importance values was constructed. The importance map used for all subsequent simulations is shown in figure 4.4. The importance map geometry is defined by a set of virtual concentric cylinder segments, with each cylinder doubling the importance value of the previous (inner) cylinder. The innermost cylinder has an importance value $I_1 = 1$ and contains a point source, while the last two outer cylinders have equal importance. This way no splitting will occur inside the collimator or crystal, as the outer cylinder covers the SPECT detector head. The virtual nature of the importance regions ensures that no photon interactions can occur when crossing boundaries. A particle history starting from a point source in the center of the cylinder

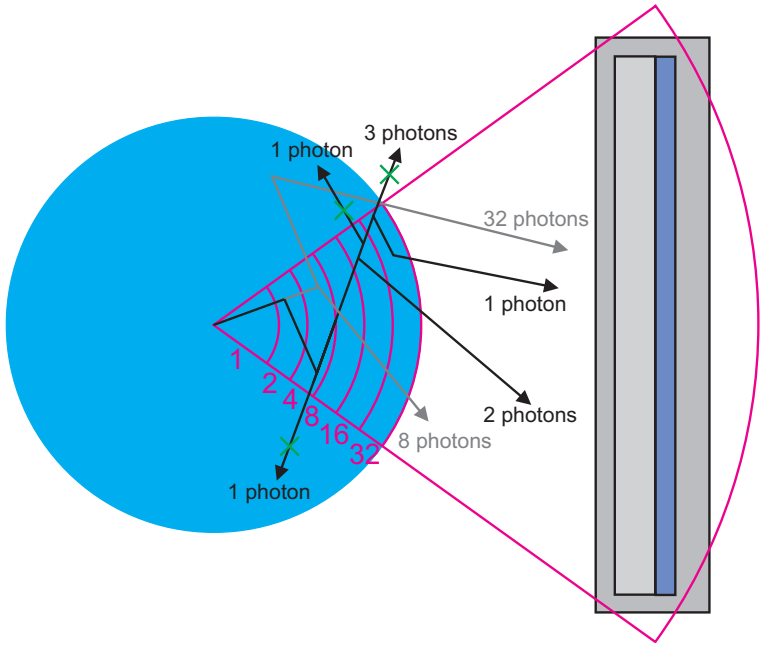


Figure 4.4: 2D view of a water phantom (cyan), a SPECT detector and the importance map with importance values (magenta) used in this study. A sample particle history (grey and black) is shown with the resulting photons after escaping the phantom and the application of RR (green).

is shown as an example. Upon reaching the region with importance value 2, the photon is split into two. In the same region, one of the two photons is subjected to Compton scatter (black path) while the other is not (grey path). The photon following the black path ultimately results in 3 photons reaching the detector. The other photons generated along this path are deleted by RR. The photon following the grey path leads to 2 photons that scatter in the region with importance value 4. One of those survives RR, resulting in 32 photons reaching the detector, while the other generates 8 photons after further splitting. The branches that are introduced this way into the particle history are discussed in the next section.

Figures 4.5 and 4.6 show the importance map in GATE in the case of three rotating SPECT detector heads. The definition and construction of the importance map has been integrated into GATE

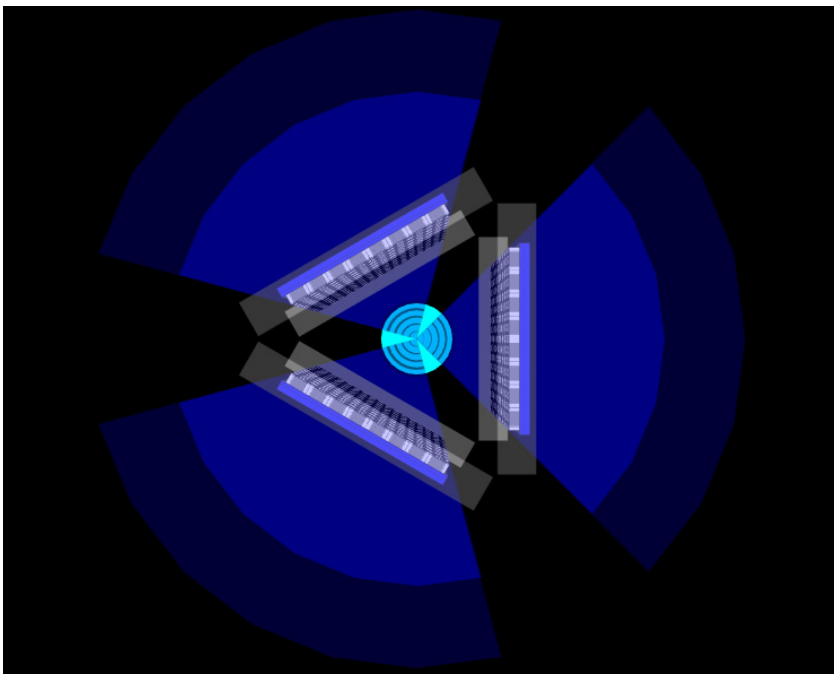


Figure 4.5: Importance map configuration for a water phantom and three SPECT detector heads.

through its script language. GATE uses an internal virtual clock to synchronise the movement and rotation of the detector heads. The virtual importance map uses the same clock so that both virtual and physical elements are synchronised during the rotation of the detector heads.

4.3 Pulse height tallies

Pulse height tallies are commonly used in medical imaging. These tallies correspond to the frequency of detected pulses in distinct energy bins, covering a relevant energy spectrum. Since each pulse is the combination of different detector hits, the tally bin is not known until the end of an event history. In an aGATE simulation this poses no problem as each detected hit can be stored and the pulse energy and centroid-weighted position can be calculated at the end of each

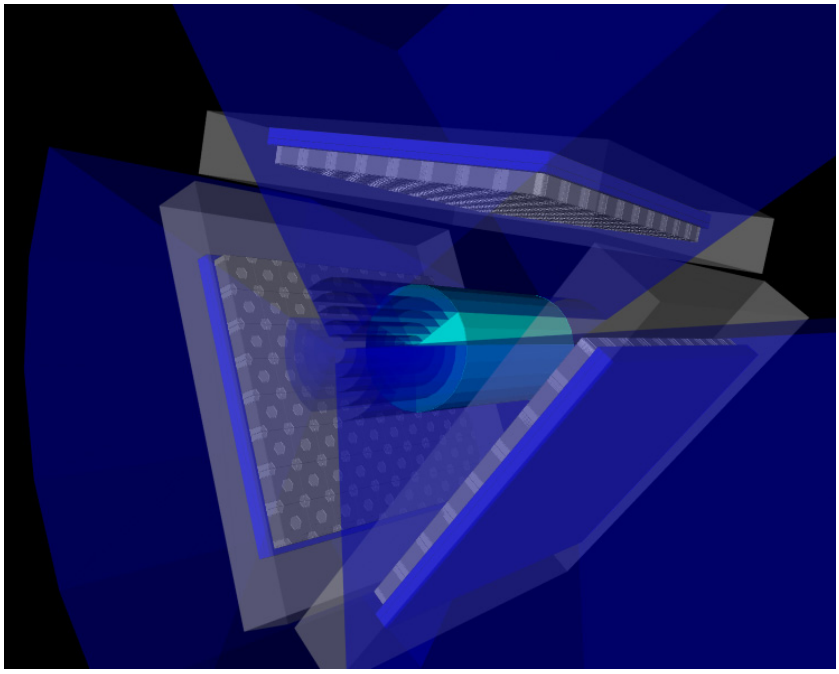


Figure 4.6: Detail of the importance map configuration for a water phantom and three SPECT detector heads.

GEANT4 event by considering all detected hits. However, as shown in the previous section, GIS introduces branches into the particle history, which results in a much more complicated pulse calculation. Simply adding all hits in a detector crystal would lead to completely wrong detection positions and energy spectrum. Therefore a new particle history has been developed within GATE. It keeps a log of all particles and their weight. At the same time a particle history tree is built by adding branches where necessary to accommodate splitting and RR. Figure 4.7 shows a simple particle history from particle t1 to t12 with both splitting (v) and physical branching (p).

Particle t1 represents the original photon. It may be subjected to Compton interactions and at some point it is split up by geometrical splitting. It should be noted that only one new photon $t2$ is generated by this process. Photons $t1$ and $t2$ now have half the weight of the original photon $t1$. Photon $t2$ is later split up into a photon and an electron as the result of a photoelectric effect. Both photon and

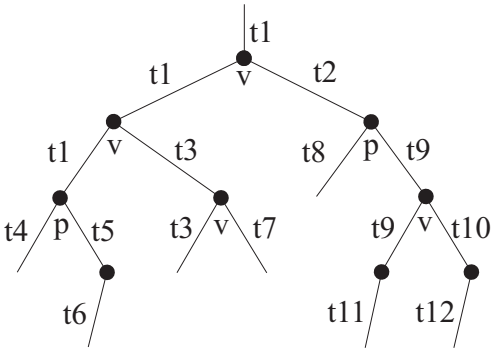


Figure 4.7: Simple particle history with branching: ”v” indicates a geometrical importance split, while ”p” indicates a physical split. Particles are marked from t1 to t12.

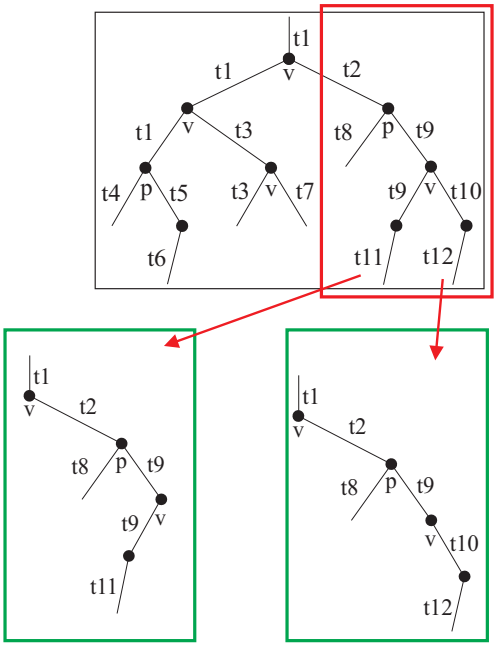


Figure 4.8: Two physically possible branches from the right side.

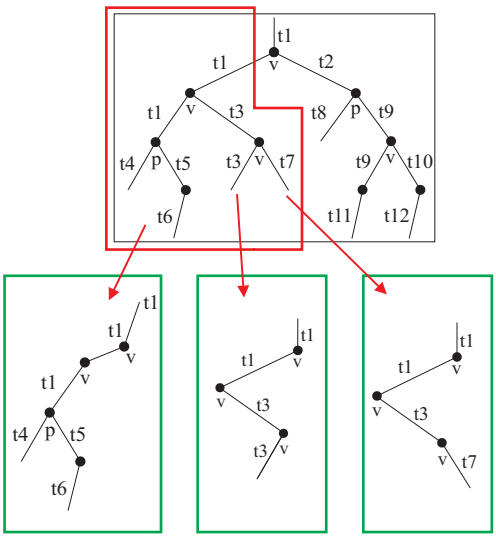


Figure 4.9: Three physically possible branches from the left side.

electron inherit the weight of their parent photon $t2$. The rest of the diagram follows the same principles. In order to retain physically meaningful pulses, one needs to extract all the physically possible paths and assign a weight to them. Each geometrical importance split introduces a new possible path. Figures 4.8 and 4.9 show the possibilities for the diagram of figure 4.7. Each path is assigned a weight according to the deconvolution approach discussed in (3). Starting at the ”leaf nodes” of the tree, we work our way up and follow each of the paths in figures 4.8 and 4.9. Whenever a geometrical importance split is encountered, we multiply the weight of the path by the splitting ratio $1/n$ with n the number of photons split into. The result for the five paths is shown in table 4.1.

It is now possible to group crystal hits together by simply looking up if they belong to the same ”physical” path and thus form an actual pulse. In this case it is possible that only paths $1, 2$ and 4 result in actual pulses, but not paths 3 and 5 in table 4.1 if there are crystal hits tied to paths $1, 2$ and 4, but not to paths 3 and 5. Note that the sum of weights of all paths in a track history tree equals 1, thus the total weight is conserved when RR is not applied.

A photon can increase its weight by surviving RR. In this case it could be scattered toward a region of higher importance and be sub-

jected to further splitting. Figure 4.10 shows how this is dealt with. The particle history adds the same photon $t1$ with its new weight 1 - as it is doubled in this case by RR - as a physical branch. This "new" photon acts as a parent for the photons produced by the splitting which follows in this case and represents in fact the same physical photon $t1$ but with a doubled weight. The total weight per history is not conserved in this way. However, RR provides an unbiased estimate of the total weight over a large number of histories as the expected value is the same as for aGATE:

$$E\left[\langle Weight\rangle_{RR}\right] = (1-r)\cdot 0 + r\cdot\frac{E\left[\langle Weight\rangle_{aGATE}\right]}{r}$$

Path	Tracks	Weight
1	$t11,t9,t8,t9,t2,t1$	$0.5*0.5=0.25$
2	$t12,t10,t8,t9,t2,t1$	$0.5*0.5=0.25$
3	$t3,t3,t1,t1$	$0.5*0.5*0.5=0.125$
4	$t6,t5,t4,t1,t1,t1$	$0.5*0.5=0.25$
5	$t7,t3,t1,t1$	$0.5*0.5*0.5=0.125$

Table 4.1: The five paths with their respective weight. The sum of weights: $0.25+0.25+0.125+0.25+0.125=1$

4.4 Evaluation study

4.4.1 Efficiency estimation and variance behaviour

Two FOMs were used to investigate the efficiency of GIS versus aGATE. The FOMs were applied to the flux D/N on the detector, with D the number of detected photons and N the number of simulated histories. In the case of aGATE the detections are scored as 1 or 0, while in the case of GIS the weight is scored. A relative FOM, shown in equation 4.1, was then used as an indicator of relative efficiency.

$$RFOM = \frac{FOM_{GIS}}{FOM_{aGATE}} \tag{4.1}$$

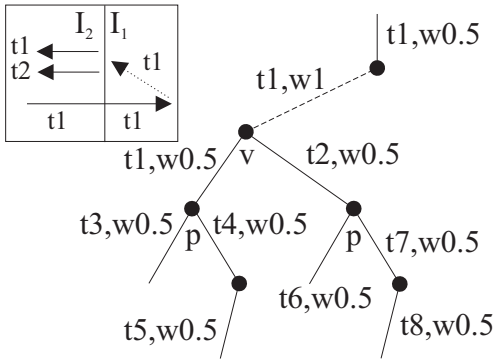


Figure 4.10: In the upper left corner a track $t1$ is shown, surviving Russian Roulette and then scattered towards a region of higher importance. The right side of this figure shows how this is logged in the track history. The additional track $t1$ with its new weight is shown as a dashed line.

Firstly, a set of 100 GIS simulations was compared with a set of 100 aGATE simulations. Each simulation consisted of a 3.5 MBq uniform ^{99m}Tc source in a 2 mm thin cylinder with a radius of 25 mm filled with water. A realistic detector setup as in section 3.4 of chapter 3 was used with a LEHR collimator. The acquisition time was 30 seconds. A 1 mm pixel size was chosen for the projection bins and a region of interest (ROI) of 30 mm x 30 mm was selected in each resulting projection for the photopeak window. The flux through each pixel in the ROI for a single simulation $\hat{x}_{pixel}$ was calculated as in equation 4.2, where $x_{i,pixel}$ is the weight of a detection in a pixel and N is the number of simulated histories. The mean flux $\hat{x}_{pixel}$ and the corresponding variance $S^2_{x_{pixel}}$ for each pixel in the ROI over the 100 simulations was calculated as shown in equation 4.3, for both aGATE and GIS. This allowed the calculation of a FOM for both cases as defined in equation 4.4, where T is the simulation time and R the relative error. Furthermore we compared the mean variances over all pixels in the ROI of both cases and scaled down the number of GIS events until the mean variances agreed. Finally it was possible to compare the distribution of the variance around the mean variance over the ROI, and the distribution of the mean (in each pixel over the 100 samples) around the mean over the ROI between aGATE and

GIS.

$$\hat{x}_{single,pixel} = \frac{1}{N}\sum_{i=1}^{N} x_{i,pixel} \tag{4.2}$$

$$\hat{x}_{pixel} = \frac{1}{100}\sum_{i=1}^{100} \hat{x}_{single,pixel} \tag{4.3}$$

$$S^2_{x_{pixel}} = \frac{1}{99}\sum_{i=1}^{100}(x_{pixel} - \hat{x}_{pixel})^2$$

$$FOM_{batch} = \frac{1}{R^2\,T} \;\; with \; R = \frac{S_{x_{pixel}}}{\hat{x}_{pixel}} \tag{4.4}$$

Secondly, a FOM based on the number of detections per second times the quality factor (QF) (2) was calculated as shown in equation 4.5. The QF indicates how much the number of detections per second is deflated due to variability in the weights. In the case of GIS with a simple importance map however, the weights are all equal. Therefore, the QF reduces to 1 as shown in equation 4.6.

$$FOM_{QF} = \frac{detections_{GIS}/s}{QF \quad detections_{aGATE}/s} \tag{4.5}$$

$$QF = \frac{(\sum_{i=1}^{N} x_i)^2}{D\sum_{i=1}^{N} x_i^2} = \frac{(Dx_{i=1})^2}{D^2 x_{i=1}^2} \tag{4.6}$$

4.4.2 Energy spectra evaluation

Energy spectra are a first indication of validity of the particle history detangling approach used for splitting and RR. Identical configurations were used each time for the GIS and the aGATE simulation. First the energy spectrum of both ^{99m}Tc and ^{67}Ga in air was compared with aGATE. A ^{99m}Tc point source with a radius of 1.5mm was used at a distance of 25cm from the collimator. The activity of the source was 150 MBq for GIS and 300 MBq for aGATE, with a total acquisition time of 30 seconds. The setup consisted of the same detector head as defined above. The LEHR collimator was replaced by a medium energy general purpose (MEGP) collimator for the simulations with ^{67}Ga. The point source with a radius of 1.5mm

was placed at a distance of 25 cm from the collimator. The activity of the source was 36 MBq for GIS and 250 MBq for the aGATE, with an acquisition time of 30 seconds. The activity values were chosen to result in a comparable number of detections, based on the number of detections per second during a short test run. As a final verification, the ^{67}Ga point source was placed in the middle of a water phantom made of a cylinder with radius 12 cm and height 34.56 cm. An equal activity of 100 MBq for both cases was used over an acquisition time of 30 seconds. In this case the activity values were chosen equal in order to verify the improved statistics with GIS.

4.4.3 Spatial resolution validation

The spatial resolution for low and medium energy simulations was compared with aGATE. For ^{99m}Tc a 30 MBq point source with a radius of 0.5 mm was placed at 3.65, 13.65, 24.65 and 38.65 cm from the detector with LEHR collimator in the case of GIS. The aGATE simulations consisted of a 20 MBq line source with a radius of 0.5 mm at 5, 15, 26 and 40 cm from the LEHR collimator as in (4). The medium energy setups consisted of a 1.5 mm point source filled with 20 MBq ^{67}Ga and placed at 3.65, 13.65, 24.54 and 38.65 cm from the detector with MEGP collimator attached for both aGATE and GIS. For the ^{99m}Tc simulations a photopeak window at 129-151 keV was used, while for ^{67}Ga two photopeak windows were used : at 83.7-102.3 keV and at 171.1-198.8 keV.

4.4.4 Sensitivity validation

The absolute sensitivity (in cps/MBq) was evaluated for a medium energy setup. A point source filled with 20 MBq (GIS) and 100 MBq (aGATE) ^{67}Ga was placed at different depths in a cylindrical water phantom with a radius of 12 cm and height 34.65 cm : 1, 5, 10, 15 and 20 cm depth. The activity values were chosen to result in a comparable number of detections. The total acquisition time was 30 seconds and two photopeak windows were used : at 83.7-102.3 keV and at 171.1-198.8 keV. A realistic detector was used with a MEGP collimator attached.

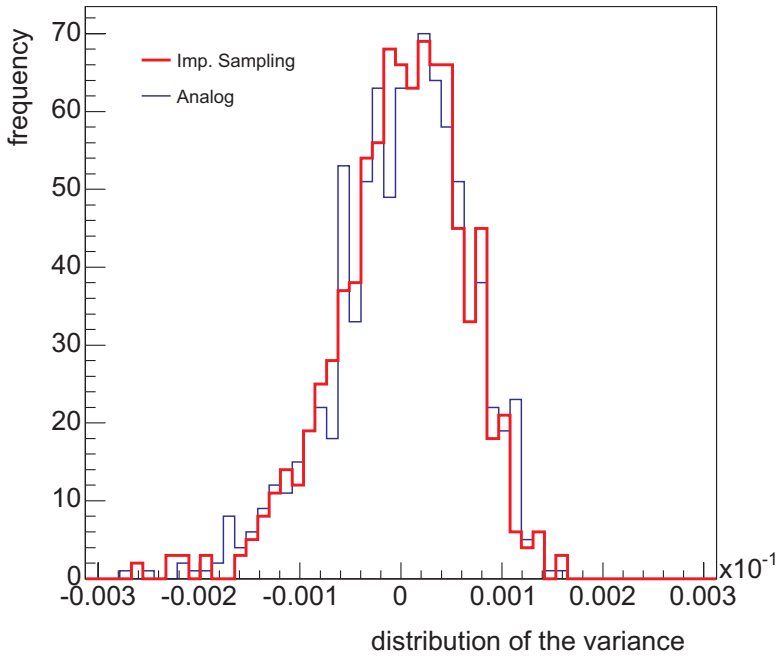


Figure 4.11: Distribution of the variance over the ROI around its mean value.

4.5 Results

4.5.1 Efficiency estimation

Figure 4.11 shows the distribution of the variance for all pixels (based on 100 simulation samples) in the ROI around its mean value for both aGATE and GIS. The distribution of the variance with GIS shows no systematic bias when compared to the analog distribution. Figure 4.12 shows the distribution of the mean for all pixels (based on 100 simulation samples) in the ROI around its mean value for both cases. Again the distribution shows no systematic bias against the analog distribution. In order to compare the distributions, a factor of approximately $RFOM_{BATCH} = 12$ was used to scale down the number of detections for GIS. The number of detections per second were 0.36 (aGATE) and 4.73 (GIS) resulting in a relative efficiency of ap-

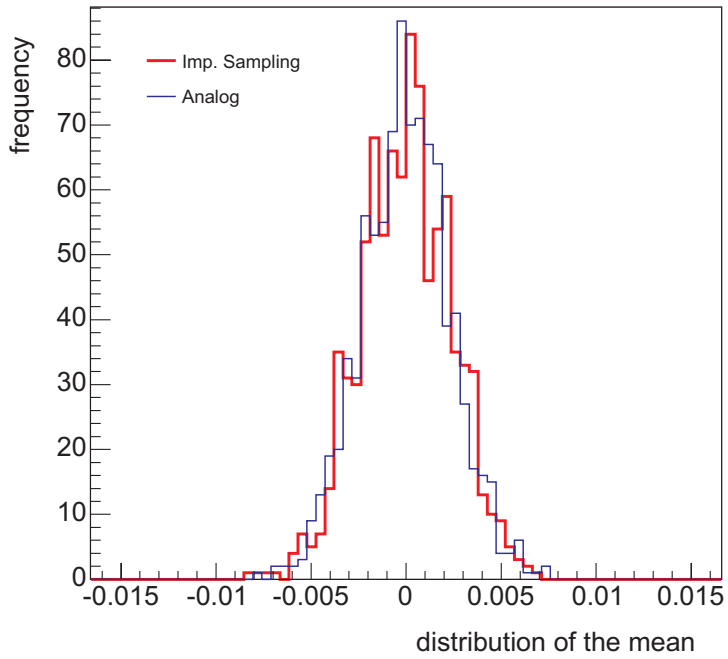


Figure 4.12: Distribution of the mean over the ROI around its mean value.

proximately $RFOM_{QF} = 13$.

4.5.2 Energy spectra evaluation

Figure 4.13 shows the energy spectrum of a ^{99m}Tc point source in air both with and without GIS. Figure 4.14 shows the same for ^{67}Ga. An excellent agreement was found between the aGATE and the GIS simulations for the spectra of both isotopes. Figure 4.15 shows the comparison between the energy spectra of aGATE and GIS simulations for a ^{67}Ga point source in a water phantom. A clearly lower variance can be observed in the case of GIS. The number of detections were 0.17 (aGATE) and 0.93 (GIS), resulting in a relative efficiency of approximately $RFOM_{QF} = 5$ in this case.

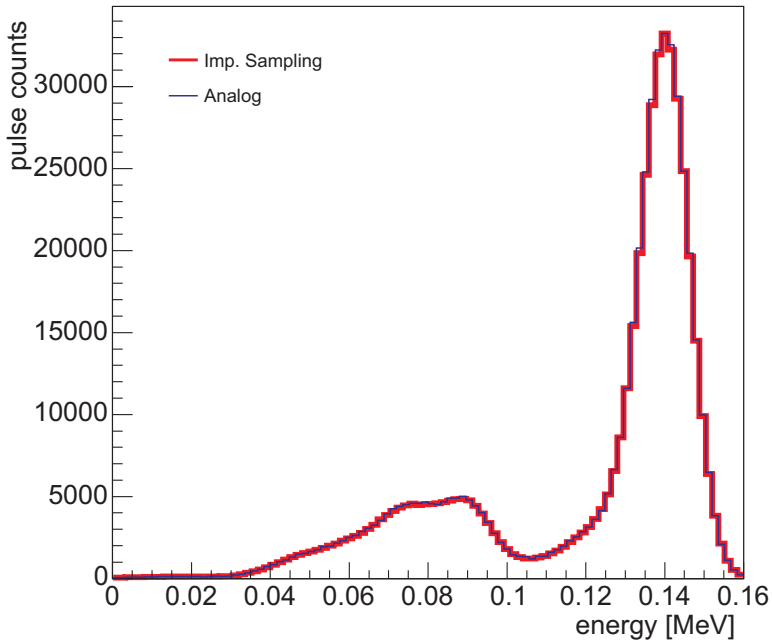


Figure 4.13: ^{99m}Tc spectra for a point source in air with and without GIS

4.5.3 Spatial resolution validation

The spatial resolution for a low energy isotope (^{99m}Tc) and a medium energy isotope (^{67}Ga) is shown in figure 4.16. A linear was drawn through the experimental values obtained from (4) and the simulation values in the aGATE case. In both the low and medium energy case a good agreement was found.

4.5.4 Sensitivity validation

Figure 4.17 shows the absolute sensitivity results for a ^{67}Ga point source in a water phantom at different depths. The results with and without GIS are in excellent agreement.

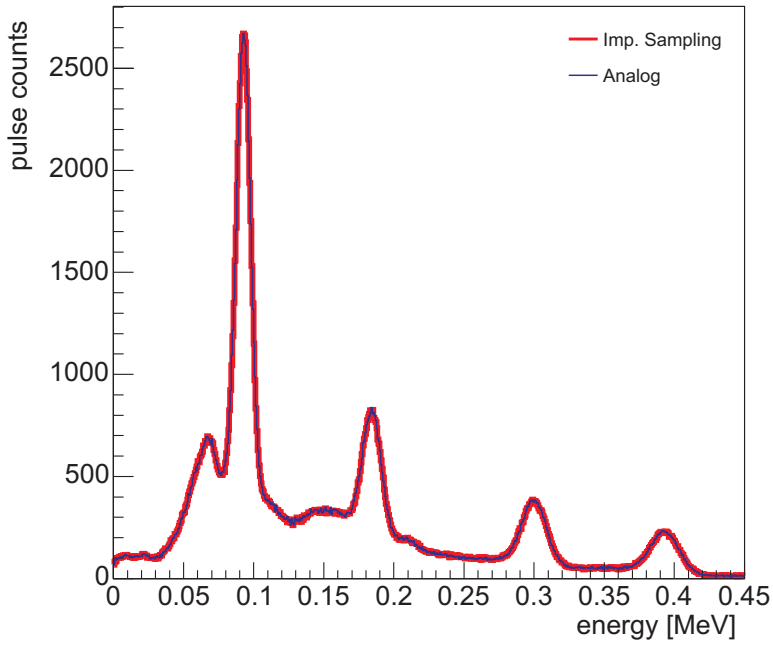


Figure 4.14: ^{67}Ga spectra for a point source in air with and without GIS

4.6 Application to PET

It has been shown previously that importance sampling can be applied to PET simulations (5; 6; 7). This is also the case for GIS. All that is required on the tracking side of the simulation is a suitable importance map. The annihilation photons will then be split accordingly. A simple extension of the particle history detangling approach would not be sufficient however. In contrast to other codes, GATE explicitly models time throughout the simulation. A direct translation to GATE would therefore pose problems for time-dependent phenomena such as detector dead time, time-of-flight, count rate, randoms rate and time-based coincidence sorting. As a result, only the optimization of photon tracking has been proposed for GATE PET (8), while no PET variance reduction techniques have been introduced. Several model-based methods could be considered to solve

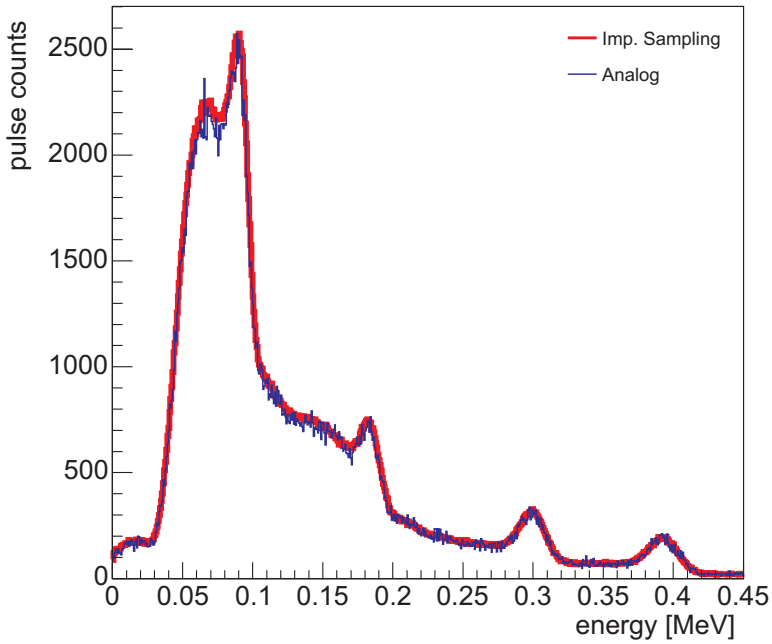


Figure 4.15: ^{67}Ga spectra for a point source in a water phantom with and without GIS

this problem.

In (9), SimSET was used to simulate true, scatter and singles rates. From that data, randoms rates and dead time were modeled. Dead time was modeled simply as a paralyzing exponential with a characteristic time of 1 μsecond obtained by measuring block dead time at moderate count rates. It was reported that the model will underestimate the dead time in a real system. A total coincidence time-window of 12.5 nano seconds was used to generate randoms.

In (10), SimSET was modified to include modeling of block detectors and crystals, random coincidences and detector dead time. Importance sampling was used for tracking photons in a phantom. Coincidence events could consist of more than two photons because of the use of importance sampling. In that case, all photons were propagated independently in the detector and every possible combination of two photons originating from the same annihilation were

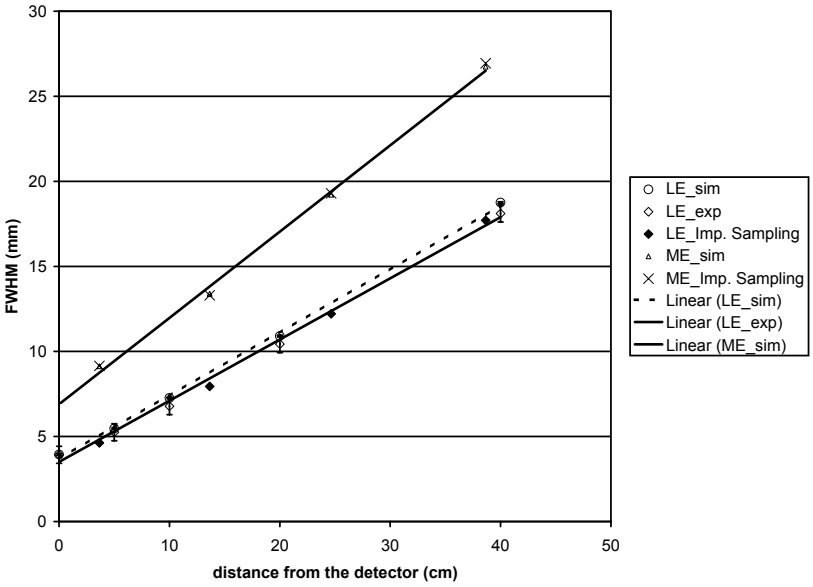


Figure 4.16: Spatial resolution comparison for low and medium energy setups. For ^{99m}Tc both simulations (with and without GIS) and experimental values are shown. For ^{67}Ga the results with and without GIS are shown. A linear was drawn through both simulation (aGATE only) and experimental values.

considered as a valid coincidence. A statistical weight equal to the product of the two photons' weights was associated with each coincidence, since the photon histories were independent. Dead time was modeled according to (11) along with the dead time free photon fluence estimated in a small run to calculate the dead time fractions in each block. Instead of randomly rejecting photons according to the computed dead time fractions, a variance reduction technique was used: all events were detected, but a statistical weight was used in order to reflect the probability of not being detected due to dead time. Since time was not explicitly modeled, it was not possible to include time-stamps with the simulation output. Instead, coincidences were simulated sequentially and photons pertaining to the same coincidence were coupled. As a result, random coincidences could also not be estimated explicitly, but were derived from single rates.

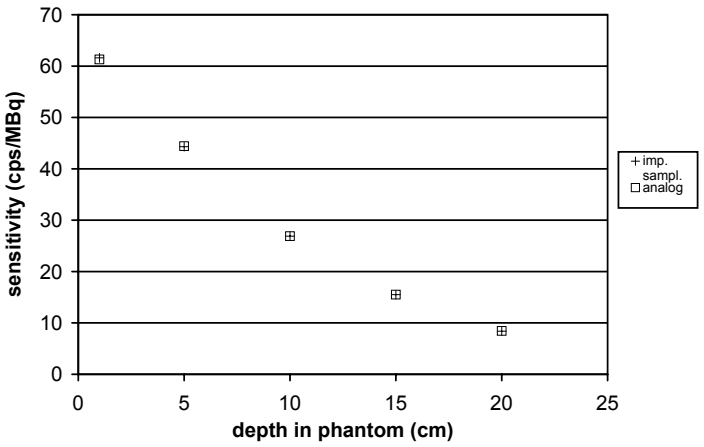


Figure 4.17: Absolute sensitivity for a ^{67}Ga point source in a water phantom at different depths in the phantom. Results for simulations with and without GIS are shown.

In (12), a similar approach was used for PET-SORTEO (13), also based on (11). Electronic dead time losses at different stages of the detection chain were modeled, based on parametric models that were fitted to experimental data. In (5), vmcPET was developed with support for Russian roulette, interaction splitting, directional biasing and other variance reduction techniques. Every possible combination of two photons originating from one or more annihilations within a considered time interval were considered as a valid coincidence. VmcPET was also coupled to GATE through a phase space file (5). In that case, no variance reduction techniques were allowed because of the time-based coincidence sorting in GATE. This might be solved in the future through complete replacement of the coincidence sorting in GATE by a parametric model in which only a series of time slices is considered. A Poissonian, time-dependent probability distribution can then be used to sample the number of decays per time slice.

The advantage of model-based solutions is undoubtedly increased performance over time-based solutions. The downside is that it limits the applicability of GATE to the models that are supported. To date no time-based solutions that introduce statistical weights throughout the GATE digitizer chain have been proposed however. Possibilities include a variance reduction technique for dead time such as proposed in (10) or variable length dead time windows based on statistical weights. Time-based solutions will likely require more computation time but can be applied to any detector model supported in GATE.

4.7 Discussion

In theory, the maximum efficiency of this technique is inversely related to the sensitivity of the detector. The larger septa of a MEGP collimator result in a larger angle of acceptance compared to a LEHR collimator. This allows for a larger flux through the collimator holes and reduces the benefit of particle splitting. The tracking overhead resulting from the increased number of particles to be tracked is besides the sensitivity the most important efficiency limiting factor. Our investigation of efficiency showed that the acceleration factors obtained for $RFOM_{BATCH}$ and $RFOM_{QF}$ were approximately equal. Therefore it can be concluded that $RFOM_{QF}$, which is based on a single simulation, can be used as an indication of efficiency for this particular type of GIS. Despite the detangling of each detected photon history and the increased tracking overhead, GIS resulted in a 5 to 13-fold increase over aGATE SPECT simulations.

The distribution of the variance and mean values revealed no systematic bias for GIS when compared to aGATE. The results shown for the energy spectra verified that the detangling of the photon histories results in correct pulse height tallies. The spectrum in figure 4.15 is a typical example of how GIS results in lower variance compared to aGATE, for an equal amount of simulated events. The spatial resolution and the sensitivity results all show an excellent agreement compared to aGATE simulations. This is the strength of GIS: its generality. It provides a moderate increase in efficiency but for any combination of SPECT isotope, phantom and detector.

The method of binary splitting used in these simulations combined with the relatively simple layout of the importance map re-

sulted in particles with equal weight. The weight only depends on the importance region where the particle ends up and is independent from the path by which the particle reached that final region. A more complex importance map based on production values could be used instead. Production values can be automatically obtained by a reversed simulation, in which the source and the detector are switched. This would likely produce a weight distribution depending on the complexity of the importance map. A weight distribution resulting from either a more complex importance map or a different splitting algorithm could increase the efficiency further. Although GIS can be used for any SPECT detector, the automatic calculation of an importance map would improve the efficiency of GIS in system design. Care has to be taken however as increased efficiency comes with higher risks compared to the current conservative approach: a higher degree of variation of the weights in a single importance region increases the variance of the tallies in that region.

4.8 Summary and original contributions

In this chapter we presented the use of GIS for SPECT. GIS is a variance reduction technique that assigns importance values to virtual regions of the simulation space. A photon crossing the boundary between two regions is then either split into multiple photons with a lower weight, or it may be subjected to a game of RR. Using this technique it is possible to increase the survival rate of photons in regions close to the SPECT detector head, and to decrease the survival rate of photons traveling away from the detector head.

Photon splitting and RR lead to a more complicated calculation of pulse height tallies. These tallies correspond to the frequency of detected pulses in distinct energy bins, covering a relevant energy spectrum. Since each pulse is the combination of different detector hits, the tally bin is not known until the end of an event history. In an analog simulation (without variance reduction) this poses no problems as each detected hit can be stored and added to a pulse at the end of each event. GIS however introduces branches into the particle history. In order to solve this problem we developed a particle history tree that includes adjustments for splitting and RR in GATE.

We validated GIS with regard to energy spectra, spatial resolution, and sensitivity for a low (^{99m}Tc) and medium (^{67}Ga) energy

isotope in air and in a water phantom. The results were compared with aGATE simulations and a good agreement was found. The variance of the flux through a detector was also estimated with a large set of simulations in order to compare its behaviour with and without GIS. A figure of merit was thus verified and it was shown that despite the detangling and increased tracking overhead, this technique can result in a 5 to 13-fold increase in efficiency over aGATE simulations.

The work presented in this chapter resulted in several conference contributions (14; 15; 16; 17; 18). Part of this work also appeared in an A1 journal paper (19).

References

[1] F.B. Brown et al. *MCNP Version 5*. Trans. Am. Nucl. Soc., 87:273–381, 2002.

[2] D. Haynor, R. Harrison, and T. Lewellen. *The use of importance sampling techniques to improve the efficiency of photon tracking in emission tomography simulations*. Med. Phys., 18:990–1001, 1991.

[3] T.E. Booth. *Monte Carlo Variance Reduction Approaches for Non-Boltzmann Tallies*. Los Alamos LA-12433-UC-910, 1992.

[4] S. Staelens, D. Strul, G. Santin, S. Vandenberghe, M. Koole, Y. D'Asseler, I. Lemahieu, and R. Van de Walle. *Monte Carlo simulations of a scintillation camera using GATE: validation and application modelling*. Phys. Med. Biol., 48(18):3021–3042, 2003.

[5] I. Kawrakow, K. Mitev, G. Gerganov, J. Madzhunkov, and A. Kirov. *Efficient photon transport in positron emission tomography simulations using VMC++*. J. Phys.: conference series, 102, 2008.

[6] R.L. Harrison, S. Dhavala, P.N. Kumar, R. Yiping Shao Manjersshwar, T.K. Lewellen, and F.P. Jansen. *Acceleration of SimSet photon history generation.* In Nuclear Science Symposium Conference Record, volume 3, pages 1835–1838, 2002.

[7] R.L. Harrison, S. Dhavala, P.N. Kumar, Y. Shao, and T.K. Lewellen. *Importance sampling in PET collimator simulations*. In Proceedings of the IEEE Nuclear Science Symposium and Medical Imaging Conference, pages 2559–2561, 2004.

[8] N.S. Rehfeld, S. Stute, J. Apostolakis, M. Soret, and I. Buvat. *Optimization of photon tracking in GATE*. Med. Phys., submitted, 2008.

[9] R.D. Badawi, S.G. Kohlmyer, R.L. Harrison, S.D. Vannoy, and T.K. Lewellen. *The effects of camera geometry on singles flux, scatter fraction and trues and randoms sensitivity for cylindrical 3D PET - a simulation study*. In Proceedings of the

IEEE Nuclear Science Symposium and Medical Imaging Conference, pages 1522–1526, 2000.

[10] B. Guerin and G. El Fakhri. *Realistic PET Monte Carlo simulation with pixelated block detectors, light sharing, random coincidences and dead-time modeling*. IEEE Trans. Nucl. Sci, 55:942–952, 2008.

[11] C. Moisan, J.G. Rogers, and J.L. Douglas. *A count rate model for PET and its applications to an LSO HR PLUS scanner*. IEEE Trans. Nucl. Sci, 44:1219–1224, 1997.

[12] C. Lartizien, C. Kuntner, A.L. Goertzen, A.C. Evans, and A. Reilhac. *Validation of PET-SORTEO Monte Carlo simulations for the geometries of the MicroPET R4 and Focus 220 PET scanners*. Phys. Med. Biol., 52:4845–4862, 2007.

[13] A. Reilhac, G. Batan, C. Michel, C. Grova, J. Tohka, D. Collins, N. Costes, and A. Evans. *PET-SORTEO: validation and development of database of simulated PET volumes*. IEEE Trans. Nucl. Sci, 52:1231–1328, 2005.

[14] J. De Beenhouwer, S. Staelens, M. Dressel, Y. D'Asseler, S. Vandenberghe, I. Lemahieu, and R. Van de Walle. *Geometrical Importance Sampling and Pulse Height Tallies in GATE*. Proceedings of the 26th Annual International Conference of the IEEE EMBS, 26:1349–1352, 2004.

[15] J. De Beenhouwer, S. Staelens, A Goedicke, B. Schweizer, Y. D'Asseler, and I. Lemahieu. *Validation of importance sampling in GATE*. In Proceedings of the 3rd European Medical and Biological Engineering Conference, page 5 pp, 2005.

[16] J. De Beenhouwer, S. Staelens, and Y. D'Asseler. *Rotational Importance Maps in GATE*. In Abstracts of the 5th FTW PhD symposium, page 2 pp, 2004.

[17] J. De Beenhouwer and S. Staelens. *Variance reduction techniques for Monte Carlo simulations in nuclear medical imaging*. In Abstracts of the 4th FTW PhD symposium, 2003.

[18] S. Staelens, J. De Beenhouwer, M. Koole, Y. D'Asseler, S. Vandenberghe, I. Lemahieu, and R. Van de Walle. *Monte Carlo simulations for SPECT: object oriented flexibility benchmark*. In Book of abstracts of the 3rd Belgian Day on Biomedical Engineering, page 30, 2003.

[19] S. Staelens, J. De Beenhouwer, D. Kruecker, L. Maigne, F. Rannou, L. Ferrer, Y. D'Asseler, I. Buvat, and I. Lemahieu. *GATE: Improving the computational efficiency*. Nucl. Instr. and Meth. A, 569:341–345, 2006.

5

Forced Detection and Convolution-based Forced Detection

5.1 Introduction

GATE is one of the most accurate and versatile Monte Carlo codes for PET and SPECT but unfortunately also one of the slowest. In chapter 4 we introduced geometrical importance sampling as an acceleration technique for GATE SPECT simulations. Although it is useful as a safe and accurate acceleration technique, it can not deliver clinically acceptable simulation times. In this chapter we accelerate GATE further with a combination of two techniques.

Firstly, we implemented both standard forced detection (FD) and convolution-based forced detection (CFD) with multiple projection sampling (MPS) in GATE. Variance reduction techniques such as forced detection have previously been applied successfully to other Monte Carlo codes in order to speed up simulations (1; 2; 3; 4). Recently, convolution-based forced detection was introduced (5). CFD

enables the fast simulation of ^{99m}Tc SPECT by combining forced detection with an analytical detector module which doesn't include septal penetration or collimator scatter. Therefore it can be used as a forward projector in a fully 3D reconstruction framework as described in (6). In addition, CFD was extended with multiple projection sampling (7), which allows the simulation of all projections simultaneously.

Secondly, as GATE is built as an upper layer to GEANT4, only a small subset of its capabilities is required by GATE. In particular, the low energy extensions and the navigator for particle tracking are used intensely. The low energy extensions are very detailed and consequently very slow. In addition, the navigator is not optimized for the geometry and phantoms used in clinical acquisitions. Therefore, an FD and CFD specialized GEANT4 navigator was developed to accelerate the detailed but slow tracking algorithms in GEANT4.

We will first describe how FD and CFD with multiple projection sampling are implemented in GATE. Secondly, we describe the modifications applied to the GEANT4 code in order to increase its efficiency. Thirdly, we present an overview of the evaluation study used for the validation of the algorithms. Finally, the calculation of the acceleration factors is explained.

5.2 Forced detection methods

5.2.1 Standard and convolution-based forced detection

The FD and CFD implementations share a common design. The FD with MPS principle is illustrated by figure 5.1. At decay, the mother particle is created. The path of the photon to the next interaction point is sampled by an analog GATE (aGATE) Monte Carlo simulation. The photon can scatter in the phantom, it can be absorbed or it can leave the phantom directly. As soon as the photon leaves the phantom, it is destroyed. In the case of figure 5.1, the particle scatters once before leaving the phantom. Photon copies are created for each projection angle at decay, and at every Compton or Rayleigh interaction along the mother photon path. The emission angle of the photon copies is selected within a solid angle subtended by a cone with the central axis perpendicular to the detector. The weight of the

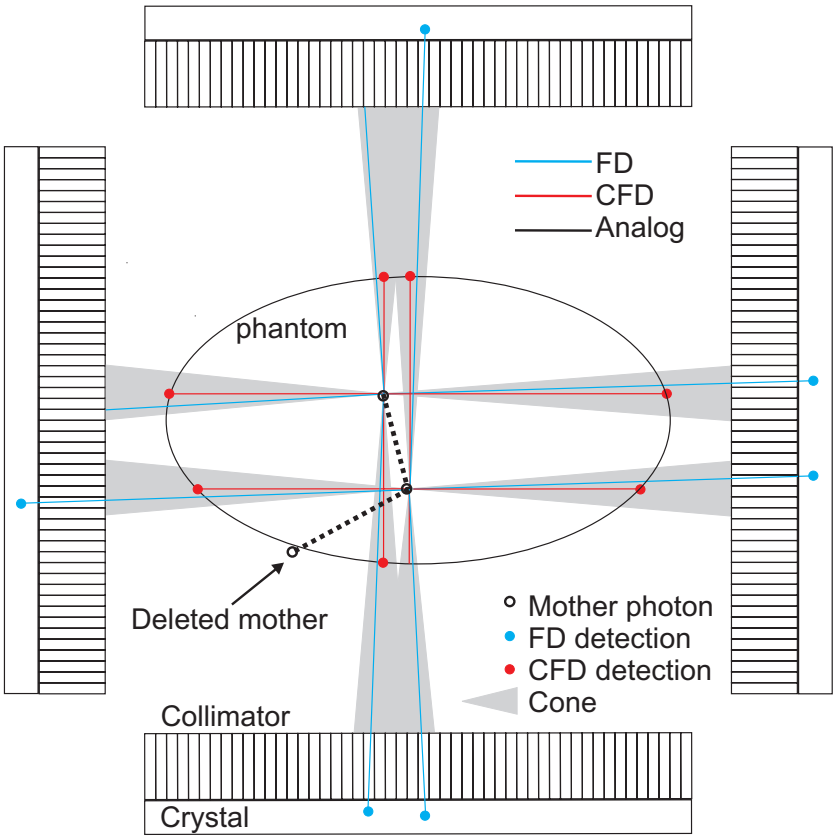


Figure 5.1: Principle of FD and CFD for four projections. The setup consists of four detectors and a simple phantom. The path of a mother particle tracked by analog GATE is shown. At decay and at the interaction site, FD and CFD photon copies are formed and their respective paths within the FD cone are shown.

copy is then adjusted to reflect the forced angle and the probability of non-interaction along its path (cfr. section 5.2.2). Each copy is then transported from the interaction point to the edge of the phantom. From that point on they are tracked further by analog Monte Carlo in order to model the detector interactions, including the generation of secondary particles. As shown by figure 5.1, the four grey cones at decay and at the scatter position indicate the solid angles. Two decay photon copies and all four scatter photon copies are finally detected in this example.

The CFD with MPS principle differs only on two accounts compared to FD. Firstly, the emission angle of the photon copies at decay and then at every Compton or Rayleigh interaction point is into a direction exactly perpendicular to the detector surface according to the central ray approach proposed in (8). Secondly, the detector is not simulated in the case of CFD but defined analytically as proposed by (8). Therefore, when the photon copies leave the phantom they are automatically detected and stored in depth dependent subprojection maps. At the end of the simulation, the projection maps are convolved with a blurring kernel and summed up to form the final projections as in (8). Since the detector is not simulated by aGATE, the generation of secondary particles is disabled with CFD.

5.2.2 Weight calculations

The weight of each photon copy needs to be adjusted at decay and at each Compton or Rayleigh interaction for (1) forced emission or scatter into a solid angle and (2) non-interaction and non-absorption between the starting point or scatter point and escape from the phantom as shown in equations 5.1 to 5.3. Here, w_p, w_c and w_r are the weight corrections for primary, Compton scattered and Rayleigh scattered photons. Contrary to other codes based on water equivalent depth (9), no approximations were used to sample the attenuation along the photon path. Instead, we chose to highly optimize the navigation code of GEANT4 as will be explained in section 5.3. The weight corrections are therefore compliant with the low energy model of GEANT4 which includes Hubbell's atomic form factor (10).

$$w_p = e^{-\int_{start}^{escape} \sigma_t(x,E)\,dx} \frac{\Omega}{4\pi} \tag{5.1}$$

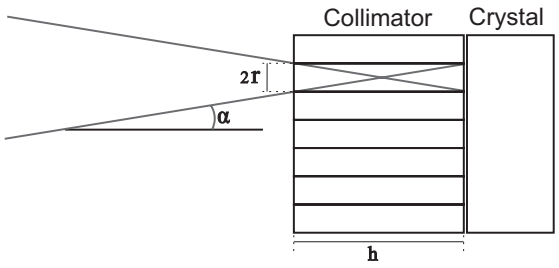


Figure 5.2: Parameters defining the half opening angle of the forced detection cone.

$$w_c = e^{-\int_{scatter}^{escape} \sigma_t(x,E)\,dx}\,\sigma_{c,SF}^{-1} \int_\Omega SF \frac{d\sigma_c}{d\Omega}\,d\Omega \qquad (5.2)$$

$$w_r = e^{-\int_{scatter}^{escape} \sigma_t(x,E)\,dx}\,\sigma_{r,FF^2}^{-1} \int_\Omega FF^2 \frac{d\sigma_r}{d\Omega}\,d\Omega \qquad (5.3)$$

At decay the emission of photons is assumed isotropic. Equation 5.1 shows the $\frac{\Omega}{4\pi}$ correction for emission into a determined solid angle Ω. The compensation for non-interaction and non-absorption between the starting point and escape from the phantom is also shown in the same equation, where $\sigma_t(x,E)$ is the total cross section at position x and energy E.

Equations 5.2 and 5.3 describe the weight corrections for Compton and Rayleigh interactions respectively. The compensation for non-interaction and non-absorption between the scatter point and escape from the phantom is shown, as well as the weight correction for scattering into a solid angle. Here, $\sigma_{c,SF}$ and σ_{r,FF^2} are the total Compton and Rayleigh cross sections also accounting for the scattering function and the squared form factor respectively. $\frac{d\sigma_c}{d\Omega}$ and $\frac{d\sigma_r}{d\Omega}$ are the differential Compton and Rayleigh cross sections per electron. The integrals in equations 5.2 and 5.3 are numerically integrated and stored in element, energy and angle dependent lookup tables. In the case of FD, the opening angle depends on collimator geometry. Figure 5.2 shows the parameters that determine the half opening angle α

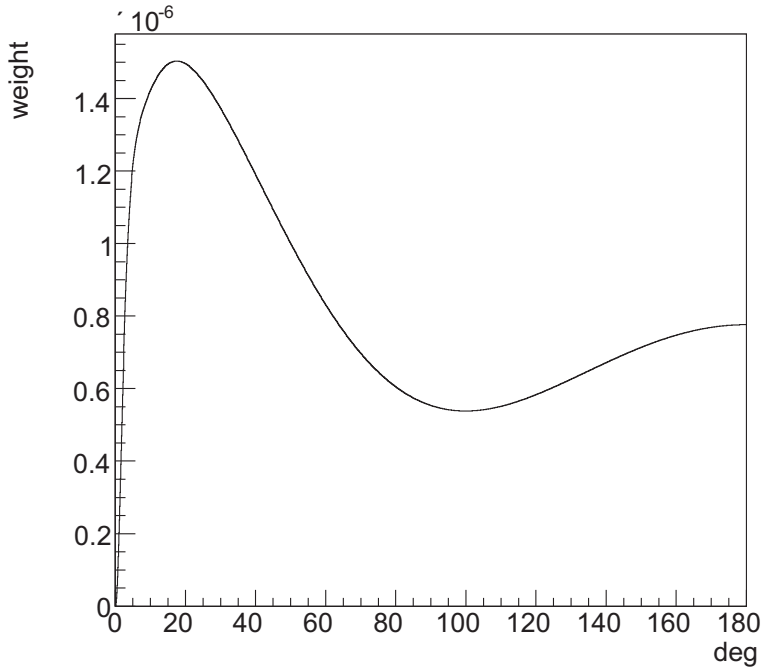


Figure 5.3: Compton scatter lookup tables for scattering into a solid angle (hydrogen at 140 KeV)

of the forced detection cone. The half opening angle depends on the collimator hole inner radius r and the height of the collimator septa h and is calculated as $\alpha = \arctan \frac{2r}{h}$. For CFD the cone half opening angle is chosen very small. The typical form of the curves stored in the lookup tables is shown in figures 5.3 and 5.4, in the case of hydrogen, for an incoming photon energy of 140 keV. The tables show the weights for scatter angles between 0 and 180 degrees, in steps of 0.1 degree. Figure 5.3 shows how the scattering function suppresses forward scattering. This is reflected in the lower weights at small scatter angles. Figure 5.4 shows how the squared form factor influences the weights in the case of Rayleigh scatter. Rayleigh scatter is strongly forward peaked at low energies, which is reflected in the high weights at small scatter angles shown in figure 5.4. The inclusion of Rayleigh scatter will therefore potentially introduce large

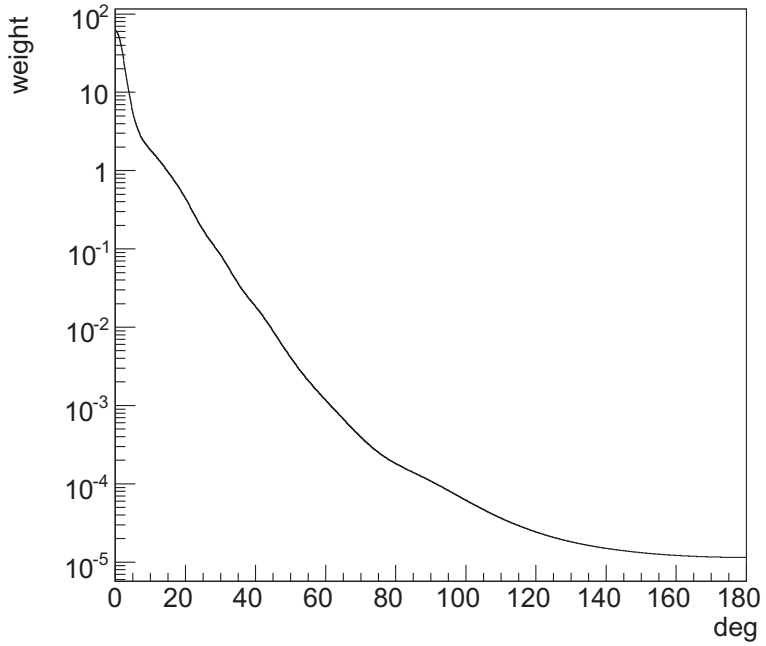


Figure 5.4: Rayleigh scatter lookup tables for scattering into a solid angle (hydrogen at 140 KeV)

weight variations, which in turn lead to increased variance. Due to this reason, the inclusion of Rayleigh scatter is optional for FD and CFD.

5.3 GEANT4 optimizations

The navigator in GEANT4 handles the tracking of particles in a very general way but it limits the efficiency of the FD and CFD algorithms. The two main reasons for this are the mean free path calculations and the determination of the volume containing a particle. This is illustrated by figure 5.5, where a mother photon is tracked along its path through an analytically defined thorax phantom by aGATE. At decay, the volume in which the photon is located is determined

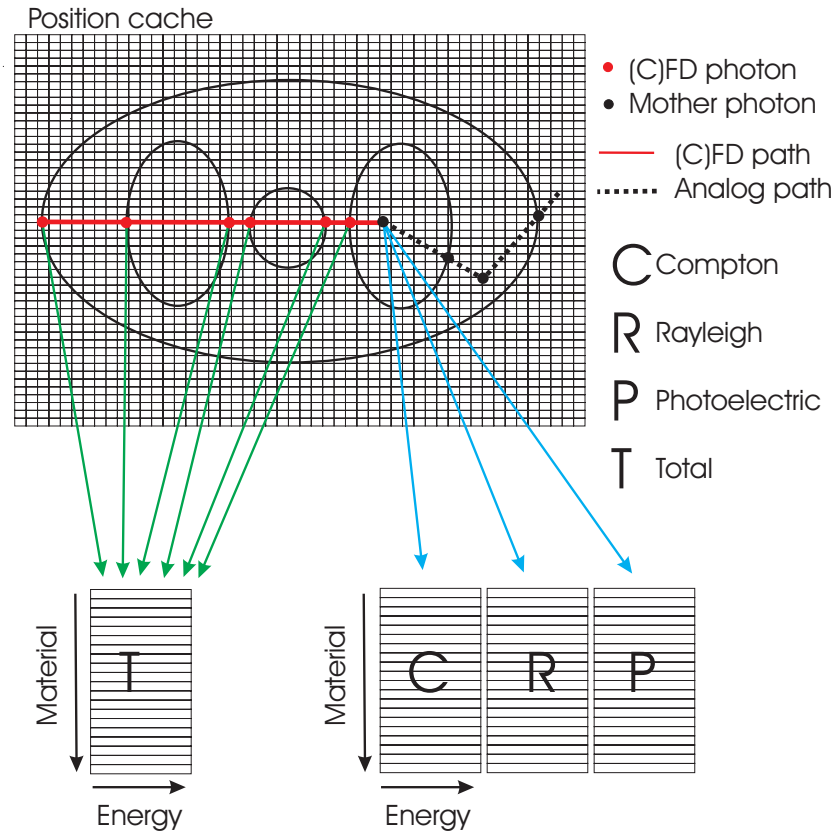


Figure 5.5: The FD and CFD optimized navigator. The position cache grid over a thorax phantom is shown with the fast mean free path tables. Sample photon paths are shown for a mother photon and an FD/CFD photon copy. The mother photon uses the tables for each process separately, while the photon copy polls the total cross section directly for each material.

and all the processes in the GEANT4 physics list are polled for their mean free path in the current material. In the case of SPECT simulations, the dominant processes are Compton and Rayleigh scatter and the photoelectric effect. Since materials in GEANT4 are built from the cross sections of the individual elements, this involves expensive log-log interpolations. The process with the smallest free path is selected, and the photon is transported to the new position. However, in the case of figure 5.5 the step is limited by geometry. At this point, the photon is located in a new volume. Therefore, the volume tree is searched for the correct volume and the processes are polled again for the mean free path in the new material. At the next position, the photon scatters and reaches the boundary of the phantom. At this point it is destroyed when using FD or CFD.

The simple tracking of the mother particle in this example therefore required at least 9 mean free path calculations and 3 expensive volume tree searches. In addition, a realistic SPECT simulation typically requires the use of a voxelized phantom, which considerably increases the use of both aforementioned expensive operations. When using FD or CFD with MPS, the tracking has to be repeated for every photon copy, depending on the number of projections that are simulated. In the case of figure 5.5 with 60 projections, this means that 60 photon copies must be tracked at decay and 60 at the scatter position, each requiring mean free path calculations and volume tree searches.

A SPECT simulation with FD or CFD shows a typical usage pattern of mean free path calculations: a limited energy range, a limited number of materials and repeated use. In addition, the weight corrections explained in section 5.2.2 require the total cross section. Therefore 4 mean free path tables are constructed at startup: one for each process and one summing the previous three. Each table is material and energy dependent, avoiding the expensive log-log interpolations and contains a cache for repeated use. This is illustrated in figure 5.5. The mother photon now uses the fast mean free path tables for each process separately. The path of 1 photon copy is shown as well. At the end of each step, the total cross section can be polled directly, for each material. Moreover, when using MPS, the cache becomes especially useful since the energy of the photon copies at decay and at Rayleigh interactions is identical.

In order to reduce the time spent on volume tree searches, an FD and CFD optimized navigator was developed. As shown in figure

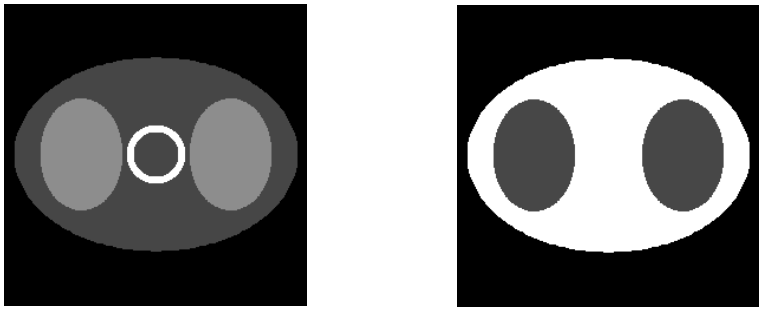
5.5, a position cache grid is placed over the phantom. At startup, each voxel in the grid is filled with a link to the volume it is positioned in. In addition, each border voxel also stores a link to the border volume. FD and CFD are based on non-absorption and non-interaction of photon copies along their path through the phantom. Thus the direction is constant for a given photon copy, and at every step a new volume is reached. The start volume of the photon copy is already known, as it is the same as the mother photon. At each of the border locations along the photon copy path in figure 5.5 where the photon crosses from one volume into another, the position cache grid is polled for the current volume. The position cache grid voxel size must be chosen smaller than the smallest distance between two borders of volumes. In the case of a voxelized phantom, the position grid voxel size and the phantom voxel size can be chosen equal.

5.4 Evaluation study

The validation of our FD and CFD approach consists of three steps. Firstly, the FD simulations are validated against aGATE simulations and compared to SimSET (version 2.6.2.6) simulations (3). Secondly, the CFD model itself is verified with a simplified version of analog GATE (sGATE). Thirdly, the CFD simulations are validated against aGATE simulations and compared to SimSET. The acronyms FD, CFD, SimSET, aGATE and sGATE will be used throughout the paper.

5.4.1 FD validation

Two simulation setups based on the Philips AXIS camera (11) with a low energy high resolution (LEHR) collimator were used in order to validate FD against aGATE and SimSET. The collimator hexagonal hole inner radius was 0.061 cm, with a septal thickness of 0.0203 cm. The AXIS camera simulations with aGATE have previously been described and validated by (12). First a single 30 second projection of a 222 MBq ^{99m}Tc point source in a cylindrical water phantom (10 cm height, 15 cm radius) was simulated. An energy window of 129 to 151 KeV was used to generate point spread functions (PSFs). Secondly, a symmetrical thorax phantom with an 109 MBq extended



(a) Activity (b) Attenuation

Figure 5.6: Activity distribution and attenuation map of a slice of the thorax phantom.

source distribution was used as shown in figure 5.6. The activity ratio was chosen as 5:10:100 for background (soft tissue), lungs and heart respectively to resemble ^{99m}Tc uptake values (13). The acquisition time was 30 seconds per projection over 60 projection angles. The projections were binned into 128x128 pixels, with 3.90625 mm per pixel and an energy window of 129 to 151 KeV. A transaxial profile of a lateral projection was drawn over 10 pixels width. Another profile was drawn over all projection angles with ten pixels width and ten slices depth. For point source simulations exactly 10^8 weighted histories were detected with FD and SimSET in order to obtain almost noiseless energy spectra and projections. For the symmetrical thorax phantom simulations an equal number of detections as in aGATE was used of approximately 10^6 weighted detections per projection. These point source and thorax simulations will be used as a means of validation throughout this paper.

Identical detectors were used for the aGATE and FD simulations. The FD simulations were performed with a cone opening angle of 5.2 degrees for the LEHR collimator.

The detector geometry used in SimSET was based on the UNC SPECT collimator module (14) and the planar SPECT detector module. The same septal thickness and inner radius was chosen as for the GATE simulations. However, the collimator module is based on a geometric transfer function for circular holes arranged in a hexagonal pattern. This configuration cannot give rise to collimator patterns and septal penetration and collimator scatter are not supported. The

planar detector consisted of the same layers as the GATE simulations, with the exception of a casing around the detector as described in (12). Variance reduction techniques were turned on when possible, including stratification and forced non-absorption. Forced detection could not be used yet in combination with the UNC collimator module.

5.4.2 CFD model verification

The CFD model itself was verified against sGATE with the point source simulation as described in section 5.4.1. The CFD simulation used a blurring kernel for a low energy high resolution (LEHR) collimator. The sGATE simulation used a simplified detector geometry consisting of only a perfectly absorbing LEHR collimator and crystal (developed with ProcessGate (15), chapter 3). Collimator scatter, septal penetration, collimator fluorescence and crystal fluorescence are accordingly not modeled. However, the hole size, septal thickness and hexagonal pattern remains intact.

5.4.3 CFD validation

Finally, CFD was validated against aGATE and SimSET using the aforementioned simulation setups (cfr. section 5.4.1).

5.5 Acceleration factors

An efficiency improvement can be realized by increasing the proportion of detected photons to the simulated photon histories, as is the case for FD and CFD. The gain in efficiency is however always less than expected by the increase of the number of detected photons due to the weights assigned to each history. The variability of the weights must therefore be taken into account for any figure of merit (FOM). We used three FOMs based on sets of 100 simulations of a point source in a water phantom for aGATE, FD and CFD respectively to measure the gain in efficiency. The FOMs are applied to the flux D/N of the detector, defined as the number of detected photons D, divided by the number of simulated photon histories. In the case

of aGATE the photons are scored as 1 or 0, while in the case of FD and CFD the weight is scored.

$$RFOM = \frac{R^2_{aGATE} T_{aGATE}}{R^2_{(C)FD} T_{(C)FD}} \quad with \quad R = \frac{st.deviation}{mean} \tag{5.4}$$

$$\begin{gathered} \hat{x} = \frac{1}{N}\sum_{i=1}^{N} x_i \\ S_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \hat{x})^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i^2 - \hat{x}^2) \\ S_{\hat{x}}^2 = \frac{S_x^2}{N} \quad and \quad R = \frac{S_{\hat{x}}}{\hat{x}} \end{gathered} \tag{5.5}$$

$$\hat{x} = \frac{D}{N} \quad and \quad S_{\hat{x}}^2 = \frac{\hat{x}(1-\hat{x})}{N-1} \tag{5.6}$$

$$QF = \frac{\left(\sum_{i=1}^{N} x_i\right)^2}{D\sum_{i=1}^{N} x_i^2} \tag{5.7}$$

$$RFOM_{QF} = \frac{detections/s}{QF \quad detections/s} \tag{5.8}$$

The first FOM is derived from the sets of 100 simulations. The mean and standard deviation of the flux are calculated. The relative error R is defined as in equation 5.4. A relative FOM (in this case $RFOM_{batch}$) taking into account the time T, also shown in equation 5.4, gives an indication of the relative efficiency of the simulations. The $RFOM_{batch}$ was calculated for both FD and CFD versus aGATE.

A second FOM was used as in (1). The sample mean, standard deviation and relative error are based on the number of simulated histories of a single simulation as introduced in section 2.4.3. In the case of FD and CFD, the flux, $\bar{x}$, is estimated by the total score of the weights divided by the number of simulated histories N as shown in equation 5.6. The variance S_x^2 associated with the distribution of the weights, x_i, is calculated as in equation 5.6. The variance of the mean and the relative error are then defined as in equation 5.6. In the case of aGATE the scoring is binary: either a photon contributes to a tally or not. In this case, the equations in 5.6 can be simplified as shown in equation 5.6. The same RFOM as in equation 5.4 was

then used to calculate the efficiency of each of the 100 simulations, expressed as $RFOM_{single}$.

Thirdly, the quality factor (QF) introduced in (2) is calculated as shown in equation 5.7. It indicates how much the number of detections per second is deflated due to the variability in the weights. It therefore gives an idea of the value of a detection with FD or CFD compared to an analog Monte Carlo simulation. The QF times the number of detections per second for FD, CFD and aGATE ($QF = 1$) can then be used to calculate $RFOM_{QF}$ as shown in equation 5.8.

5.6 Results

5.6.1 FD validation

Figure 5.7 shows the FD versus aGATE and SimSET energy spectra comparison for the point source in a water phantom. No energy blurring was used and the results were normalized to the maximum. Between 0 and 50 KeV the contributions are mainly due to detector backscatter. The FD and aGATE spectra match very well, while the backscatter is slightly underestimated by SimSET. The FD underestimation in the 50 keV to 90 keV range can be mainly attributed to collimator fluorescence, due to the limited cone opening angle. The SimSET underestimation in this range is due to the lack of a collimator fluorescence model in the energy spectrum. The FD model closely matches the analog simulations from 90 to 140 KeV. This includes the crystal fluorescence peak and the effects of the GEANT4 low energy model which are especially visible in the 130 to 140 KeV range. Neither of these effects are modeled by SimSET in the energy spectrum. Close to the primary peak however, the discrepancies are due to both penetration and collimator Rayleigh scatter as these effects are underestimated by FD and not modeled by SimSET. Figure 5.8 shows the percent contributions of collimator penetration, collimator fluorescence and collimator scatter to the total number of detections and to the number of detections within the photopeak window for the analog simulation. The energy spectrum of these contributions is shown in figure 5.9.

The PSFs of the FD versus aGATE and SimSET comparison for the point source in a water phantom are shown in figure 5.10. The full width at half maximum (FWHM) and the full width at tenth maxi-

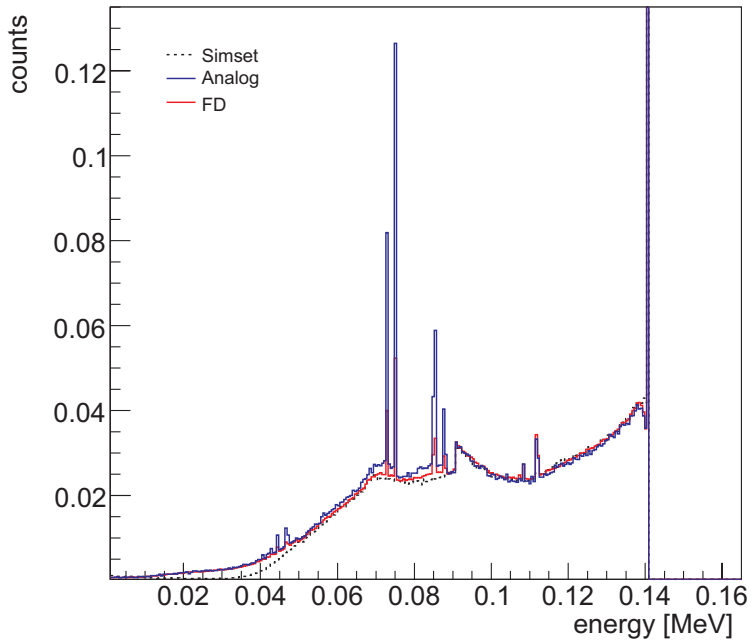


Figure 5.7: FD vs aGATE and SimSET energy spectra comparison, scale normalized to the maximum. (Zoomed)

mum (FWTM) for all point source simulations are shown in table 5.1. The FWHM and the FWTM of FD and aGATE are in very good agreement. The SimSET FWTM deviates relatively far from the aGATE simulation value.

5.6.2 CFD model verification

The results of the CFD model verification with sGATE as described in section 5.4.2 are shown in figures 5.11 to 5.12. Figure 5.11 shows the energy spectra comparison of the point source simulation for CFD versus sGATE. No energy blurring was applied and the results were normalized to the maximum. The total energy spectrum and 5 scatter orders are compared. The CFD model agrees very well with the sGATE simulation both for the total energy spectrum and the in-

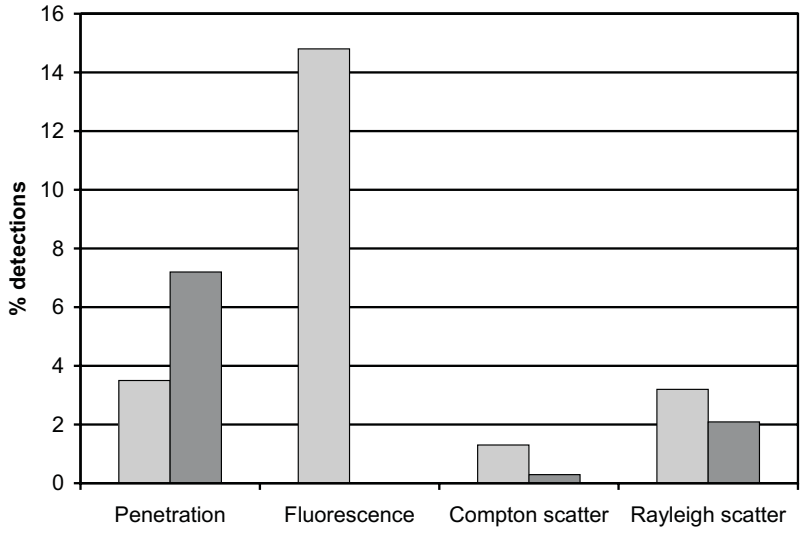


Figure 5.8: Influence of septal penetration, fluorescence, Compton scatter (collimator) and Rayleigh scatter (collimator) as computed by aGATE. Percent contributions to the total number of detections (light grey) and to the total number of detections within the photopeak window (dark grey).

dividual scatter orders. Only very small discrepancies can be noticed in the energy spectrum. These are due to the collimator hole inner radius and septal thickness and the finite number of collimator holes in the analog simulation, since the CFD model assumes a collimator hole at every detector position.

The corresponding PSFs are shown in figure 5.12. The FWHM and the FWTM for all these simulations is also shown in table 5.1. There is a near perfect agreement between the CFD and sGATE PSFs.

5.6.3 CFD validation

Figures 5.13 and 5.14 show the CFD, aGATE and SimSET energy spectrum comparison for the point source simulation of section 5.4.3.

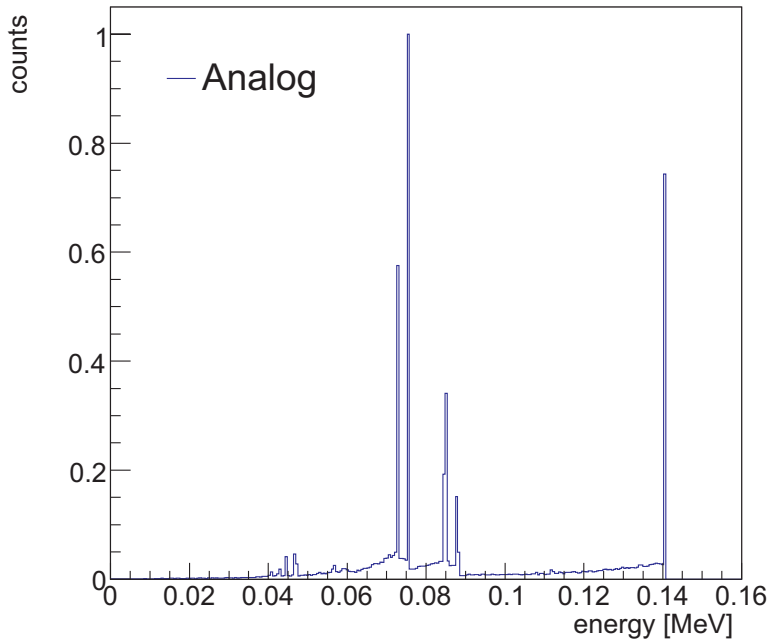


Figure 5.9: Influence of septal penetration, fluorescence, Compton scatter (collimator) and Rayleigh scatter (collimator) as computed by aGATE. Energy spectrum normalized to the maximum.

The limits of the CFD model regarding the energy spectrum become apparent in this case. The large overestimation of primary photons by 41% implies an underestimation of scattered photons. There are two main reasons for this behaviour. First, the detector model assumes a hole at each collimator position and applies only a simple Gaussian blurring kernel to reflect a geometrical filter, which has no effect on the energy spectrum. Therefore, both septal penetration and detector interactions are not modeled, including fluorescence, Compton and Rayleigh scatter in the collimator of which the contributions are shown in figures 5.8 and 5.9. Secondly, only attenuation paths through the phantom that are perpendicular to the detector are sampled by forced photons. Therefore the SimSET energy spectrum with the planar detector model agrees much better with the aGATE

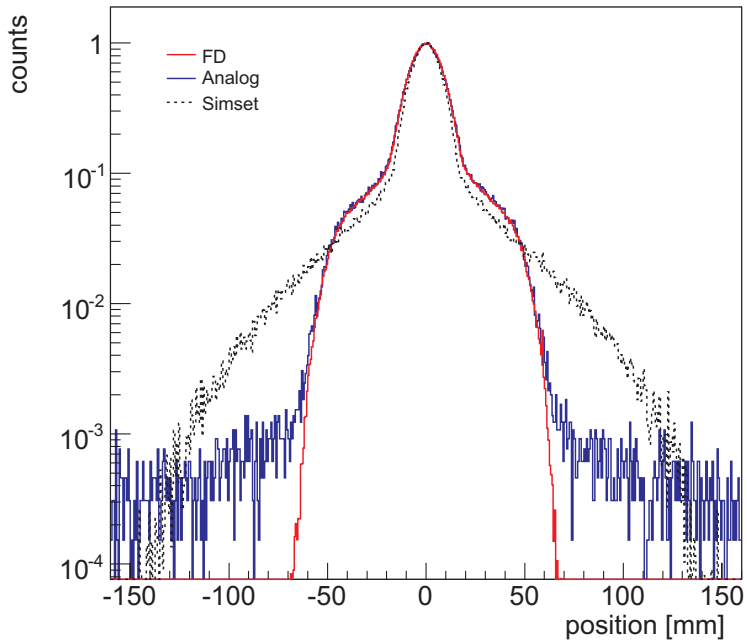


Figure 5.10: FD vs aGATE and SimSET PSFs, log scale normalized to the maximum.

simulation.

However, the agreement between the CFD and aGATE PSFs is much better than between the SimSET and aGATE PSFs, as shown in figure 5.15. The FWHM and FWTM values are also shown in table 5.1.

5.6.4 Symmetrical thorax phantom

Figure 5.16 shows transaxial profiles of the thorax phantom simulations for the aGATE versus FD, CFD and SimSET simulations respectively. There is a a close match for FD, CFD and SimSET versus the analog simulation, with slightly larger deviations for SimSET. Figure 5.17 depicts the sinograms (central axial slice) of the analog simulation and the SimSET, FD and CFD simulations. A pro-

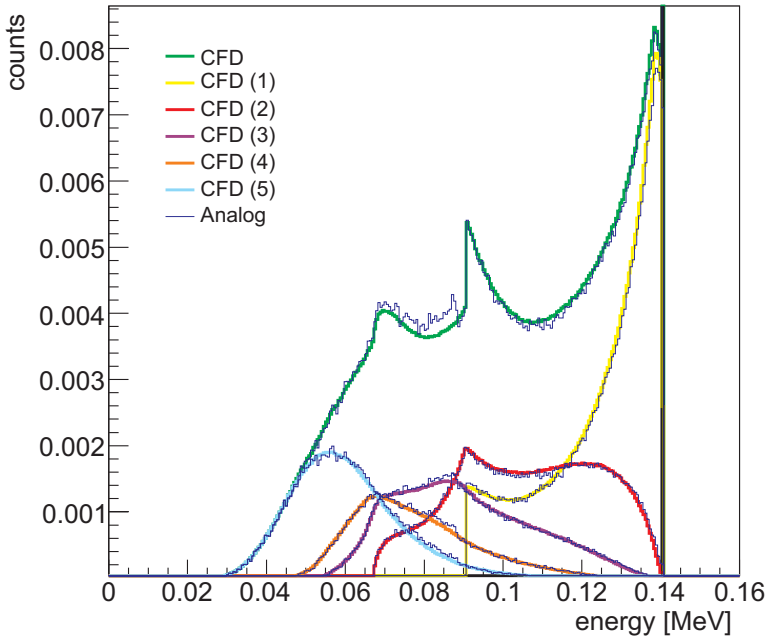


Figure 5.11: CFD model verification: CFD vs sGATE energy spectra comparison. The total energy spectrum and scatter orders (1) to (5) are compared.

file comparison between aGATE, FD, CFD and SimSET is shown in figure 5.19 together with the blurred energy spectrum in figure 5.18. The energy spectrum shows a slight overestimation of the primary peak for FD and increasingly larger overestimations for SimSET and CFD respectively. Spatially, the discrepancies with aGATE are smallest for FD, slightly larger for CFD and largest for SimSET.

5.6.5 Acceleration

Table 5.2 shows the FD and CFD RFOMs and QFs for the FD and CFD point source simulations. The RFOM based on a single simulation, $RFOM_{single}$, is the average of the RFOMs calculated for each

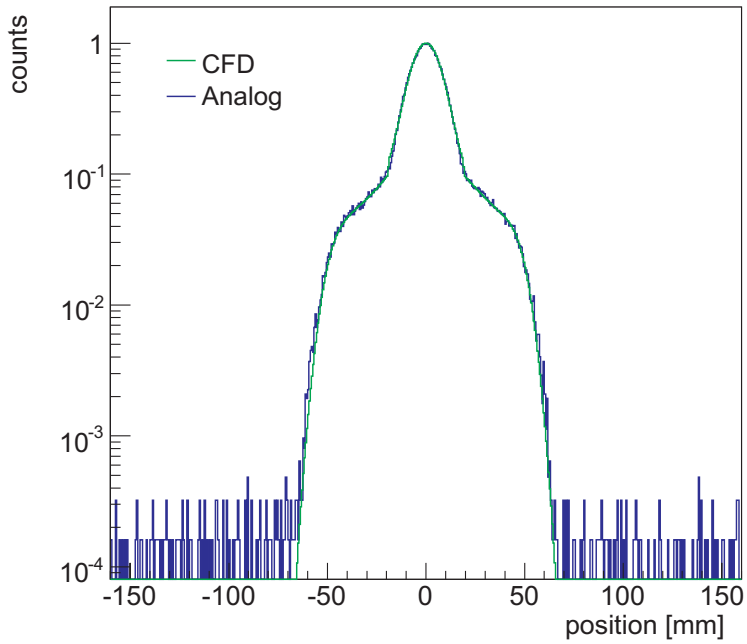


Figure 5.12: CFD model verification: CFD vs sGATE PSFs, log scale normalized to the maximum.

of the 100 simulations. An example of the spread of the RFOMs in the case of FD is shown in figure 5.20. The QFs and the RFOMs based on the QFs in table 5.2 are also an average over the set of 100 simulations.

All acceleration factors are a combination of the GEANT4 optimization and FD/CFD. The $RFOM_{single}$ derived from a single simulation largely agrees with the other RFOMs and can therefore be used as an approximate indication of the relative efficiency. FD is approximately three order of magnitude times faster than aGATE. CFD is over 900 times faster than FD and approximately six orders of magnitude faster than aGATE. The QFs indicate that a detection generated by FD or CFD is approximately worth 0.5 detections with aGATE when Rayleigh scatter is not included in the simulation. The inclusion of the Rayleigh scatter process results in a significant drop

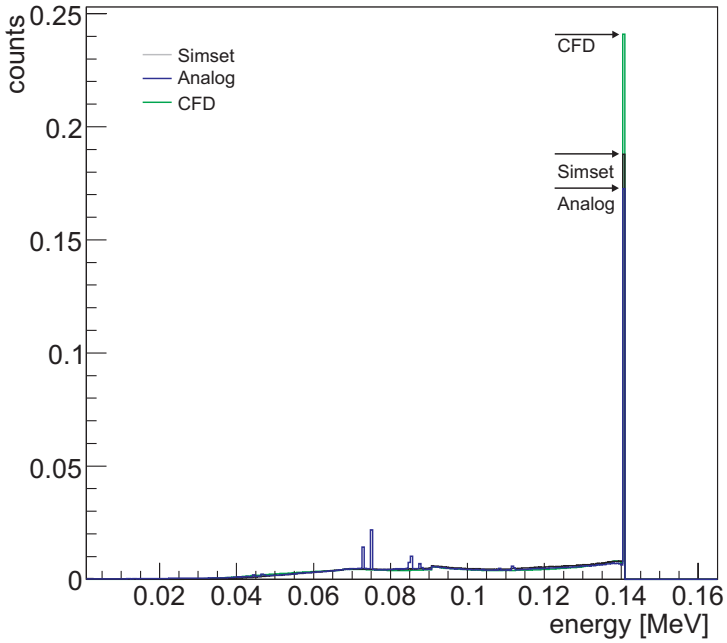


Figure 5.13: CFD vs aGATE and SimSET energy spectra comparison.

	FD	CFD	aGATE	sGATE	SimSET
FWHM	20.0	18.7	20.0	18.7	18.1
FWTM	41.3	38.7	41.3	39.4	33.7

Table 5.1: FWHM and FWTM for FD, CFD, aGATE, sGATE, SimSET

in the QF for both FD and CFD. Longer simulation times are therefore required in order to include Rayleigh scatter. Another more efficient method may be the inclusion of weight windows (16) in order to control the weight variations.

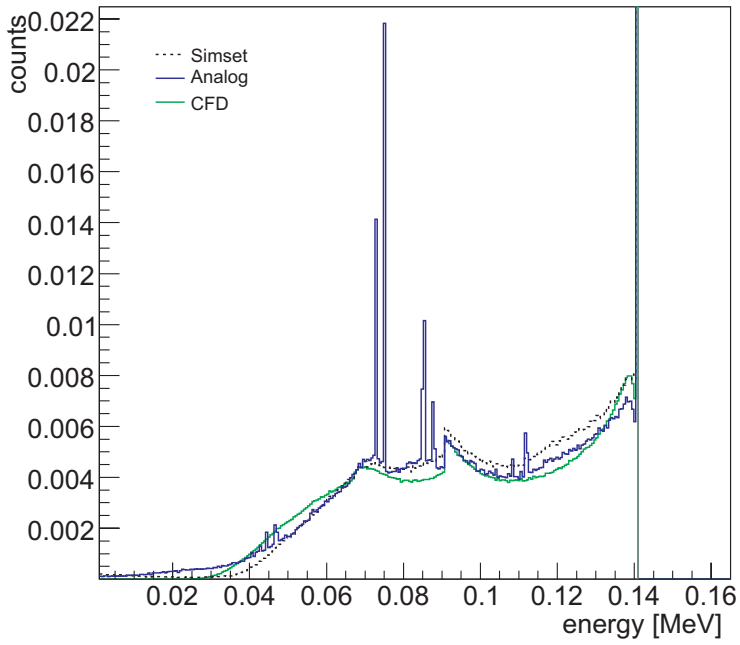


Figure 5.14: CFD vs aGATE and SimSET energy spectra comparison (Zoomed).

	FD	CFD
$RFOM_{batch}$	1690	1652487
$RFOM_{single}$	1747	1682219
$RFOM_{QF}$	1828	1788500
QF	0.48	0.53
$QF_{Rayleigh}$	0.25	0.18

Table 5.2: Relative FOMs and QFs: FD and CFD versus aGATE

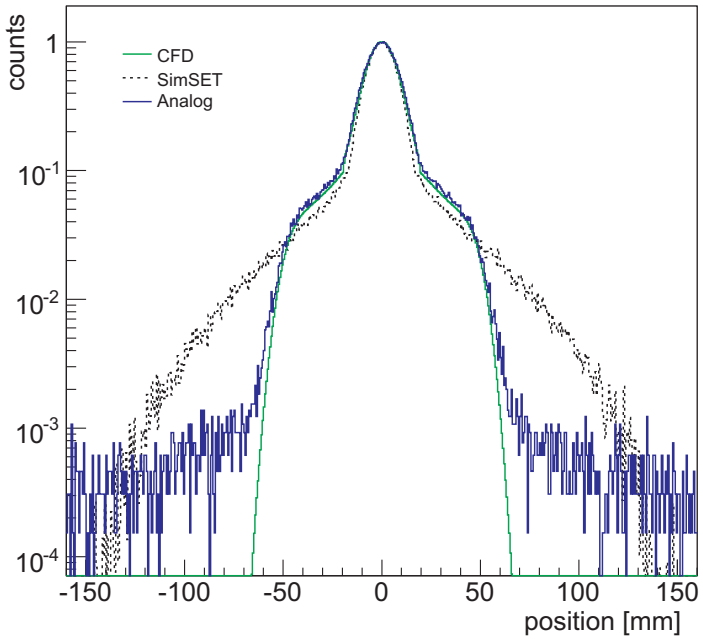


Figure 5.15: CFD vs aGATE and SimSET PSFs, log scale normalized to the maximum.

5.7 Application to PET

Previously, forced detection has been applied to PET simulations with SimSET (3). The principle remains the same as in SPECT: photon copies are forced into a direction that depends on the solid angle of acceptance of the detector. Relatively simple detector geometry was used. The detector was modeled as a series of adjacent regular right cylinders with transaxial layers. Photon interactions including scatter, absorption, and septa penetration were simulated, but there were no blocks and hence no gap effects. In addition, time was not modeled throughout the simulation.

A direct translation of the forced detection methods in this chapter to PET would therefore consist of two parts. Firstly, an extension of the physics code is required. Positron annihilation must be

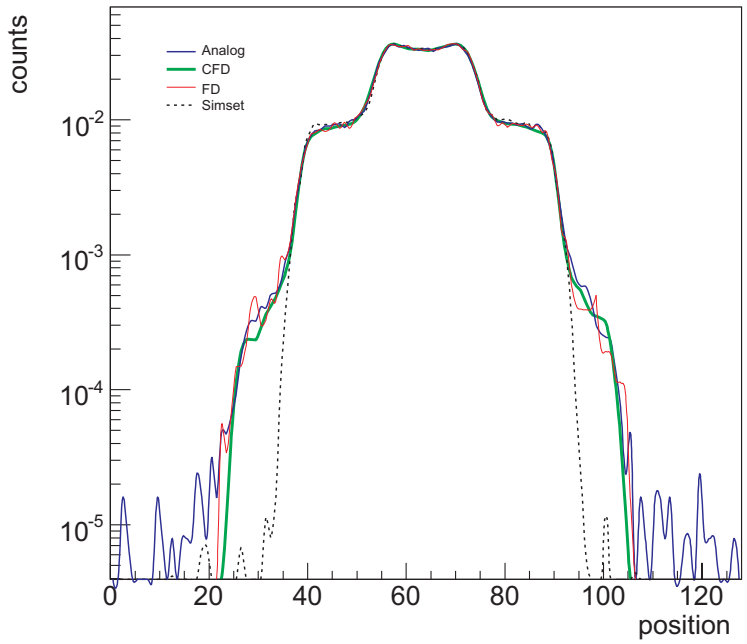
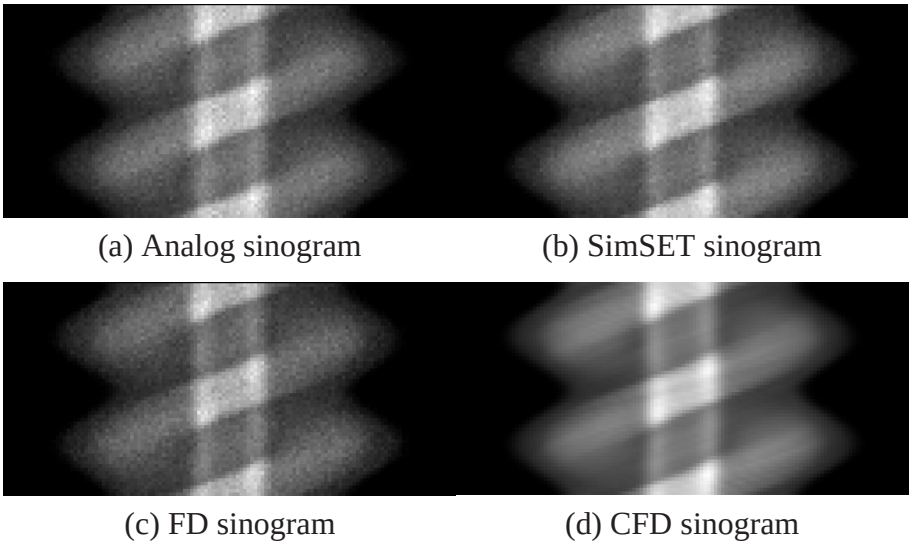


Figure 5.16: Transaxial profiles of the symmetrical thorax phantom projections for aGATE, FD, CFD and SimSET (log scale).

modified in order to simulate the forced detection of unscattered annihilation photons. This is only a simple modification, since the emission angle is sampled uniformly, after which acolinearity is applied. In addition, a recalculation of the probability tables is required to accommodate the acceptance angle of the PET detector in question. Secondly, the same problems apply as discussed in section 4.6. Therefore a complete revision of the digitizer chain is required to support statistical weights. Forced detection was primarily developed to solve the sensitivity problem in SPECT simulations. Due to the higher sensitivity of PET, a smaller gain in efficiency is to be expected.



(a) Analog sinogram (b) SimSET sinogram

(c) FD sinogram (d) CFD sinogram

Figure 5.17: Sinograms of the symmetrical thorax phantom simulations.

5.8 Discussion

A realistic SPECT simulation of the MCAT phantom takes approximately 7 years and 111 days with analog GATE, 4 days and 15 hours with SimSET, 35 hours with FD and 130 seconds with CFD on a single pentium 4 2.8Ghz CPU.

Analog GATE is the most detailed and versatile simulator and can be used for both PET and SPECT. However, the simulation time is too long to be clinically applicable.

SimSET is much faster than aGATE. Although it deviates from aGATE when using the UNC collimator, it does supports both PET and SPECT for various other isotopes than ^{99m}Tc. The limited SimSET acceleration result is mainly due to the absence of forced detection in combination with the UNC SPECT collimator module. The layered detector model and the inclusion of Rayleigh scatter in the phantom also introduced a significant slowdown.

Our FD solution is faster than SimSET but currently only supports ^{99m}Tc SPECT. The FD and CFD models share the same code, therefore it can be concluded that mainly the detector interactions slow down FD in GATE. FD is however both spectrally and spa-

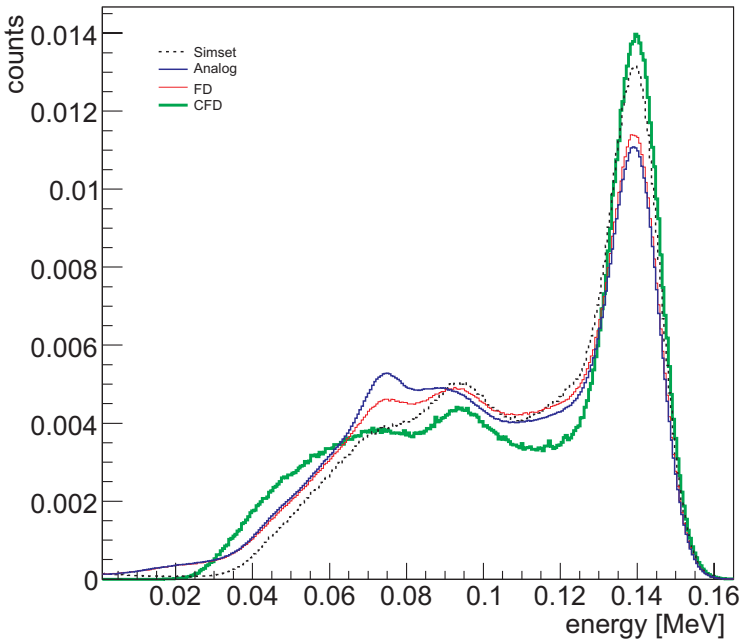


Figure 5.18: Comparison of FD and CFD sinograms vs aGATE and SimSET: Energy spectrum

tially in excellent agreement with aGATE. Although the limited cone opening angle results in a slight underestimation of scatter, this has a negligible spatial effect for ^{99m}Tc. The FD and CFD solutions presented in this paper can also readily be combined with the GATE cluster platform to shorten the simulation time even further. This enables fast, realistic high count ^{99m}Tc SPECT simulations with GATE which can be used for detector design and prototyping. An additional septal penetration module will be developed in the future in order to simulate isotopes with higher energy peaks.

CFD is clearly the fastest solution for ^{99m}Tc SPECT and fast enough to be used as a forward projector in a reconstruction algorithm. Due to the lack of collimator scatter and penetration, there is a FWTM discrepancy for CFD. This shows that a penetration and collimator Rayleigh scatter model would contribute most in order to

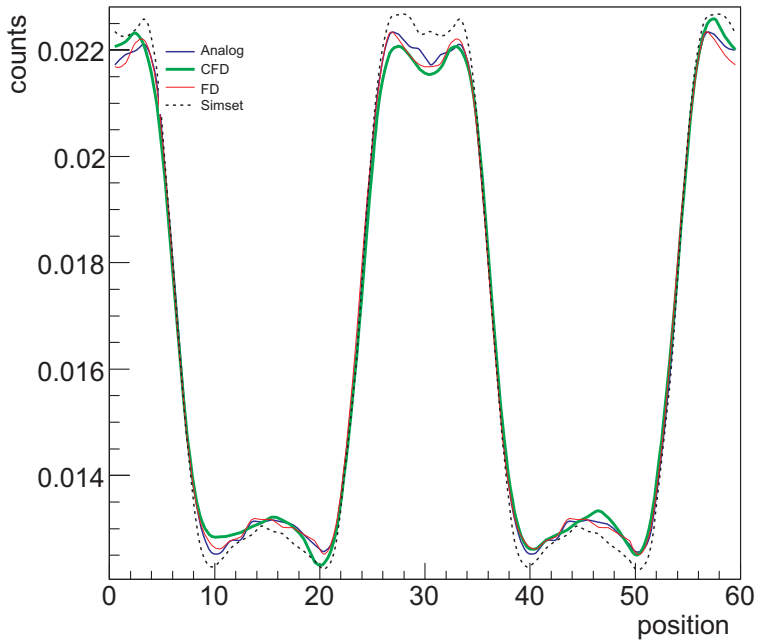


Figure 5.19: Comparison of FD and CFD sinograms vs aGATE and SimSET: Profile

obtain increased accuracy for CFD. Future work will therefore include the extension of CFD in GATE with our previously developed septal penetration module (17).

5.9 Summary and original contributions

In this chapter we investigated the acceleration of ^{99m}Tc SPECT simulations with forced detection (FD) and convolution-based forced detection (CFD) for GATE. SPECT simulations are computationally very intensive for two main reasons. Firstly, without variance reduction most photons are tracked in vain, since only 0.02% of all emitted photons actually reach the crystal surface when a low energy high resolution (LEHR) collimator is used. Secondly, the particle

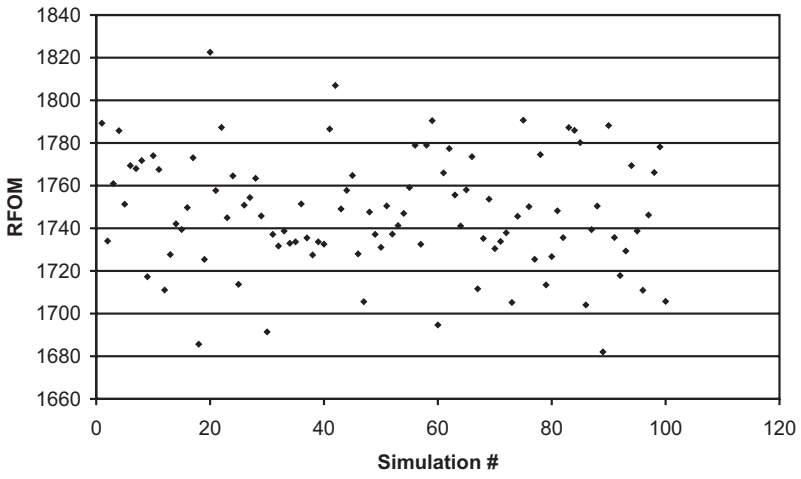


Figure 5.20: The RFOMs of FD versus aGATE for 100 independent simulations of FD and aGATE.

navigation algorithms in GEANT4 are not optimized for efficiency.

FD is a variance reduction technique that increases the proportion of photons that result in a detection. At decay and at each interaction point a photon copy of the original photon is transported through the phantom in a direction sampled within a solid angle toward the SPECT detector head. A weight is assigned to each photon copy to compensate for the forced direction and the non-absorption between the starting point or scatter point and escape from the phantom. Contrary to other Monte Carlo codes, no approximations were used to sample the attenuation along the photon path. The weights are therefore compliant with the low energy model used in GEANT4, including atomic form factors for Compton and Rayleigh scatter. With standard forced detection all detector interactions are simulated by analog Monte Carlo. In the case of convolution-based forced detection the detector is modeled analytically as proposed by Beekman et al in (5). In addition, both FD and CFD were extended with multiple projection sampling which enables the generation of all projections simultaneously.

Subsequently we discussed optimizations for the GEANT4 particle navigation algorithms. The calculation of the photon mean free

path was accelerated with on-the-fly generated tables, specifically for the materials and energy range required for the simulation. In addition, an FD and CFD optimized navigator was introduced to handle the tracking of photon copies using a-priori information of the photon paths and the use of a position cache grid placed over the attenuation geometry.

The validation of our FD and CFD approach consisted of three steps. Firstly, the FD simulations were validated against analog GATE. A point source in a water phantom and a symmetrical thorax phantom with an extended source distribution were simulated. FD was found to be both spectrally and spatially in excellent agreement with analog GATE. Although the limited cone opening angle results in a slight underestimation of scatter, this has a negligible spatial effect for ^{99m}Tc. Secondly, the CFD model was validated against a simplified Gate simulation using a perfectly absorbing collimator and crystal. Only very small discrepancies could be noticed in the energy spectrum due to the collimator hole pattern modeled with the simplified GATE simulation. With regard to the point spread functions, a near perfect agreement was obtained. Thirdly, CFD was validated against analog GATE using the same simulation setups as for FD. Spatially, a very good agreement was found, although a FWTM discrepancy resulted from the lack of collimator scatter and penetration modeling. Spectrally, a large deviation from analog GATE was observed. This was expected, as the detector model in CFD is defined only spatially.

We can conclude that both FD and CFD were implemented and validated in GATE for ^{99m}Tc SPECT. Although CFD is well over 900 times faster than FD, it is limited by the analytical description of the detector. As detector interactions are not modeled, the resulting energy spectra may deviate relatively far from an analog GATE simulation. This is in contrast to the FD simulations, where the photopeak window was almost perfectly modeled. Spatially, both FD and CFD closely match the analog GATE simulations. This work accelerates GATE from three to six orders of magnitude and it can be combined with the GATE cluster framework (see (18) and chapter 6).

The work presented in this chapter resulted in the publication of a conference proceeding (19) and an A1 journal paper (20).

References

[1] F.B. Brown et al. *MCNP Version 5*. Trans. Am. Nucl. Soc., 87:273–381, 2002.

[2] D. Haynor, R. Harrison, and T. Lewellen. *The use of importance sampling techniques to improve the efficiency of photon tracking in emission tomography simulations*. Med. Phys., 18:990–1001, 1991.

[3] R.L. Harrison, S. Dhavala, P.N. Kumar, R. Yiping Shao Manjersshwar, T.K. Lewellen, and F.P. Jansen. *Acceleration of SimSet photon history generation.* In Nuclear Science Symposium Conference Record, volume 3, pages 1835–1838, 2002.

[4] M. Ljungberg. *The SIMIND Monte Carlo program*. In M. Ljungberg, S.E. Strand, and M. A. King, editors, Monte Carlo calculations in nuclear medicine: Applications in diagnostic imaging, pages 145–163, Bristol, 1998. Institute of Physics Publishing.

[5] F. J. Beekman, H. W. de Jong, and E. Slijpen. *Efficient SPECT scatter calculation in non-uniform media using correlated Monte Carlo simulation*. Phys. Med. Biol., 44:183–192, 1999.

[6] F.J. Beekman, H.W.A.M. de Jong, and S. van Geloven. *Efficient fully 3-D iterative SPECT reconstruction with Monte Carlo-based scatter compensation*. IEEE Trans. Med. Imaging, 21(8):867–877, 2002.

[7] S. Lui et al. *Accelerated SPECT Monte Carlo Simulation using Multiple Projection Sampling and Convolution-based Forced Detection.* In Proceedings of the 2006 IEEE Medical Imaging Conference, pages 3142–3147, 2006.

[8] H. W. A. M. de Jong, E. T. P. Slijpen, and F. J. Beekman. *Acceleration of Monte Carlo SPECT simulation using convolution-based forced detection*. IEEE Trans. Nucl. Sci, 48(1):58–64, 2001.

[9] F.J. Beekman and M.A. Viergever. *Fast SPECT Simulation Including Object Shape Dependent Scatter*. IEEE Trans. Med. Imaging, 14:271–282, 1995.

[10] J. H. Hubbel. *Summary of Existing Information on the Incoherent Scattering of Photons particularly on the Validity of the Use of the Incoherent Scattering Function*. Radiat. Phys. Chem., 50(1):113–124, 1997.

[11] Philips Medical Systems, 595 Miner Road, Cleveland, OH 44143, USA.

[12] S. Staelens, D. Strul, G. Santin, S. Vandenberghe, M. Koole, Y. D'Asseler, I. Lemahieu, and R. Van de Walle. *Monte Carlo simulations of a scintillation camera using GATE: validation and application modelling*. Phys. Med. Biol., 48(18):3021–3042, 2003.

[13] J. Xiao, T . C. de Wit, S. Staelens, and F. J. Beekman. *Evaluation of 3D Monte Carlo-Based Scatter Correction for Tc99m Cardiac Perfusion SPECT*. J. Nucl. Med., 47(10):1662–1669, 2006.

[14] B.M. Tsui and G.T. Gullberg. *The geometric transfer function for cone and fan beam collimators*. Phys. Med. Biol., 35:81–93, 1990.

[15] J. De Beenhouwer, S. Staelens, S. Vandenberghe, J. Verhaeghe, and I. Lemahieu. *Process level discrimination for GATE: assessment of contamination in SPECT and spurious activity in PET*. Med. Phys., submitted, 2008.

[16] D. Haynor, R. Harrison, T. Lewellen, A. Bice, C. Anson, S. Gillespie, Miyaoka R., K. Pollard, and J. Zhu. *Improving the efficiency of emission tomography simulations using variance reduction techniques*. IEEE Trans. Nucl. Sci, 37(2):749–753, 1990.

[17] S. Staelens, T. de Wit, and F.J. Beekman. *Fast hybrid SPECT simulation including efficient septal penetration modelling (SP-PSF)*. Phys. Med. Biol., 52:3027–3043, 2007.

[18] J. De Beenhouwer, S. Staelens, D. Kruecker, L. Ferrer, Y. D'Asseler, I. Lemahieu, and F.R. Rannou. *Cluster computing software for GATE simulations*. Med. Phys., 34(6):1926–1933, 2007.

[19] J. De Beenhouwer, S. Staelens, S. Vandenberghe, and I. Lemahieu. *Acceleration of GATE SPECT Simulations*. In Proceedings of the IEEE Nuclear Science Symposium and Medical Imaging Conference, pages 3649–3655, 2007.

[20] J. De Beenhouwer, S. Staelens, S. Vandenberghe, and I. Lemahieu. *Acceleration of GATE SPECT Simulations*. Med. Phys., 35(4):1476–1485, 2008.

6

Cluster Computing Software for GATE Simulations

6.1 Introduction

In chapters 4 and 5 we explored the possibility to accelerate GATE with variance reduction techniques such as geometrical importance sampling, forced detection and convolution-based forced detection. In this chapter we investigate the acceleration of GATE through the use of parallelization, a method that does not introduce any approximations and that can readily be combined with the acceleration techniques developed in the previous chapters. Monte Carlo simulations are excellently suited for parallelization, showing a theoretical linear speed-up as function of the number of processing nodes. However, this is usually not achieved due to many factors such as the setup time for each individual job, the output handling and the cluster specific system overhead such as remote disk access. Parallelization has been successfully implemented in other Monte Carlo packages. Previous work (1) has reported on the parallelization of the Monte Carlo N-Particle code (MCNP) (2) and a speed increase factor that approached the number of processors using the Message Passing In-

terface (MPI) was reported (3). The feasibility and efficiency of performing MCNP5 calculations with a small, heterogeneous computing cluster built from personal computers has been explored in (4). A nearly linear speed increase for up to 10 processors was obtained. In (5) the different steps of porting a Monte Carlo package to the PowerPC are described. A linear decrease in computing time was achieved for 12 computing nodes. A parallel environment for SimSET (6) has been described in (7), based on a client-server model to partition the events that takes load balancing into account. SIMIND (8) has successfully been ported to an IBMSP2 (9) parallel computer with the events equally distributed among the processors using MPI.

In order to reduce the overall computing time of GATE experiments, a parallel computing platform for running such simulations in a cluster of computers is developed which significantly shortens the setup time and provides fast data output handling. In (10) the submission of GATE simulations on a grid has been described. The presented solution suffered however from the trade-off between job splitting and grid pay-off depending on the queuing policy of the grid software. The approach described here is platform independent in the sense that the simulations are virtually separated so that the user obtains a number of fully resolved independent job execution macros accompanied by a platform specific submit file. Moreover, the software is fully automated and requires no interaction whatsoever from the user. This chapter focuses on investigating the scalability of GATE jobs on small to medium sized local clusters running on various platforms such as openMosix (11), Condor (12), openPBS (13) and Xgrid (14).

The parallelized simulations are made up of 3 steps : the job distribution, the actual simulations (on a number of CPUs) and the output handling. We will first describe the job distribution, second the data output handling and third the mathematical model used to measure the speed increase. A twofold evaluation study is then presented to validate our approach. First we investigate the performance and scalability with a number of benchmark simulations. Secondly, we describe a case study used to evaluate the potential gain in image quality.

6.2 Job distribution

6.2.1 Job splitter

There are a number of possible parallelization techniques, such as activity based splitting and event based splitting. However, the approach used should not alter physical properties such as singles rates, random coincidence events and system deadtime. The most general scheme for splitting PET and SPECT simulations that does not involve any approximation nor simplification, is the time-domain decomposition. Time-domain decomposition based on equal time intervals is inefficient however for isotopes with a relatively short (compared to the acquisition time) half life. The first time interval of a simulation with an aquisition time equal to the half life of the isotope already requires twice as much computation time as the last time interval. Therefore, an experiment starting at t_s with initial source activity A_0 and ending at t_S is partitioned into N simulations, such that each simulation processes an equal amount of decays and thus requires an equal amount of computation time. This principle is shown in equation 6.1 where λ is the decay constant supplied to the job splitter. The amount of decays of the simulation ending at time t_n must equal n/N times the total amount of decays.

$$A_0 \int_{t_s}^{t_n} e^{-\lambda t} dt = \frac{n}{N} A_0 \int_{t_s}^{t_S} e^{-\lambda t} dt \tag{6.1}$$

The time intervals can be therefore be calculated by using formula 6.2.

$$t_n = \frac{\ln\left(\frac{N-n}{N} e^{-\lambda t_s} + \frac{n}{N} e^{-\lambda t_S}\right)}{-\lambda} \tag{6.2}$$

Using a set of predefined random seeds it is possible to initialize the GEANT4 random generator to produce independent streams for all the simulations in the experiment. The generation of the seeds is accomplished by the HepJamesRandom engine of CLHEP (15; 16) which implements the algorithm RANMAR by Marsaglia-Zaman described in (17) and section 2.4.1.4. A partition of n streams is built by initializing the same random generator by n different seeds (18). This enables 900.000.000 independent streams with a period of approximately 10^{43}.

In order to apply time-domain splitting to GATE, it must be integrated with its virtual clock synchronization. A typical GATE simulation is divided in time slices defined by a time length (the equivalent of a GEANT4 run). Although the activity is updated continuously throughout the simulation, the geometry is only updated at the beginning of each time slice. For example, a single detector head

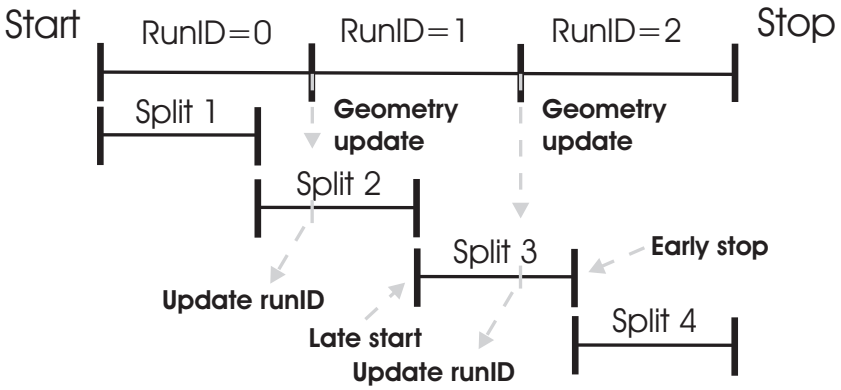


Figure 6.1: A GATE simulation consisting of 3 runs. The parallelization introduces virtual time slices that update the geometry based on the original time schedule.

SPECT simulation consisting of a 1800 second acquisition with a time slice length of 30 seconds would result in 60 time slices and thus 60 projections as the gantry does not move between projections. If the time-domain based decomposition would be restricted to the number of time slices, this would limit the maximum number of CPUs that could be used. To overcome this limitation, virtual time slices are introduced which can span a shorter, equal or longer time length than that of a single time slice in the original simulation, while the original geometry updates remain in place. Figure 6.1 illustrates this novel approach for a simulation consisting of 3 time slices that is split into four virtual time slices. In this case the virtual time slices do not coincide with the original time slices and the geometry updates need to take place at the end of the original time slices with rundID 0 and 1. Virtual slice 3 for instance starts with the geometry of the time slice with runID 1 but with less source activity as it is synchronized to the correct point in time. It is therefore actually a late start of this original time slice. The geometry update at the end

of the time slice with runID 1 occurs during this virtual time slice. At this time the virtual slice is updated with a new runID to accommodate the geometry updates of the original time schedule. Figure 6.1 shows how the third virtual slice ends before the time slice with rundID 2 is finished. It is therefore also an early stop of the time slice with runID 2.

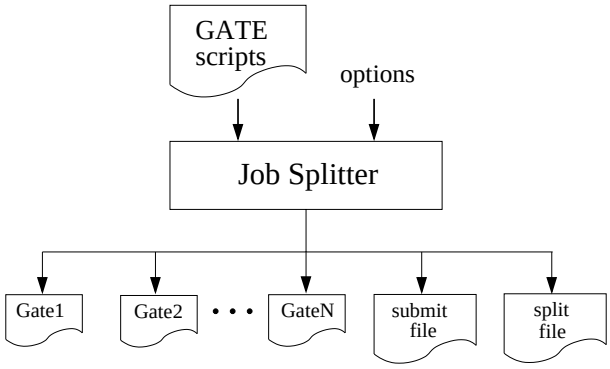


Figure 6.2: Schematic overview of the job splitter

In order to subdivide the simulation into virtual slices as described above, a job splitter program was developed. As shown in figure 6.2, it uses the simulation macro to produce a collection of fully resolved macro files, each describing the simulation during a certain virtual time slice. The non-parameterized macro files can be run with a simple GATE command and therefore provide the highest level of compatibility with any kind of cluster platform. A cluster platform specific submit file is also generated. It contains all the commands neccessary to launch the simulation on the cluster. The supported cluster platforms currently include Condor, openMosix, OpenPBS and Xgrid and this list can easily be extended. Finally, a split file is generated that contains all information about the partitioned simulation such as data storage directories and time information to facilitate the merging of the output files and error handling.

6.2.2 Job setup time

In (19) accurate GATE models of low energy high resolution (LEHR) and medium energy general purpose (MEGP) collimators were de-

scribed. Since the job splitter returns fully resolved macros, the acquisition geometry needs to be built for each individual simulation. However, a typical high resolution SPECT collimator can have more than 80000 air holes, which are by default all separate GEANT4 daughter volumes inside a mother collimator volume. The setup time becomes high, up to 70 minutes for a triple head camera. In order to reduce the overhead caused by this, a parameterized collimator setup was developed. It is based on the GEANT4 replica system in which a single volume represents multiple copies of a volume (the air holes) within its mother volume (the collimator itself). This can be done because the air hole locations follow a well defined translational symmetry. SPECT collimator geometries are built by this novel approach in less than a second.

6.3 Output data handling

During a simulation, data output is collected in the form of ROOT files (20) that contain detailed information about the particle interactions such as energy, position and the eventID. The data size can vary from a few kilobytes to several gigabytes depending on the type and duration of the acquisition. Figure 6.3 shows the functionality of the output merger which uses the ROOT files from the parallelized simulations as input. The information of the split file, generated by the job splitter (section 6.2.1), is used to merge the ROOT files into a single output file. The eventIDs will not be the same as in a single CPU simulation however, since this is reset at the beginning of each new run and each virtual time slice. This is solved by using the last eventID of each subsimulation as an offset for the next subsimulation. In case any simulation failed to complete, the output merger will detect the incomplete or missing output file. Using the information in the split file and the fully resolved macros, it is possible to restart only that specific part of the experiment.

In (21) it was shown that the output file merging is responsible for a large amount of the total computation time. This was mainly due to the large data transfers from hard drive to memory and vice versa while merging the output data into a single file with corrected eventIDs. A fast file merger was developed to overcome this limitation. Its functionality is shown in figure 6.4. The fast merger opens each output file in succession. The corrected eventIDs are stored lo-

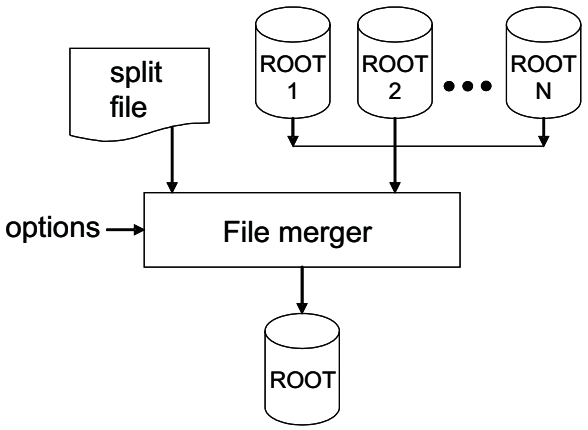


Figure 6.3: The output merger is used to merge the ROOT output data into a single output file.

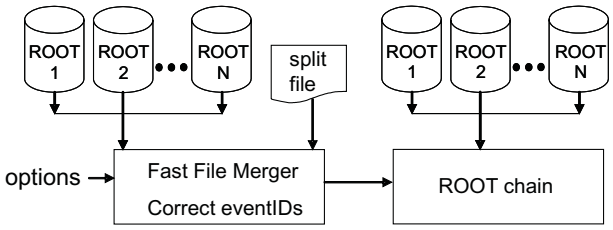


Figure 6.4: Schematic overview of the novel merging approach

cally in each output file by creating a new branch in the ROOT tree structure. The ROOT chain functionality can then be used to access the data for analysis (22). A chain is a list of ROOT files containing the same tree. The chain links the output files together and each file is opened in succession when a query is executed.

6.4 Mathematical model

In order to investigate the scalability of GATE simulations on a cluster, an acceleration factor (AF) is defined by equation 6.3. It consists

of the sum of the single CPU setup time $T_{s,setup}$ and the single CPU simulation time $T_{s,sim}$, divided by the sum of the split time T_{split}, the average setup time to build the geometry simultaneously on each CPU $\widehat{T}_{p,setup}$, the average parallel simulation time $\widehat{T}_{p,sim}$ and the total merge time T_{merge}.

An upper limit for the AF which does not take the split and merge time into account is given by AMdahl's law (AM) (23) as in equation 6.4 where p is the number of CPUs.

$$AF = \frac{T_{s,setup} + T_{s,sim}}{T_{split} + \widehat{T}_{p,setup} + \widehat{T}_{p,sim} + T_{merge}} \tag{6.3}$$

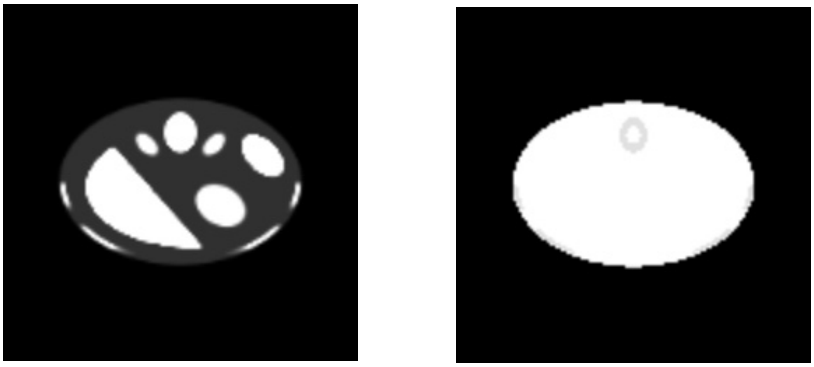
$$AM = \frac{T_{s,setup} + T_{s,sim}}{T_{s,setup} + T_{s,sim/p}} \tag{6.4}$$

6.5 Evaluation study

6.5.1 Benchmark simulations

The cluster was based on openMosix and consists of 38 nodes (24) with 17 dual XEON 2.4Ghz processors and 21 dual XEON 2.8Ghz processors, each with 2GB memory. Four series of benchmarks were run on this cluster. The first benchmark series was based on the Allegro (25; 26) PET scanner, serving as an example of a human PET scanner. A cylindrical water phantom was used with a 5 cm radius and 10 cm height. A 15 kBq $^{13}NH_3$ point source was placed in the center of this water phantom, in order to investigate the scalability when using an isotope with a short half life of approximately 10 minutes. The total acquisition time was 600 seconds, spanning one half life. The generated data output rate was 4.2 KB/s on a single CPU. The second benchmark was based on the same PET system with the crystals replaced by a perfect absorber to reflect high sensitivity. A 15 kBq planar flood source limited to 2D emission was used for a total acquisition time of 600 seconds. The generated data output rate was 174 KB/s on a single CPU. The third benchmark was based on the Axis SPECT system with a low energy high resolution (LEHR) collimator setup (26). The same phantom was used as for the Allegro benchmark, with a 15 kBq ^{99m}Tc point source for a total acquisition

time of 600 seconds. The fourth benchmark was based on the same Axis system but with the collimator modeled using the novel parameterized collimator implementation (cfr. 6.2.2). The data output rate for both implementations of the Axis system was 0.1 KB/s on a single CPU.



(a) Activity (b) Attenuation

Figure 6.5: a: Activity distribution for the slice of the MCAT phantom. b: Attenuation map of the slice.

The job splitter (cfr. 6.2.1) was used to generate macros and submit files to run the simulations on a linearly increasing number of CPUs and both output merging approaches (cfr. 6.3) were used to handle the output data.

6.5.2 Case study

In order to evaluate the improvement in image quality, projections of the activity distribution in a slice of the Mathematical Cardiac Torso (MCAT) (27) were simulated using a single CPU and on a 60 CPU cluster. A ^{67}Ga source of 37 MBq was used with an acquisition time of 60 seconds. The activity distribution map and the attenuation map are shown in figure 6.5. The single CPU simulation contained 1/60th of the activity of the 60 CPU case in order to maintain a comparable simulation time. The detector was binned at 128x128 (4.6875 mm) and both photopeaks were used: 93 kev (20% energy window) and 185 kev (15% energy window). The job splitter was used to run the simulations on the cluster and the data output was merged using

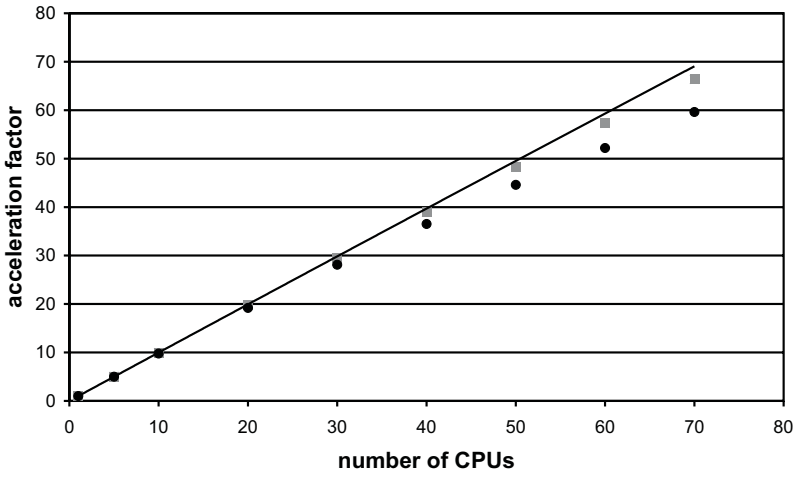


Figure 6.6: Scalability for the Allegro PET system: AFs with original merger (●); AFs with the fast merger (▪); AM (—)

the fast merger. The images were reconstructed using 10 iterations of standard MLEM (28), post-smoothed with a gaussian (4 pixels wide).

6.6 Results

Figure 6.6 shows the measured acceleration factors (AF) when compared to a single CPU for the Allegro PET system. The measured AFs are shown for both merger implementations as wel as the AMs. The AF for the maximum number of CPUs of 70 was 66. The overhead as a percent contribution to the total application time and as a function of the number of jobs is depicted in figure 6.7. The overhead by the setup and split time can be neglected but the single file output merger is responsible for over 11% overhead when using 70 CPUs. This merge time overhead drops to less than 2% when using the fast merger. Figures 6.8 and 6.9 show the corresponding results for the high data output rate PET system. The maximum AF obtained with the single file output merger remains below 8 for 70 CPUs. The merge time overhead rises over 89% for 70 CPUs. The fast merger significantly reduces this overhead to 14% for 70 CPUs resulting in

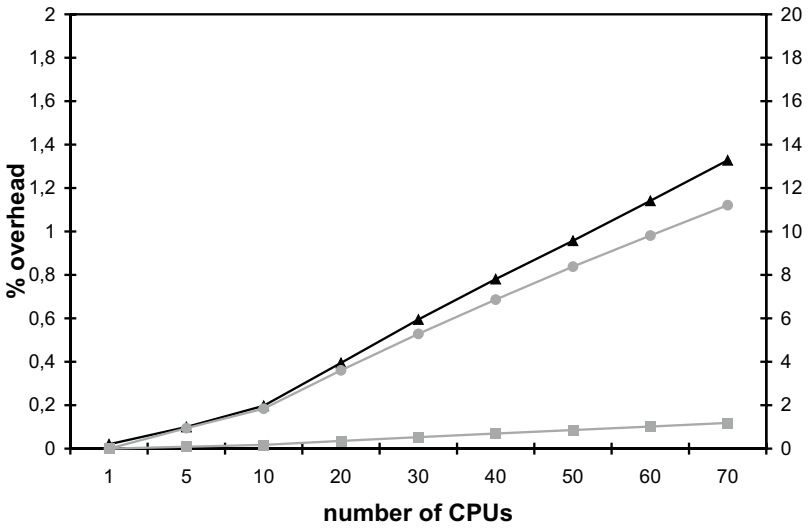


Figure 6.7: Overhead for the Allegro PET system. Left axis: setup overhead (▲). Right axis: merge overhead (●); fast merge overhead (■).

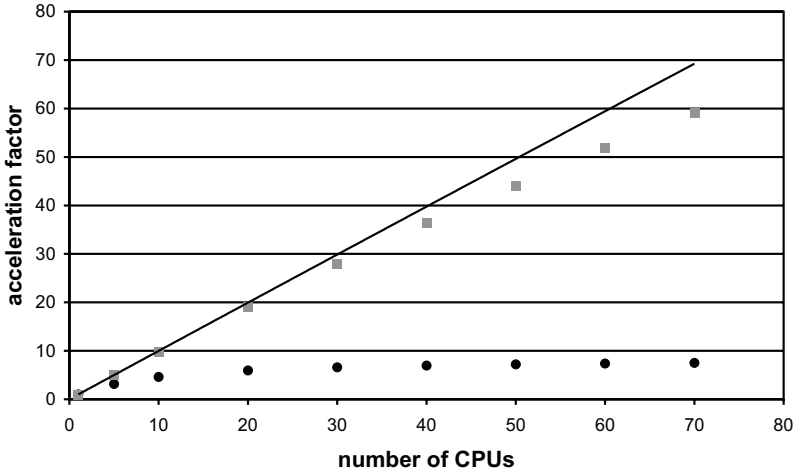


Figure 6.8: Scalability for the PET system with high data output rates: AFs with original merger (●); AFs with the fast merger (■); AM (—)

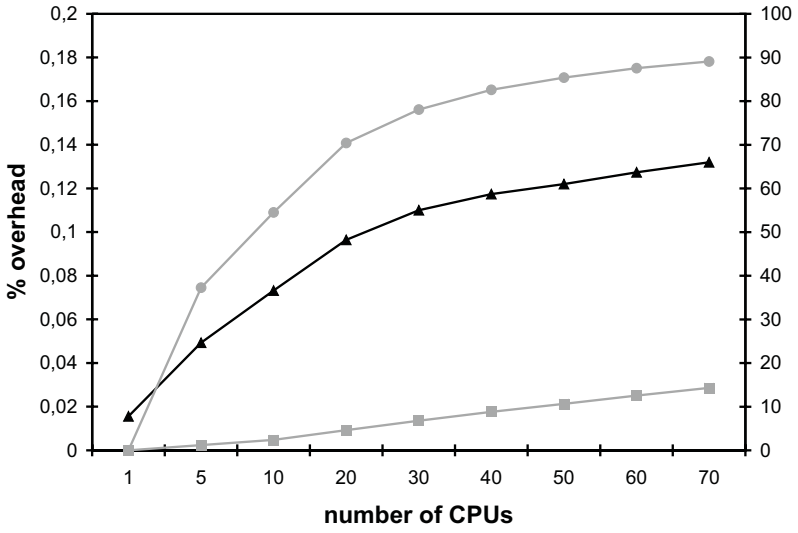


Figure 6.9: Overhead for the PET system with high data output rates. Left axis: setup overhead (▲). Right axis: merge overhead (●); fast merge overhead (■).

an AF of 59 for 70 CPUs.

Figure 6.10 shows the AFs and AMs for the Axis system with both the parameterized and the non-parameterized collimator setup. The maximum obtained AF for the non-parameterized collimator setup is 4.5 for 70 CPUs. Figure 6.11 shows that the setup overhead is 94.8% for 70 CPUs in this case. The maximum obtained AF for the parameterized collimator setup was 65.6 for 70 CPUs and the setup overhead shown in figure 6.12 decreased to 5.6% by using the parameterized collimator setup. The results for the fast merger overhead completely coincide with the default merger overhead (not shown) due to the low data output rate.

The reconstructed slice of the MCAT phantom with one CPU and with sixty CPUs is shown in figure 6.13a and b respectively. The single CPU simulation time was 659 minutes and 56 seconds. The average simulation time per CPU on the cluster was 655 minutes and 19 seconds, with a fast merge time of 32 minutes and 34 seconds.

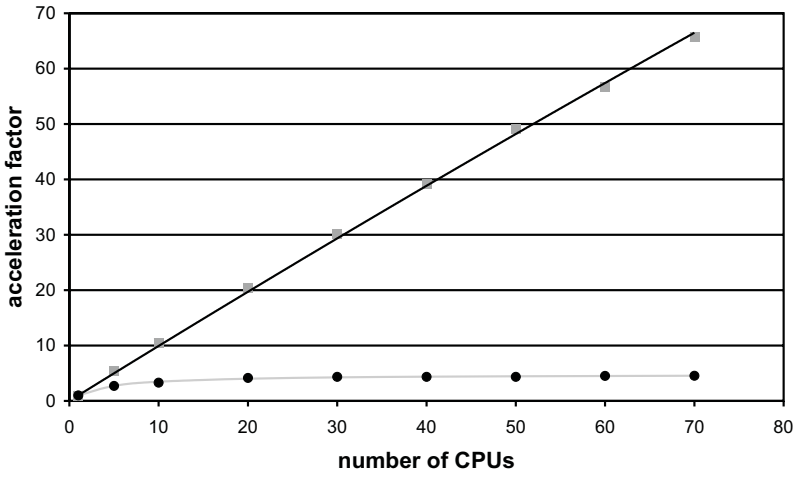


Figure 6.10: Scalability for the Axis SPECT system with the parameterized and non-parameterized collimator setup: AFs and AM for the parameterized setup (■, —); AFs and AM for the non-parameterized setup (●, —)

6.7 Discussion

If only the actual simulations were considered without setup time or merge time overhead, then the AFs would scale almost linearly. The time splitting is based on an equal amount of decays for each simulation. Therefore, the same computation time is needed for each CPU and no extra overhead is introduced even when using an isotope with a short half life.

PET acquisitions are typical examples of simulations with a high data output rate due to a higher sensitivity and because singles and coincidences can both be stored. The split time is negligible compared to the merge time and the same is true for the setup overhead introduced by the GATE scripts. The merge time however depends on the amount of data to be merged. This is why the percentual merge time overhead becomes larger for an increasing amount of CPUs, as the simulation time decreases as a function of the number of CPUs. The merge time is thus independent of the number of CPUs used in the simulation. The scalability is strongly dependent on the type of merger used. Small deviations from the AMs were already

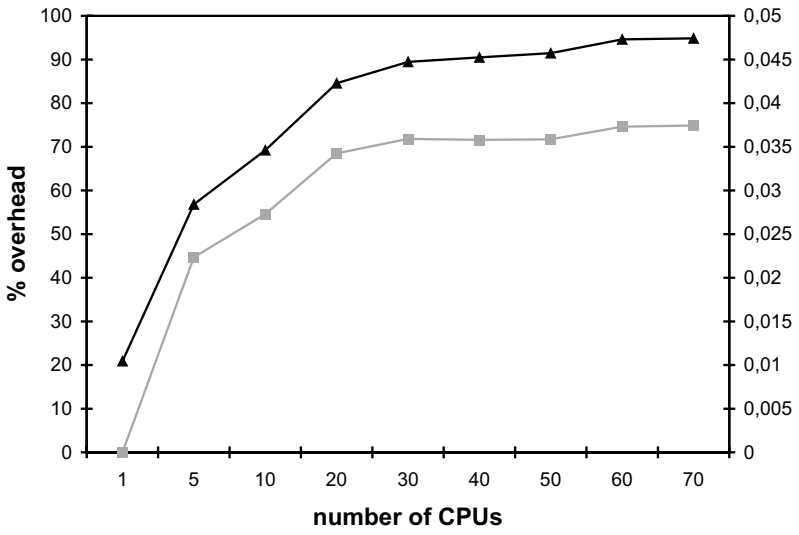


Figure 6.11: Overhead for the Axis SPECT system with the standard collimator setup. Left axis: setup overhead (▲). Right axis: merge overhead (■).

observed for the Allegro benchmark when using the default merger. The deviations were far more prominent in the high data output experiment consisting of a high sensitivity PET setup. The overhead for the latter system rose up to 90% for 70 CPUs resulting in an AF of less than 10. However, the fast merger significantly reduces this overhead, which results in good scalability. Although the merge time overhead has been significantly reduced, it is still higher than for the Allegro due to the much larger amount of data output.

SPECT acquisitions are typical examples of low sensitivity systems with relatively small data output sizes for a long simulation time. However, it could be concluded that the measured AFs deviate significantly from the ideal behaviour. The reason for this is the large collimator setup time required by the GATE scripts which, in contrast to the PET experiments, cannot be neglected. The use of our parameterized collimator setup greatly reduces this overhead, which results in excellent scalability. It was also shown that Amdahl's law can be used for both Axis setups to predict the scalability.

It is useful to determine an optimal number of CPUs to be used

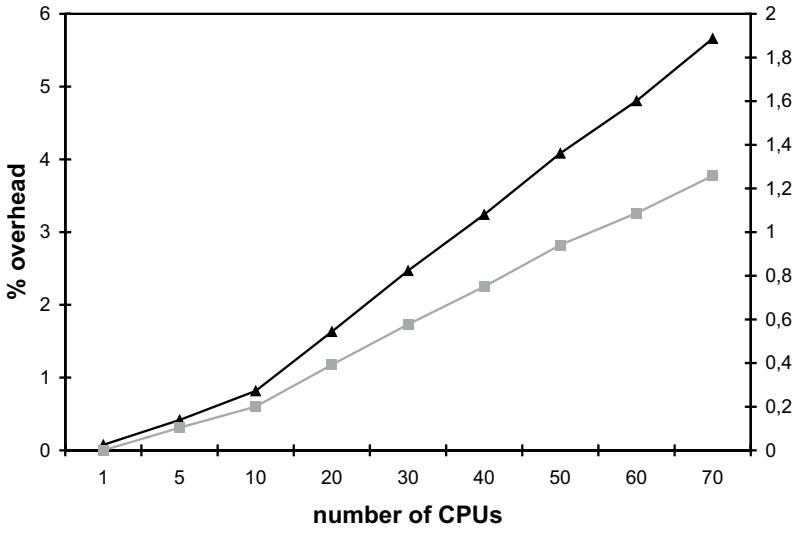


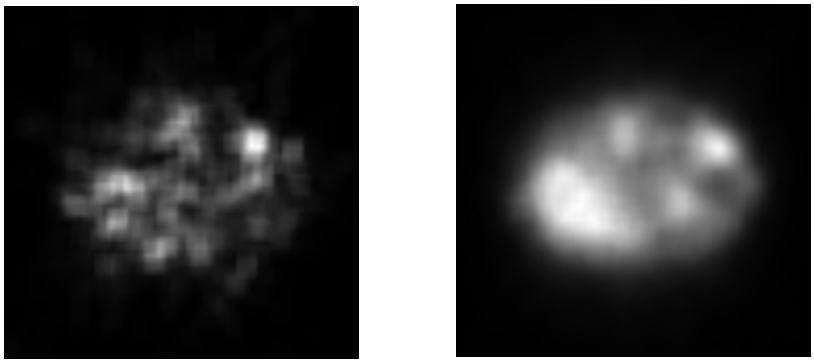
Figure 6.12: Overhead for the Axis SPECT system with the parameterized collimator setup. Left axis: setup overhead (▲). Right axis: merge overhead (■).

for large clusters. A general Acceleration Estimate (AE) for larger clusters can be derived by inclusion of the merge time T_{merge} and the split time T_{split} into Amdahl's law, as in equation 6.5.

$$AE = \frac{T_{split} + T_{s,setup} + T_{merge} + T_{s,sim}}{T_{split} + T_{s,setup} + T_{merge} + T_{s,sim/p}} \tag{6.5}$$

Both the total simulation time and the data output sizes of a GATE simulation are fairly predictable on a single CPU by running a short test simulation. The merge step only depends on the output size to be merged and the split time can be neglected as it is in the order of seconds. Figure 6.14 shows for example how the AEs are a close match to the AFs measured for the high sensitivity PET benchmark.

The development of the fast file merger and the parameterized collimator setup results in excellent overall scalability for both PET and SPECT systems. In addition, the GATE cluster framework is compatible with the variance reduction techniques discussed in chapters 4 and 5 and the optimized navigation algorithms discussed in chapter 5.



(a) One CPU (b) Sixty CPUs

Figure 6.13: Reconstruction of a slice of the MCAT phantom. a: Single CPU case. b: Sixty CPU case.

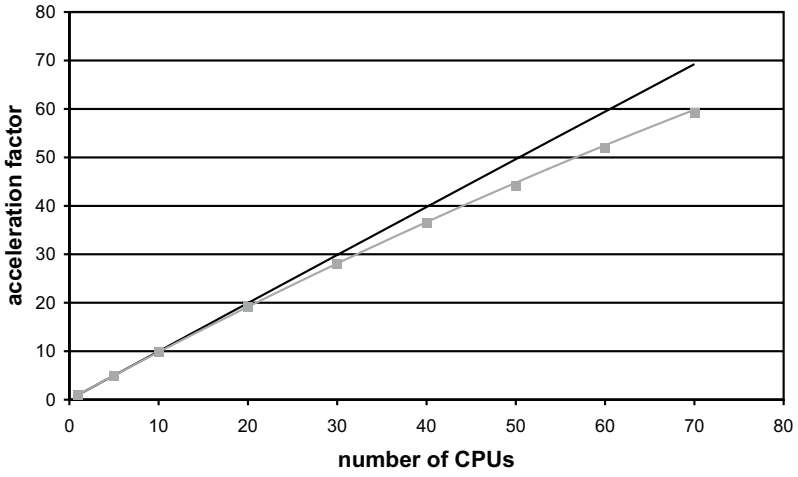


Figure 6.14: AFs (▪), AEs (—) and AM (**—**) for the PET system with high data output rates and fast merger.

The reconstructed slices of the MCAT phantom clearly show the usefulness of our framework. The image generated with the 60 CPU cluster contains much less noise for a comparable simulation time. Using our framework it was possible to finish the simulations overnight.

6.8 Summary and original contributions

In this chapter we investigated the acceleration of GATE Monte Carlo simulations through the use of parallelization. A distributed computing framework for executing GATE simulations on a cluster of computers was designed and developed. Our approach is composed of two steps: job distribution and output data handling.

A time-domain partitioning scheme with virtual time slices is used for job distribution. In this way, equal computation time is assigned to each CPU, even for isotopes with a short half life. Independent random number streams are produced to guarantee that each subsimulation is statistically independent. This approach retains all experimental parameters such as singles, randoms and detector deadtime. In addition, a parameterized collimator model was developed to reduce the setup time required by SPECT simulations, which would otherwise impose a limit on the scalability.

Large data files created by each of the sub simulations contain detailed information about particle interactions such as energy and detection position. This data is however related to virtual time slices. Therefore it is necessary to merge the data according to the real time schedule. We found the output file merging to be responsible for a large amount the total computation time. To overcome this limitation in scalability, we developed a fast data output merger that creates a chain of the already existing files with the required time schedule corrections.

The scalability of our method was investigated with benchmark simulations on an increasing number of CPU's. Firstly, a PET setup consisting of a $13^N H_3$ point source in a water phantom was simulated. An acceleration factor of 66 on 70 CPU's was observed with a drop in file merger overhead from 11% to only 2%. Secondly, a high sensitivity PET setup with a planar flood source was simulated. A maximum acceleration factor of 59 on 70 CPU's was found, with a drop in overhead from 89% to 14%. We found that the percent merge time overhead becomes larger for an increasing amount of CPU's. Thirdly, we investigated the influence of the parameterized collimator model. Two SPECT setups of a water phantom with a ^{99m}Tc point source and a parameterized and a non-parameterized collimator were simulated. An acceleration factor of 65.6 on 70 CPU's was found with a drop in setup time overhead from 94.8% to 5.6%.

The potential gain in image quality was investigated with a case

study. Projections of the ^{67}Ga activity distribution in a slice of a torso phantom were simulated using a single CPU and on a 60 CPU cluster. The results were compared for approximately equal simulation time. Therefore the single CPU simulation consisted of $1/60th$ of the activity of the 60 CPU case. Visual inspection revealed much less noise in the 60 CPU case, clearly pointing out the usefulness of our framework.

Due to the developments in this chapter both PET and SPECT GATE simulations now scale almost linearly with cluster size.

The work presented in this chapter resulted in the publication of two conference proceedings (21; 29) and an A1 journal paper (30). Part of this work also appeared in another A1 journal paper (31).

References

[1] L. Deng and Z.S. Xie. *Parallelization of MCNP Monte Carlo neutron and photon transport code in parallel virtual machine and message passing interface*. J. Nucl. Sci. Tech., 36(7):626–629, 1999.

[2] F.B. Brown et al. *MCNP Version 5*. Trans. Am. Nucl. Soc., 87:273–381, 2002.

[3] Jeffrey M. Squyres and Andrew Lumsdaine. *A Component Architecture for LAM/MPI*. In Proceedings, 10th European PVM/MPI Users' Group Meeting, number 2840 in Lecture Notes in Computer Science, pages 379–387, Venice, Italy, 2003. Springer-Verlag.

[4] S. M. Robinson, R. J. McConn Jr, R. T. Pagh, J. E. Schweppe, and E. R. Siciliano. *Characteristics of multiprocessing MCNP5 on small personal computer clusters*. J. Instrum., 1:P06001, 2006.

[5] H. Zaidi, C. Labbé, and C. Morel. *Implementation of an environment for Monte Carlo simulation of fully 3D positron tomography on a high-performance parallel platform*. Parallel Comput., 24:1523–1536, 1998.

[6] R.L. Harrison, S. Dhavala, P.N. Kumar, R. Yiping Shao Manjersshwar, T.K. Lewellen, and F.P. Jansen. *Acceleration of SimSet photon history generation.* In Nuclear Science Symposium Conference Record, volume 3, pages 1835–1838, 2002.

[7] M.G. Thomason, R.F. Longton, J. Gregor, G.T. Smith, and R.K. Hutson. *Simulation of emission tomography using grid middleware for distributed computing*. Comput. Methods Programs Biomed., 75(3):251–258, 2004.

[8] M. Ljungberg. *The SIMIND Monte Carlo program*. In M. Ljungberg, S.E. Strand, and M. A. King, editors, Monte Carlo calculations in nuclear medicine: Applications in diagnostic imaging, pages 145–163, Bristol, 1998. Institute of Physics Publishing.

[9] Y.K. Dewaraja, M. Ljungberg, A. Majumdar, A. Bose, and K.F. Koral. *A parallel Monte Carlo code for planar and SPECT imaging: implementation, verification and applications in (131)I SPECT*. Comput. Methods Programs Biomed., 67(2):115–124, 2002.

[10] J. Montagnat, F. Bellet, H. Benoit-Cattin, V. Breton, L. Brunie, H. Duque, Y. Legr, I.E. Magnin, L. Maigne, S. Miguet, J.M. Pierson, L. Seitz, and T. Tweed. *Medical images simulation, storage, and processing on the European DataGrid testbed*. Journal of Grid Computing, 2(4):387–400, 2004.

[11] http://openmosix.sourceforge.net.

[12] D. Thain, T. Tannenbaum, and M. Livny. *Distributed computing in practice: the Condor experience.* Concurrency - Practice and Experience, 17:323–356, 2005.

[13] http://www.openpbs.org.

[14] http://www.apple.com/acg/xgrid/.

[15] CLHEP - A Class Library for High Energy Physics, http://cern.ch/clhep.

[16] F. James. *A Review of Pseudorandom Number Generators*. Computer Phys. Comm., 60:329–344, 1990.

[17] G. Marsaglia and A. Zaman, Toward a Universal Random Number Generator, Florida State University FSU-SCRI-87-50 (1987).

[18] M. Traore and D. Hill. *The Use of Random Number Generation for Stochastic Distributed Simulation: Application to Ecological Modeling*. In European Simulation Symposium, volume 13, pages 555–559, 2001.

[19] S. Staelens, D. Strul, G. Santin, S. Vandenberghe, M. Koole, Y. D'Asseler, I. Lemahieu, and R. Van de Walle. *Monte Carlo simulations of a scintillation camera using GATE: validation and application modelling*. Phys. Med. Biol., 48(18):3021–3042, 2003.

[20] R. Brun and F. Rademakers. *ROOT - An Object Oriented Data Analysis Framework*. In Nucl. Inst. & Meth. in Phys. Res. A, volume 389, pages 81–86, 1997.

[21] J. De Beenhouwer, D. Kruecker, S. Staelens, L. Ferrer, A.F. Chatziioannou, and F.R. Rannou. *Distributed Computing Platform for PET and SPECT Simulations with GATE*. In Proceedings of the IEEE Nuclear Science Symposium and Medical Imaging Conference, volume 94, pages 2437–2440, 2005.

[22] ftp://root.cern.ch/root/doc/chapter12.pdf.

[23] G.M. Amdahl. *Validity of single-processor approach to achieving large-scale computing capability*. In Proceedings of AFIPS Conference, pages 483–485, Reston, VA., 1967.

[24] Manufactered by Supermicro Computer, Inc , 980 Rock Ave. San Jose, CA 95131, USA.

[25] S. Surti and J.S. Karp. *Imaging characteristics of a 3-dimensional GSO whole-body PET camera*. J. Nucl. Med., 45:1040–1049, 2004.

[26] Philips Medical Systems, 595 Miner Road, Cleveland, OH 44143, USA.

[27] B.M.W. Tsui, J.A. Terry, and G.T Gullberg. *Evaluation of cardiac cone-beam SPECT using observer performance experiments and ROC analysis*. Investigative Radiology, 28(12):1101–1112, 1993.

[28] L. Shepp and Y. Vardi. *Maximum likelihood reconstruction for emission tomography*. IEEE Trans. Med. Imaging, 1(2):113–122, 1982.

[29] J. De Beenhouwer, S. Staelens, Y. D'Asseler, and I. Lemahieu. *Optimizing the Scalability of Parallelized GATE Simulations*. In Proceedings of the IEEE Nuclear Science Symposium and Medical Imaging Conference, pages 3904–3908, 2006.

[30] J. De Beenhouwer, S. Staelens, D. Kruecker, L. Ferrer, Y. D'Asseler, I. Lemahieu, and F.R. Rannou. *Cluster computing software for GATE simulations*. Med. Phys., 34(6):1926–1933, 2007.

[31] S. Staelens, J. De Beenhouwer, D. Kruecker, L. Maigne, F. Rannou, L. Ferrer, Y. D'Asseler, I. Buvat, and I. Lemahieu. *GATE: Improving the computational efficiency*. Nucl. Instr. and Meth. A, 569:341–345, 2006.

7
General Conclusion

In this final chapter we summarize the main contributions of this work and we discuss some aspects for future study.

Chapter 2 gave an overview of Monte Carlo simulations in nuclear medicine. A thorough literature study was presented, introducing the basic principles and applications of both PET and SPECT. The Monte Carlo principle itself was explained and the relevance of Monte Carlo simulations to nuclear medicine was discussed subsequently. It has applications in detector design, quantification, correction methods for image degradations, detection tasks etc. The benefits and drawbacks of the current Monte Carlo simulators were discussed and GATE was introduced as our Monte Carlo simulator of choice. The GEANT4 Monte Carlo model that forms the basis of GATE was then presented. Low energy photon interaction models were explained and particle transportation was discussed. After a detailed explanation of the basic features of GATE, the final part of the chapter focused on the Achilles' heel of GATE: efficiency. It became apparent that the acceleration of GATE could only be achieved through a combination of efficient data analysis, dedicated variance reduction techniques, fast navigation algorithms and parallelization. These topics formed the basis of the subsequent chapters.

Chapter 3 focused entirely on the analysis capabilities of GATE

with the development of an analysis framework. With this framework it is possible to analyse any process, from the decay of isotopes to particle interactions and detections in any detector element for any type of phantom in a GATE Monte Carlo simulation. An assessment was made of spurious activity in ^{124}I-Bexxar PET and of contamination in ^{131}I-Bexxar SPECT. In the case of PET we have shown that for the first time with any Monte Carlo simulator, it is possible to detect spurious coincidences for realistic phantoms. Optimized energy thresholds were derived and used for the simulation of a torso phantom. The energy thresholds improved image quality but some spurious activity remained. We have also shown that the spurious activity could not be assumed to be spatially uniform. As a result, the standard reconstruction and correction techniques are not adequate. In the future, a dedicated single scatter algorithm that incorporates spurious activity could provide better results. In the case of SPECT our framework allowed an assessment of contamination by an accurate classification of detections. It was found that the correction algorithms used could not correct for septal penetration. The comparison of PET and SPECT showed that ^{124}I with optimized energy thresholds offer better image quality when standard reconstruction techniques are used. Our framework was not only used as a validation tool and efficient analysis module. It is currently being used as the primary analysis tool for nearly all Monte Carlo simulations at our Medical Image and Signal Processing group. Its applications range from the fast simulation of Bremsstrahlung to the investigation of contamination for different isotopes with realistic phantoms and prototype detectors. In addition it will be released to the GATE community in the near future.

In chapter 4 we introduced Geometrical Importance Sampling (GIS) as a general acceleration technique for GATE SPECT simulations. For a low energy high resolution collimator, only 0.02% of all emitted photons actually reach the crystal surface. The majority of photons is therefore tracked in vain as they do not contribute to the result. GIS is a variance reduction technique that actually solves this problem by increasing the survival rate of photons in regions close to the SPECT detector head, and by decreasing the survival rate of photons traveling away from the detector head. As a result, branches were introduced into the particle history. A solution for this problem was developed based on a history tree with adjustments for splitting and Russian Roulette. A validation study was carried out with re-

gard to energy spectra, spatial resolution, and sensitivity for a low (^{99m}Tc) and medium energy isotope (^{67}Ga) in a water phantom and in air. A good agreement with analog simulations was found. A figure of merit was verified based on a study of the variance of a flux on a detector and it was shown that this technique can result in a 5 to 13-fold increase over analog simulations. The main benefit of GIS lies in its generality. It provides an increase in efficiency for any combination of SPECT isotope, phantom and detector. A relatively simple importance map was used in this study, which led to an equal weight for all detected photons. The particle history detangling is however prepared for weight distributions as a result of more complex importance maps. The derivation of an optimal importance map and the investigation of the potential increase in efficiency is the subject of future research, along with the adaption of GIS to PET.

In chapter 5 the acceleration of ^{99m}Tc SPECT simulations was investigated. Our approach consisted of two steps. *In the first step*, forced detection (FD) and convolution-based forced detection (CFD) was used to increase the proportion of photons that result in a detection. Our approach is unique in two ways. *Firstly*, photons are transported through the phantom in a direction sampled within a solid angle toward the SPECT detector head. Therefore, in the case of FD all detector interactions are simulated by analog Monte Carlo. In the case of CFD, an analytical detector model is used. *Secondly*, for both FD and CFD no approximations were used to sample the attenuation along the photon path. The weights are therefore compliant with the low energy model used in GEANT4. *In the second step*, the particle navigation algorithms in GEANT4 were optimized for efficiency with FD and CFD. The validation showed that CFD is over 900 times faster than FD but that it is limited by its analytical description of the detector. Spatially both FD and CFD closely match analog GATE simulations and an increase in efficiency from three to six orders of magnitude was realized. This allows for the simulation of a realistic acquisition of a complex torso phantom within 130 seconds. While FD will be used in simulation studies, CFD will be used as a forward projector in iterative reconstruction. The next step in the development of these techniques is the extension to higher energy isotopes such as ^{111}In and ^{131}I and to fan beam collimators.

In chapter 6 we investigated the acceleration of GATE Monte Carlo simulations through the use of parallelization. Our approach is composed of two steps: job distribution and output data handling.

The job distribution is based on a time-domain partitioning scheme that retains all experimental parameters which guarantees the statistical independence of each subsimulation. A parameterized collimator model was also developed to reduce the setup time required by SPECT simulations. The data output is handled by a fast output merger that creates a chain of the already existing files with the required time schedule corrections. A validation study of our framework showed an excellent scalability with an acceleration factor of approximately 66 on 70 CPU's for both PET and SPECT. Even in the case of a high sensitivity PET simulation the acceleration was 59 on 70 CPU's. The potential gain in image quality was investigated with the simulation of a ^{67}Ga activity distribution in a torso phantom on one CPU and on 60 CPUs with equal simulation time. Visual inspection revealed much less noise in the 60 CPU case, clearly pointing out the usefulness of our framework. Due to the developments in this chapter both PET and SPECT GATE simulations now scale almost linearly with cluster size. In addition, the cluster framework is compatible with the developments in all previous chapters.

The focus of this dissertation was on the acceleration of GATE Monte Carlo simulations in such a way that the applicability of GATE as a general PET and SPECT Monte Carlo simulator is retained. A suitable acceleration technique can be chosen, depending on the required speed and accuracy. The acceleration of GATE has been one of the most important research aspects of the OpenGATE collaboration. To our knowledge, the work in this dissertation presents the first use of GATE within clinically useful simulation times. A fast GATE simulator does not only enhance the practical usability of GATE. It allows GATE to be incorporated into reconstruction algorithms which in turn will lead to better reconstructed image quality and the continued position of GATE as golden standard in PET and SPECT simulators.

A

Integration into GATE

A.1 Structural overview

Each of the techniques discussed in chapters 3 to 6 was incorporated into a single version of GATE in order to retain the general applicability of GATE as a PET and SPECT Monte Carlo simulator. Therefore it is possible to define a specific phantom and detector model first and then choose a suitable acceleration method depending on the trade-off between the required speed and accuracy. Only a few additional lines of macro code are required. System design can be performed by aGATE. If enhanced analysis is mandatory, it can be obtained with ProcessGATE. If on the other hand a faster simulation is desired, it can be obtained with GIS. SPECT protocol optimization can be performed by FD, while CFD can be used for model-based reconstruction. In addition, the parallelization in chapter 6 can be used in combination with the methods of all previous chapters.

An overview of the structure of the code is shown in figure A.1. It is a simplified representation of the logical connections between the most important building blocks. When FD or CFD are enabled through a macro command, the regular GATE physics code is replaced by a specialized versions for FD or CFD, that incorporate the

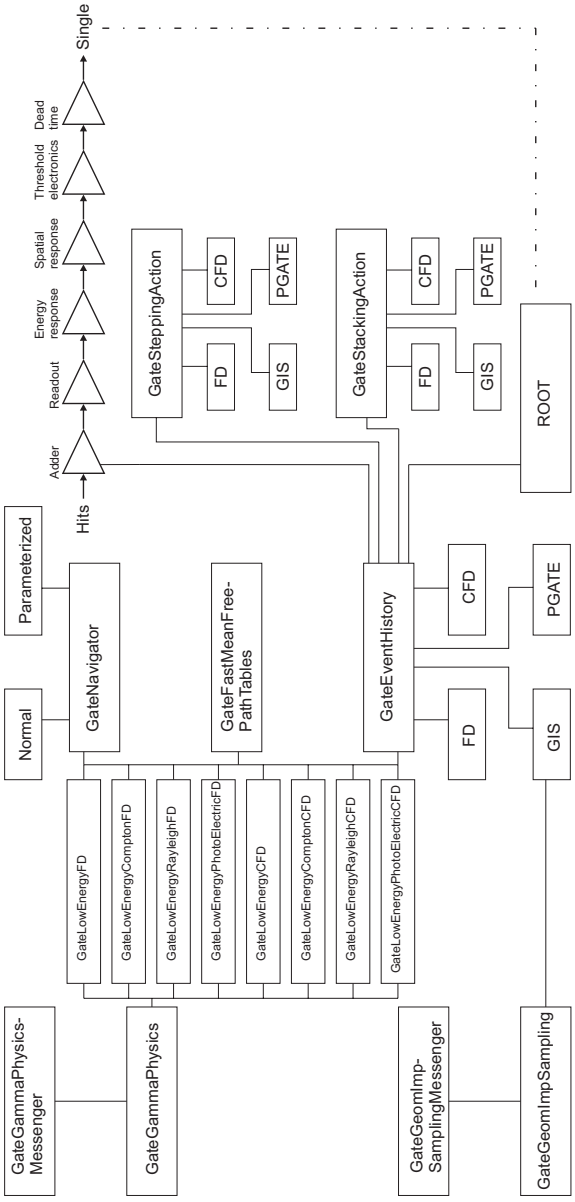


Figure A.1: Logical structure of the relationship between the different classes developed for GIS, ProcessGATE, FD and CFD.

weight calculations and probability tables discussed in chapter 5. Implementations are available for Compton scatter and Rayleigh scatter. The classes 'GateLowEnergyFD/CFD' are used to create copy photons at decay. The classes 'GateLowEnergyPhotoElectricFD/CFD' implement the photoelectric effect for FD/CFD. In the case of FD, secondary particles are not generated for copy photons inside the phantom. The generation of secondary particles is enabled however for tracking inside the detector. In the case of CFD, the generation of secondary particles is disabled completely. Each of the physics classes uses the fast mean free path tables discussed in chapter 5. In the case of FD and CFD, the navigation is performed by aGATE for mother photons and for FD copy photons in the detector. Copy photons in the phantom are handled by the specialized navigator discussed in chapter 5. The specialized navigator has two separate implementations for tracking in normal geometry (spheres, cylinders, boxes,..) or parameterized geometry (voxel grids, collimator holes,..). In the case of GIS, a specialized GEANT4 navigator is used for tracking in virtual geometry as discussed in chapter 4.

The central part of the code is formed by 'GateEventHistory'. This part of the code gathers information from three fronts. Firstly, macro commands are interpreted through messenger classes. This way, a specialized version of this class is used whenever FD, CFD, GIS or ProcessGATE are enabled. Secondly, all FD/CFD physics classes and the GIS classes are connected to this central point. This way, specific information can be stored such as information concerning the copy photons or split photons (weight, ID, parent photon,..). Thirdly, information is gathered during the tracking of each particle through the geometry. This is facilitated by FD, CFD, GIS or ProcessGATE versions of 'GateSteppingAction' and 'GateStackingAction'. In the case of FD, CFD and GIS this enables step by step control over particles after they were created by the physics classes. In the case of ProcessGATE this enables the detailed analysis information to be stored temporarily as discussed in chapter 3.

At the end of each event, the information gathered from the three fronts is processed by each of the four possible versions of 'GateEventHistory'. For ProcessGATE, this allows the user-defined algorithms to classify detections as discussed in chapter 3. For GIS and FD, this enables particle history detangling as discussed in chapters 4 and 5. For CFD, this enables the depth dependent subprojection maps discussed in chapter 5. After processing, the information may be dis-

persed either directly to ROOT output or it may first be used in the GATE digitizer chain. The adder module then uses this information to form pulses from hits as discussed in chapter 2.

A.2 Example macros

A.2.1 Macro commands for ProcessGATE

```
/gate/processGate/detectFluorescence
/gate/processGate/detectPenetration
/gate/processGate/detectCollimatorScatter
#/gate/processGate/detectFullSpect
```

The first three commands enable the detection of collimator fluorescence, penetration and Compton and/or Rayleigh scatter. The fourth command replaces the other commands and enables the detection of penetration in the collimator and fluorescence and scatter in all phantom and detector volumes.

A.2.2 Macro commands for GIS

```
/gate/gis/enable
#cylinders for the first head
/gate/gis/insertCylinder
/gate/gis/setHeight 70. cm
/gate/gis/setInnerRadius 0. cm
/gate/gis/setOuterRadius 2. cm
/gate/gis/setStartAngle -45.  deg
/gate/gis/setSpanningAngle 90. deg
/gate/gis/setTranslation 0. 0. 0. cm
/gate/gis/setSpeed  2. deg/s
/gate/gis/setPoint1 0 0 0 cm
/gate/gis/setPoint2 0 0 1 cm
/gate/gis/setImportance 1
```

After explicitly enabling GIS, the virtual importance map is defined for each detector head. In this case cylinders are used. The macro commands above are used for the first cylinder. Consecutive cylinders can be defined after each other. The usual properties of a cylinder are set as defined in the GEANT4 manual: height, inner and

outer radius, start angle and spanning angle. In addition, the rotation speed is set in order to comply with the movement of the detector heads over time. Lastly, the importance value is set. Currently all integer values are accepted, but only binary splitting was validated in chapter 4.

A.2.3 Macro commands for FD

```
/gate/physics/gamma/enableFD
/gate/physics/gamma/selectPhotoelectric lowenergy
/gate/physics/gamma/selectCompton lowenergy
/gate/physics/gamma/selectGammaConversion inactive
/gate/physics/gamma/selectRayleigh lowenergy
```

After explicitly enabling the FD physics, it is possible to enable or disable Compton or Rayleigh scatter.

```
/gate/forceddetection/enable
/gate/forceddetection/setParallelBeam 1
/gate/forceddetection/setnHeads 60
/gate/forceddetection/setOpeningAngle 2.6
/gate/forceddetection/setAngle 0
/gate/forceddetection/setStepAngle 6
/gate/forceddetection/setCollimatorFrontFace -240.6835
/gate/forceddetection/setCollimatorBackFace -250.6835
/gate/forceddetection/setVoxelTracking 1
/gate/forceddetection/setVoxelTrackingGridSize 128
/gate/forceddetection/setVoxelTrackingPixelSize 3.17
```

The commands above first explicitly enable FD. The second command allows the choice between parallel beam and fan beam collimators. Only parallel beam collimators have been validated for now however. Next, the number of projections to be used with multiple projection sampling can be chosen, along with the FD cone half opening angle. The 'setAngle' command defines the starting angle of the detector heads. The 'StepAngle' command allows to set the degrees between each projection angle. The collimator front and back face is used to position copy photons in front of the collimator for analog tracking. Lastly, the 'VoxelTracking' command allows to choose between tracking in simple volumes or in a voxelized grid. In case of a voxelized grid, the grid size and the pixel size in mm must be set.

A.2.4 Macro commands for CFD

```
/gate/physics/gamma/enableCFD
/gate/physics/gamma/selectPhotoelectric lowenergy
/gate/physics/gamma/selectCompton lowenergy
/gate/physics/gamma/selectGammaConversion inactive
/gate/physics/gamma/selectRayleigh lowenergy
```

After explicitly enabling the CFD physics, it is possible to enable or disable Compton or Rayleigh scatter.

```
/gate/cfd/enable
/gate/cfd/setAngle 0
/gate/cfd/setOpeningAngle 0.1
/gate/cfd/setnHeads 60
/gate/cfd/setStepAngle 6
/gate/cfd/setCollimatorFrontFace -347.6835
/gate/cfd/setVoxelTracking 1
/gate/cfd/setVoxelTrackingGridSize 128
/gate/cfd/setVoxelTrackingPixelSize 3.17
```

The rest of the CFD code must be explicitly enabled. The CFD commands are equivalent to their FD counterparts.

A.3 GATE cluster

No macro commands are required to use the parallelization of chapter 6. The system was designed to be completely transparent to the user. Therefore, any macro using aGATE, ProcessGATE, GIS or FD may simply be submitted to the job splitter. In order to use Gate in cluster mode you need 3 components:

- The job splitter
- The file merger
- A cluster aware version of Gate

A.3.0.1 Installation of the job splitter

The job splitter can be installed in the same directory as Gate. Two environment variables need to be added to the evironment file used to compile Gate:

- export GC_DOT_GATE_DIR=/somedir/
- export GC_GATE_EXE_DIR=/somedir/bin/Linux-g++/

The first variable indicates the location of a hidden directory called .Gate. The directory will contain the splitted macros for each simulation. Even when splitting the same macro several times, a new directory will be created for each instance. In normal circumstances one does not need to look into it. In case of an error, it can be used to run only a specific part of a simulation again. The second environment variable indicates the location of the job splitter executable. As the Gate environment file will be used to compile the job splitter source code, the executable will likely be located in the same directory as the Gate executable.

To install, load the Gate/Geant4 environment variables and unpack the job splitter source code into its own directory (bash example):

$ *source environment_variables*
$ tar -zxvf jobsplitter.tar.gz
$ cd jobsplitter
$ gmake

A.3.0.2 Installation of the file merger

To install, unpack the file merger source code into its own directory and compile (bash example):

$ *tar -zxvf filemerger.tar.gz*
$ cd filemerger
$ gmake

The file merger executable is located in the current directory.

A.3.0.3 Preparing your macro

The cluster software should be able to handle all GATE macros. However, only ROOT is currently supported as an output format. Please disable other output formats as they cannot yet be merged. If an isotope with a short half life compared to the acquisition time is simulated, then it may be useful to specify the half life in your macro as follows:

- /gate/cluster/setTimeSplitHalflife 6600. s

This way, the load will be approximately equal for each CPU.

A.3.0.4 Using the job splitter

In order to view information on general usage, just run the job splitter executable without any options:
$ *./bin/Linux-g++/gjs*

gjs [-options] your_file.mac
options (in any order):
-a value alias : use any alias
-numberofsplits n : the number of job splits; default=1
-noseeds : do not create random seed files in the .Gate/job directory
-clusterplatform name: the cluster platform, name is one of the following:
openmosix - condor - openPBS - xgrid
This executable is compiled with openmosix as default
-openPBSscript : template for an openPBS script
see the example that comes with the source code
overrules the environment variable below
-condorscript : template for a condor submit file
see the example that comes with the source code
-v : verbosity 0 1 2 3 - 1 default
Environment variables:
GC_DOT_GATE_DIR : indicates the .Gate directory for splitted mac files
GC_GATE_EXE_DIR : indicates the directory with the Gate executable
optional GC_PBS_SCRIPT : the openPBS template script

Usage (bash):
export GC_DOT_GATE_DIR=/home/user/gatedir/
export GC_GATE_EXE_DIR=/home/user/gatedir/bin/Linux-g++/
Examples:
gjs -numberofsplits 10 -clusterplatform openmosix macro.mac
gjs -numberofsplits 10 -clusterplatform openmosix -a /somedir/rootfilename ROOT_FILE macro.mac
gjs -numberofsplits 10 -clusterplatform openPBS -openPBSscript /somedir/script macro.mac
gjs -numberofsplits 10 -clusterplatform xgrid macro.mac
gjs -numberofsplits 10 /somedir/script macro.mac
WARNING: just like in a normal GATE macro you do not use filename extensions for output formats (ROOT,INTERFILE,..)

The supported platforms are currently: openMosix, openPBS, Condor and Xgrid. In the case of openMosix:

```
$ ./bin/Linux-g++/gjs -numberofsplits 5 -clusterplatform
        openmosix macro.mac
```

The job splitter will subdivide the simulation macro into fully resolved, non-parameterized macros. In this case there are 5 such macros. They are placed under the .Gate directory, as specified by the GC_DOT_GATE_DIR environment variable. A list of all the data output options is given after successful completion:

ROOT output is enabled
ASCII output is disabled
INTER output is disabled
LMF output is disabled
ECAT output enabled if ECAT + sinogram selected
but no filename is given; using a default one
SINO output enabled if ECAT system selected
but no filename is given; using a default one

If an alias was expected for output files and it was not supplied then this will be mentioned in the output options list. A standard name will be supplied automatically, as well as appropriate numbering. The .Gate directory now has a subdirectory called macro, that contains:

macro1.mac
macro2.mac
macro3.mac
macro4.mac
macro5.mac
seed1.rndm
seed2.rndm
seed3.rndm
seed4.rndm
seed5.rndm
macro.split

The 5 macros are listed as well as 5 random seeds to initialize the random engine. The .split file contains information about the splitted simulation and can be used to merge the data after the simulation. The current directory, from which the jobsplitter was called, now contains the cluster submit file. In order to run the splitted simulation on the cluster, one only needs to execute this file: $ *chmod +x macro.submit*
$ *./macro.submit*

The .Gate directory supports automatic numbering. If the same macro is used repeatedly, then the subsequent directories will be numbered.

A.3.0.5 Using the file merger

The file merger can be run with either the original macro or the split file as input. In order to view information on general usage, just run the file merger executable without any options:

$ *./gjm*

gjm [-options] your_file.mac or your_file.split
You may either give the name of your gate macro file or the name of a split file created by gjs.

options:
-outDir path : where to save the output files default is PWD
-v : verbosity 0 1 2 3 - 1 default

-f : forced output - an existing output file will be overwritten
-cleanonly : do only a the cleanup step i.e. no merging
erase work directory in .Gate and the files from the parallel jobs
-cleanonlyTest : just tells you what will be erased by the -cleanonly
-clean : merge and then do the cleanup automatically
-fastMerge : merge and then do the cleanup automatically"¡¡endl;
environment variable:
GC_DOT_GATE_DIR : points to the .Gate directory

In order to merge the output files into a single file, just supply the macro file or the split file to the file merger. The output file can be used as a usual single CPU output file. For some older ROOT versions it may be necessary to use the -maxRoot option to limit the ouput file size in MB:

$ *./gjm macro.split or ./gjm -maxRoot 1900 macro.mac*

In case a single output file is not required, it is possible to use the option 'fastMerge'. This way the eventIDs in the ouputfiles are corrected locally. A ROOT chain, which is a list of files containing the same tree, is then required to link the output files together for analysis. A chain for the Singles could be made as follows (in a file called chain.c):

```
{
gROOT->Reset();
TChainchain("Singles");
chain.Add("rootf1.root");
chain.Add("rootf2.root");
chain.Add("rootf3.root");
chain.Add("rootf4.root");
chain.Add("rootf5.root");
}
```

Once all files are added to the chain, one can use the chain as a regular Ttree.

Journal Papers

[1] J. De Beenhouwer, S. Staelens, D. Kruecker, L. Ferrer, Y. D'Asseler, I. Lemahieu, and F.R. Rannou. *Cluster computing software for GATE simulations*. Med. Phys., 34(6):1926–1933, 2007.

[2] J. De Beenhouwer, S. Staelens, S. Vandenberghe, and I. Lemahieu. *Acceleration of GATE SPECT Simulations*. Med. Phys., 35(4):1476–1485, 2008.

[3] J. De Beenhouwer, S. Staelens, S. Vandenberghe, J. Verhaeghe, and I. Lemahieu. *Process level discrimination for GATE: assessment of contamination in SPECT and spurious activity in PET*. Med. Phys., submitted, 2008.

[4] S. Staelens, K. Vunckx, J. De Beenhouwer, F. Beekman, Y D'Asseler, J. Nuyts, and I. Lemhieu. *GATE simulations for optimization of pinhole imaging*. Nucl. Instr. and Meth. A, A569:359–363, 2006.

[5] S. Staelens, J. De Beenhouwer, D. Kruecker, L. Maigne, F. Rannou, L. Ferrer, Y. D'Asseler, I. Buvat, and I. Lemahieu. *GATE: Improving the computational efficiency*. Nucl. Instr. and Meth. A, 569:341–345, 2006.

[6] E. Rault, S. Vandenberghe, R. Van Holen, J. De Beenhouwer, S. Staelens, and I. Lemahieu. *Comparison of image quality of different iodine isotopes (I-123, I-124 and I-131)*. Cancer Biotherapy and Radiopharmaceuticals, 22(3):423–430, 2007.

[7] E. Rault, S. Vandenberghe, S. Staelens, J. De Beenhouwer, and I. Lemahieu. *Fast Simulation of Yttrium-90 Bremsstrahlung with GATE*. Nucl. Instr. and Meth. A, submitted, 2008.

[8] R. Van Holen, S. Vandenberghe, S. Staelens, J. De Beenhouwer, E. Rault, and I. Lemahieu. *Reduced Influence of High Energy Contamination in 123I Imaging by Using Rotating Slat Collimators*. Nucl. Instr. and Meth. A, submitted, 2008.

[9] R. Van Holen, S. Vandenberghe, S. Staelens, J. De Beenhouwer, and I. Lemahieu. *Fast 3D iterative image reconstruction for SPECT with rotating slat collimators*. Phys. Med. Biol., accepted for publication, 2008.

Conference Proceedings

[1] J. De Beenhouwer, D. Kruecker, S. Staelens, L. Ferrer, A.F. Chatziioannou, and F.R. Rannou. *Distributed Computing Platform for PET and SPECT Simulations with GATE*. In Proceedings of the IEEE Nuclear Science Symposium and Medical Imaging Conference, volume 94, pages 2437–2440, 2005.

[2] J. De Beenhouwer, S. Staelens, M. Dressel, Y. D'Asseler, S. Vandenberghe, I. Lemahieu, and R. Van de Walle. *Geometrical Importance Sampling and Pulse Height Tallies in GATE*. Proceedings of the 26th Annual International Conference of the IEEE EMBS, 26:1349–1352, 2004.

[3] E. Rault, S. Vandenberghe, S. Staelens, J. De Beenhouwer, and I. Lemahieu. *Investigation of Yttrium-90 Bremsstrahlung with GATE*. In Journal of Nuclear Medicine, abstract book of the 54th SNM annual meeting, volume 48, page 422, 2007.

[4] E. Rault, S. Vandenberghe, R. Van Holen, J. De Beenhouwer, and I. Lemahieu. *Comparison of Quantification and Image Degrading Factors with Different Iodine Isotopes*. In European Journal of Nuclear Medicine, abstracts of the annual congress of the EANM, volume 33, page 374, 2006.

[5] E. Rault, S. Vandenberghe, S. Staelens, J. De Beenhouwer, and I. Lemahieu. *Investigation of Yttrium-90 Bremsstrahlung with GATE*. In Abstract book of the Workshop Quantitative Imaging and Dosimetry in Nuclear Medicine, page 25, 2007.

[6] R. Van Holen, S. Vandenberghe, E. Rault, J. De Beenhouwer, and I. Lemahieu. *Reducing High Energy Contamination in SPECT Using a Rotating Slat Collimator*. In Book of Abstracts 4th International Conference on Imaging Technologies in Biomedical Sciences, page 65, 2007.

[7] E. Rault, S. Vandenberghe, R. Van Holen, J. De Beenhouwer, S. Staelens, and I. Lemahieu. *Comparison of Quantification*

and Image Quality for Different Iodine isotopes (I-123, I-124 and I-131). In Abstract book of the 22nd Annual BHPA Symposium on Physics in Medicine, page 20, 2007.

[8] J. De Beenhouwer, S. Staelens, Y. D'Asseler, and I. Lemahieu. *Optimizing the Scalability of Parallelized GATE Simulations*. In Proceedings of the IEEE Nuclear Science Symposium and Medical Imaging Conference, pages 3904–3908, 2006.

[9] J. De Beenhouwer, R. Van Holen, S. Vandenberghe, S. Staelens, Y. D'Asseler, and I. Lemahieu. *Graphics hardware accelerated reconstruction of SPECT with a slat collimated strip detector*. In Proceedings of the 2006 International Conference on Image Processing, Computer Vision and Pattern Recognition, page 7 pp, 2006.

[10] S. De Clercq, S. Staelens, J. De Beenhouwer, Y. D'Asseler, and I. Lemahieu. *LROC assessment of non-linear filtering methods in Ga-67 SPECT imaging*. In Proceedings of SPIE/Medical Imaging, pages 6146–6160, 2006.

[11] J. De Beenhouwer, S. Staelens, A Goedicke, B. Schweizer, Y. D'Asseler, and I. Lemahieu. *Validation of importance sampling in GATE*. In Proceedings of the 3rd European Medical and Biological Engineering Conference, page 5 pp, 2005.

[12] A Goedicke, B Schweizer, S. Staelens, and J. De Beenhouwer. *Fast simulation of realistic SPECT projections using forced detection in GEANT4*. In Proceedings of the 3rd European Medical and Biological Engineering Conference, page 6 pp, 2005.

[13] W. Deckers, J. Verhaeghe, J. De Beenhouwer, Y. D'Asseler, and I. Lemahieu. *Head-motion compensation during reconstruction from list-mode PET data*. In Proceedings of the 3rd European Medical and Biological Engineering Conference, page 4 pp, 2005.

[14] S. Staelens, J. De Beenhouwer, D. Kruecker, Y. D'Asseler, L. Maigne, F. Rannou, L. Ferrer, I. Buvat, and I. Lemahieu. *GATE : improving the computational efficiency*. In Book of abstracts 3rd International Conference on Imaging Technologies in Biomedical Sciences, page 17, 2005.

[15] J. De Beenhouwer, S. Staelens, S. Vandenberghe, Y. D'Asseler, and I. Lemahieu. *Graphics hardware accelerated optimal rotator for iterative reconstruction*. In Proceedings of the 2005 International Conference on Mathematics and Engineering Techniques in Medicine and Biological Sciences, pages 219–224, 2005.

[16] S. Staelens, S. Vandenberghe, J. De Beenhouwer, S. De Clercq, Y. D'Asseler, I. Lemahieu, and R. Van de Walle. *Simulation study comparing the imaging performance of a solid date detector with a rotating slat collimator versus parallel beam collimator setups*. In Proceedings of SPIE Medical Imaging 2004, volume 5372, pages 301–310, 2004.

[17] J. De Beenhouwer, S. Staelens, S. Vandenberghe, and I. Lemahieu. *Acceleration of GATE SPECT Simulations*. In Proceedings of the IEEE Nuclear Science Symposium and Medical Imaging Conference, pages 3649–3655, 2007.

[18] E. Rault, S. Vandenberghe, R. Van Holen, J. De Beenhouwer, and I. Lemahieu. *Comparison of quantification and image degrading factors for different iodine isotopes (I-123, I-124 and I-131)*. In Abstracts 7e UGent-FirW doctoraatssymposium, page 38, 2006.

[19] J. De Beenhouwer, S. Staelens, and Y. D'Asseler. *Rotational Importance Maps in GATE*. In Abstracts of the 5th FTW PhD symposium, page 2 pp, 2004.

[20] J. De Beenhouwer and S. Staelens. *Variance reduction techniques for Monte Carlo simulations in nuclear medical imaging*. In Abstracts of the 4th FTW PhD symposium, 2003.

[21] S. Staelens, J. De Beenhouwer, M. Koole, Y. D'Asseler, S. Vandenberghe, I. Lemahieu, and R. Van de Walle. *Monte Carlo simulations for SPECT: object oriented flexibility benchmark*. In Book of abstracts of the 3rd Belgian Day on Biomedical Engineering, page 30, 2003.