FACULTY OF ENGINEERING
AND ARCHITECTURE

Timing Capabilities and Image Reconstruction for the Walk-Through PET: A Large Flat-Panel PET System with Monolithic Detectors

Jens Maebe

Doctoral dissertation submitted to obtain the academic degree of
Doctor of Biomedical Engineering

Supervisors
Prof. Stefaan Vandenberghe, PhD - Prof. Christian Vanhove, PhD
Department of Electronics and Information Systems
Faculty of Engineering and Architecture, Ghent University

September 2024

GHENT
UNIVERSITY

ISBN 978-94-6355-873-0
NUR 954
Wettelijk depot: D/2024/10.500/78

Members of the Examination Board

Chair

Prof. Hennie De Schepper, PhD, Ghent University

Other members entitled to vote

Prof. Jan Aelterman, PhD, Ghent University
Prof. Yves D'Asseler, PhD, Ghent University
Prof. Bjorn De Sutter, PhD, Ghent University
Stefan Gundacker, PhD, RWTH Aachen University, Germany
Prof. Suleman Surti, PhD, University of Pennsylvania, USA

Supervisors

Prof. Stefaan Vandenberghe, PhD, Ghent University
Prof. Christian Vanhove, PhD, Ghent University

Acknowledgements

Doing my PhD has been a long, but fun and interesting, journey. I've learned a great deal of things, and met many wonderful people. While it was certainly not without its ups and downs, I can confidently say that I am happy to have done it. But now at the end, comes the most difficult part of all: writing the acknowledgements.

First, I would like to thank my promotors Stefaan Vandenberghe and Christian Vanhove. **Stefaan**, thank you for giving me the opportunity to do research in a topic I have really grown to enjoy. You gave me the necessary guidance, while still providing me with plenty of freedom in terms of both the overall direction of my research, and how to tackle specific issues at hand. I'm sure this played a large part in my enjoyment of the PhD. Outside of work, you made our group feel like friends, organizing many activities for us. I learned a lot from you, not only about PET, but also about all sorts of nice places and restaurants in Belgium! **Chris**, although you were less directly involved in my research, you always had time to help with questions or practical issues, even before you officially became my promotor. Thank you for helping with all the lab measurements, whether ours or our students, throughout the years. Your relaxed vibe made it very easy and pleasant to work with you! I would also like to thank **Roel** Van Holen, who was initially one of my co-promotors, but unfortunately had to stop when he left UGent. It was great to have you there, even after leaving UGent, to fill in these specific knowledge gaps. I've really enjoyed the few conferences we were at together!

Second, I would like to thank **Hennie** De Schepper for taking the time to chair the PhD jury. Also, a thanks to the members of the jury, **Suleman** Surti, **Stefan** Gundacker, **Yves** D'Asseler, **Bjorn** De Sutter and **Jan** Aelterman, for reading my dissertation and providing valuable feedback. A special thanks to Yves for assisting with all the scanner experiments we did in the past.

Saskia, you have been a great secretary, making all of our lives easier! I've really enjoyed our lunchtime talks. **Jurgen**, thank you for providing all the necessary practical support. It was nice knowing that when we had an issue, we could hop by and probably get a fix within 15 minutes. Thank you to the professors from IBiTech, **Pieter**, **An**, **Patrick** and **Charlotte**, for creating such a great research group and work environment!

Furthermore, there are some people who, while not directly part of our research group, have been a great help throughout these last few years. **Karel**, your electronics expertise has been invaluable, especially during the later parts of my PhD, when we started doing experimental detector work. Thank you for taking the time for us. **Nadia**, I am really glad Stefaan found you! Your physician's point of view was always helpful, and I am glad you could often join us outside of work, so that we got to know each other better. **Joel**, your extensive expertise in PET, and ability to look at problems from a different angle, has provided me (and the rest of the group) with invaluable feedback.

Florence, our (slightly competitive) academic genius! More so than a great colleague, you've been a great friend, and I'm really happy for the memories we've made. I've had a great time working, joking, traveling, and spending time with you. You were always checking up on everyone. Sometimes, I don't understand where you get the energy for it, but I am really grateful that you do! **Maya**, I still remember when you were that colleague from Qatar who I only knew from online meetings. I'm happy you were able to come to Belgium, because otherwise I would have never gotten to know you the way I do now. Hopefully one day, we'll get the chance to go to Lebanon as well! **Rabia**, I haven't known you for as long as I do the others, but you've been a great addition to the team, and I hope we'll get to know each other further in the future. Over chocolates, as you say! **Meysam**, we've probably worked together the most of all, and although we could surely annoy each other from time to time because of it (sorry), I can say I've had a great time with you around. It was nice to have someone to confide in. **Boris**, thank you for being another relaxed and easy-going person in our group. You've been a great colleague, travel companion and friend throughout!

Melissa, you've been a great friend, and a great source of snacks and drinks! If only the rats didn't take up so much of your time, we could spend some more time with you. **Jolan**, you were great at bringing people together, and/or luring them to the kitchen for a break. **Emma**,

I hope that one day, you too will learn the joy of drinking hot chocolate in the mornings. **Gert**, although I often only saw you once every month or so, it was nice to regularly catch up with each other. **Zoé**, writing these acknowledgements have made me realize I should talk to you more often! **Amith**, you're a kind, warm, nurturing, beautiful, and probably the most fun... Rather, not boring and have a funny accent, allegedly. In all seriousness though, you've been a great friend throughout the years, who was always (*only applicable past 10 AM) there for me, either to have some good old fun, or to talk about more serious topics. I'm grateful you always took the time to pass by our office! **Tim**, you're one of my top three colleagues from BioMMedA! Also, thank you for all your assistance on the administrative aspects at the end of my PhD. **Sarah**, next time we go kayaking, we'll beat Maya! Maybe in Sweden, she won't be acclimatized to the weather there. **Wouter**, I'm amazed how you ever got any work done, given that you were always either sporting, eating or digesting, but thank you for making the restaurant trips more enjoyable! To all the other colleagues in **BioMMedA**, thank you for all the good memories! It wouldn't have been the same with just MEDISIP.

There are also a number of past colleagues I would like to thank. **Mariele**, thank you for keeping me company during those times when it was basically just the two of us in the office! I'm glad we got along so well. **Paulo**, you're quite possibly the most energetic person I've ever met! Our time together in the lab was relatively short, but very memorable. I'm happy you came back to Belgium, so that we still get to see each other. **Jens** (Mincke), thank you for being a cheerful personality in our office, and showing us all those unconventional purposes you can use PET for. **Ashkan**, I'm glad that due to some administrative shenanigans, it was decided that the best office place for a veterinary graduate working on a jiu-jitsu glove, was at MEDISIP. You've been a great friend and unofficial colleague! **Charlotte**, although you were often sitting elsewhere, whenever you would pass by our office, you would bring all of your positivity with you. **Milan**, you were a great colleague, and I'll miss running into you in Infinity lab.

A special thank you to my parents, **Pascal** and **Paula**, for supporting me throughout my PhD. Also a thank you to my **friends** and **family** who, while they had no direct involvement in my PhD, have been a great support throughout my life, and have shaped who I am today.

Finally, a thank you to the Research Foundation Flanders (FWO) for funding this work (file number 79876).

Summary

This dissertation is focused on enhancing the time resolution in monolithic detectors for positron emission tomography (PET), and the development of an image reconstruction framework for a new, flat-panel, long axial field-of-view (AFOV) PET system based on such detectors. The studies performed in this dissertation are based on simulation, using the GATE (GEANT4 Application for Tomographic Emission) Monte Carlo software. PET is a functional imaging technique, where a tracer, labelled with a positron emitting radioactive isotope, is administered to the patient to study metabolic processes in the body. PET is, for example, commonly used in oncology, where fluorodeoxyglucose (FDG) is used as a tracer. It accumulates in cells with high energy consumption, such as cancerous cells, so that visualization of the tracer distribution allows for detection of malignant tumors.

In Chapter 2, we start by giving some background information regarding the physical principles behind PET. We delve deeper into PET detector technology, and the image reconstruction process for PET, as these are the main focuses of this dissertation. Potential sources of image degradation (resolution limitations or noise production) during the PET acquisition process are summarized, and we describe how these should be handled in the image reconstruction. Monte Carlo simulation, and the GATE software, are also introduced as a tool for designing and evaluating new systems and algorithms. In Chapter 3, we take a look at two recent evolution in PET: long AFOV PET for improved image quality by making systems more sensitive, and the use of deep learning in PET. Since deep learning has a wide range of applications within PET, we again limit our discussion to those aspects relevant for this dissertation: deep learning for PET detector signal processing, and deep learning for image reconstruction and image post-processing.

Timing capabilities of monolithic detectors

As the isotope decays, pairs of back-to-back gamma photons are produced and then captured by the PET detectors, from which the tracer distribution can be reconstructed. In this dissertation, we focus on newly developed monolithic detector technology. Monolithic detectors consist of a single monolithic crystal coupled to an array of multiple silicon photomultipliers (SiPMs), and offer high spatial resolution in addition to depth-of-interaction (DOI) decoding capabilities. They offer superior system resolutions compared to current clinical PET systems, which are based on pixelated detectors (where the crystal is subdivided into smaller pixels).

For the scintillation material, there are two main choices for PET detectors: L(Y)SO (lutetium(-yttrium) oxyorthosilicate) or BGO (bismuth germanate). LYSO has superior scintillation characteristics, with a faster scintillation pulse and more optical photons generated per gamma interaction, which generally translates to better spatial and time resolutions for the detector. Time-of-flight (TOF) information about the difference in detection time of the two gamma photons can be incorporated in the image reconstruction, better localizing where along the line-of-response (LOR) the annihilation happened, thereby improving the image signal-to-noise ratio. Therefore, the detector time resolution is an important parameter for scanner performance, making L(Y)SO is the material of choice for TOF-PET. BGO, on the other hand, is considerably cheaper and despite its slow scintillation characteristics, does offer the potential for improved time resolution through near-instantaneous generation of Cerenkov photons. These Cerenkov photons are, however, very few ($\sim$ 18 photons generated per 511 keV gamma absorption), and are therefore difficult to detect.

In Chapter 4, we perform a simulation study to better understand and optimize the time resolution in monolithic detectors. We investigate the effect of detector geometry (monolithic/pixelated), scintillation material (LYSO/BGO) and crystal surface finishes on the photon collection efficiency (number of optical photons detected per gamma event) and predicted time resolutions. Both scintillation and Cerenkov photon production are considered, and we investigate the potential of Cerenkov-based time estimation in monolithic BGO. As expected, we observe that reflective side surfaces increase the photon collection efficiency. More interestingly, the surface finish on the side of the SiPMs also has a large impact on the photon collection efficiency, with a rough surface finish

providing better collection efficiency than a polished surface finish. We note that for monolithic BGO, the increase in collection efficiency does translate to improved time resolution (gains of 10 to 15%), whereas for monolithic LYSO, we note little difference. We also conclude that it is very difficult to make use of the Cerenkov photons to further improve time resolution in monolithic BGO detectors. This is due to the spread of light in monolithic crystals, resulting in almost no SiPMs detecting more than a single Cerenkov photon per SiPM. Consistently measuring these Cerenkov photons therefore requires triggering below the single photoelectron level, which is difficult to do due to SiPM dark counts (thermal, rather than optical, excitations).

In Chapter 5, we seek to further improve time resolution in monolithic (LYSO) detectors using deep learning based time estimation. Conventionally, the gamma interaction time in a monolithic detector can be estimated by leading edge discrimination on the SiPM signals, providing a single timestamp per SiPM. The first few recorded timestamps can then be averaged to obtain an estimated gamma interaction time. In doing so, we however lose both valuable information stored in the SiPM signals, and any spatial information regarding the distribution of timestamps. Deep learning time prediction based on convolutional neural networks (CNNs) can, however, retain and make use of this information for better time resolution. We investigate two options: (i) using the matrix of SiPM timestamps, obtained from leading edge thresholding, as an input to a 2D CNN for gamma time prediction, and (ii) using the matrix of the leading edge portion of digitized SiPM signals as an input to a 3D CNN for gamma time prediction. The 2D CNN therefore incorporates spatial information about the SiPM timestamps, and the 3D CNN further includes additional temporal information. Based on our simulations, the 2D CNN improved time resolution by 17%, from a TOF resolution of 177 ps full width half maximum (FWHM) for timestamp averaging to 151 ps FWHM for the 2D CNN. The 3D CNN showed an improvement of 26%, with a TOF resolution of 141 ps FWHM. In addition, the neural network methods showed better spatial uniformity for the time resolution, with less variance based on the gamma interaction position. Although the 3D CNN performed best, obtaining digitized SiPM signals requires more expensive electronics, making the 2D CNN a more feasible approach in practice.

Image reconstruction for the Walk-Through PET

In the second part of the dissertation, we focus on developing an image reconstruction framework for the Walk-Through PET (WT-PET): a flat-panel, long (106 cm) AFOV PET system based on monolithic detector technology. The WT-PET is being developed at our research group, with the aim to provide a high-resolution, high-throughput and cost-effective alternative to existing long AFOV PET systems. Long AFOV PET systems offer increased sensitivity (probability of detecting the gamma photons), which translates to better image quality and therefore better diagnostic capabilities. These systems are however very expensive due to the vast number of detectors required. The WT-PET reduces this cost by using a dual-panel, rather than cylindrical, geometry. This allows detectors to be placed closer to the patient, thereby reducing the number of detectors required to achieve a certain sensitivity. The system achieves high spatial resolution through the use of monolithic detectors, and patients are scanned in an upright position, to increase throughput by reducing time spent on positioning the patient on the bed.

The unique configuration of the WT-PET, in terms of the flat-panel geometry and use of monolithic detectors, requires custom image reconstruction software. In Chapter 7, we develop and introduce our iterative image reconstruction framework. We evaluate the system spatial resolution based on point source reconstructions, and obtain a spatial resolution of below 2 mm FWHM across most of the FOV. Reconstructions are further evaluated using the image quality (IQ) phantom, and the digitized humanoid XCAT phantom. We observe that good image quality can be obtained in short, 30 second acquisitions, although we do observe some limited angle artifacts due to the flat-panel geometry of the system.

During a PET acquisition, the gamma photons may scatter within the patient, resulting in mispositioned LORs that manifest as additive noise in the reconstruction. In Chapter 7, we have artificially removed these scattered coincidences based on the ground-truth labels obtained from the Monte Carlo simulation, but in practice, these scatters should be estimated and corrected for during the image reconstruction procedure. Conventionally, the single scatter simulation (SSS) algorithm is used for this, which estimates scatter contributions based on a combination of the activity distribution and attenuation coefficient map, the physical models of photon scattering, and the scanner geometry and detector parameters. One of the drawbacks of SSS is the need for an attenuation map, which is usually derived from an accompanying computed tomography (CT) scan.

However, for dose reduction purposes it can be beneficial to skip the CT, in which case SSS becomes unsuitable. An alternative to SSS is energy based scatter correction, where scatter contributions are estimated from the dual energy distribution of scattered and unscattered coincidences, and a model of the detector energy response. The attenuation map is no longer required, providing a CT-less scatter correction algorithm. We implement this algorithm within our reconstruction framework in Chapter 8, and again evaluate on the IQ and XCAT phantoms. We observe that we can accurately correct for scatter with energy-based scatter correction, showing minimal image degradation compared to the ideal case of artificially removing scattered coincidences from the simulation data.

The attenuation map is also required for attenuation correction in PET: as gamma photons are attenuated (absorbed or scattered) in the body, fewer coincidences are detected from annihilations deeper within the patient. This effect needs to be corrected for during the image reconstruction, to obtain an accurate estimation of the tracer distribution. In Chapter 7, this was done using the ground-truth attenuation map from the simulation, which as mentioned, would be derived from a CT in practice. Therefore, if we wish to skip the CT for dose reduction purposes, we also require a CT-less attenuation correction method. We investigate this in Chapter 9, where we study two specific methods: (i) the use of external gamma transmission sources to derive an attenuation map, and (ii) simultaneous estimation of the activity and attenuation maps from the TOF emission data itself. Visually, differences between reconstructions using the ground-truth attenuation map and the ones based on CT-less attenuation correction methods are minimal. Quantitatively, we do observe some systematic differences for the simultaneous estimation of activity and attenuation, with an overestimation of activity in the lung and arms, and an underestimation of activity in the abdominal region. We expect that inclusion of a regularization term in the reconstruction to penalize unrealistic attenuation distributions could improve the results, but this was not explored in this work.

Finally, in Chapter 10 we implement a deep learning based approach to correct for the limited angle artifacts arising from the flat-panel geometry of the system. We do this by training a U-Net (a CNN architecture commonly used for image-to-image translation) to predict artifact-free images from the original reconstructions containing limited angle artifacts. Rather than implementing the neural network as a post-processing step,

we incorporate the network within the iterative reconstruction procedure. This makes the process more robust by incorporating the measured LORs at every iteration, ensuring data consistency. Training and testing is done using the humanoid XCAT phantoms, showing good results. Incorporating the network in the reconstruction procedure provided sharper images with better contrast recovery at the cost of increased noise, compared to the implementation as a post-processing step. We do note that further testing (e.g., on experimental data) would be required for better evaluation of the algorithm, as the simulated training data may be too simplistic in nature, and not entirely representative of experiment.

Conclusion

The WT-PET provides a promising alternative to other long AFOV PET systems, with a number of advantages enabled by the use of monolithic detector technology and the flat panel design. The unique configuration, however, also has some limitations, which we have aimed to address within this dissertation. We improved the time resolution of the monolithic detectors, developed CT-less image reconstruction methods for the WT-PET, and provided a framework for limited angle artifact correction in the system.

Samenvatting

Dit proefschrift is gericht op het verbeteren van de tijdresolutie in monolithische detectoren voor positronemissietomografie (PET), en de ontwikkeling van een raamwerk voor beeldreconstructie voor een nieuw, vlakpaneel, lang axiaal gezichtsveld (axial field-of-view, AFOV) PET-systeem gebaseerd op dergelijke detectoren. De studies in dit proefschrift zijn gebaseerd op simulatie met behulp van de GATE (GEANT4 Application for Tomographic Emission) Monte Carlo software. PET is een functionele beeldvormingstechniek waarbij een tracer, gelabeld met een positronemitterend radioactief isotoop, aan de patiënt wordt toegediend om stofwisselingsprocessen in het lichaam te bestuderen. PET wordt bijvoorbeeld veel gebruikt in de oncologie, waar fluorodeoxyglucose (FDG) als tracer wordt gebruikt. Het stapelt zich op in cellen met een hoog energieverbruik, zoals kankercellen, zodat visualisatie van de verdeling van de tracer het mogelijk maakt om kwaadaardige tumoren op te sporen.

In hoofdstuk 2 geven we eerst wat achtergrondinformatie over de fysische principes achter PET. We gaan dieper in op PET-detectortechnologie en het beeldreconstructieproces voor PET, aangezien dit de belangrijkste aandachtspunten zijn in dit proefschrift. Potentiële bronnen van beelddegradatie (resolutiebeperkingen of ruisproductie) tijdens het PET-acquisitieproces worden samengevat, en we beschrijven hoe deze moeten worden behandeld in de beeldreconstructie. Monte Carlo-simulatie en de GATE-software worden ook geïntroduceerd als hulpmiddel voor het ontwerpen en evalueren van nieuwe systemen en algoritmen. In hoofdstuk 3 kijken we naar twee recente ontwikkelingen in PET: lange AFOV PET voor verbeterde beeldkwaliteit door systemen gevoeliger te maken, en het gebruik van deep learning in PET. Omdat deep learning een breed scala aan toepassingen heeft binnen PET, beperken we onze discussie weer tot de aspecten die relevant zijn voor dit proefschrift: deep learning voor PET-detectorsignaalverwerking, en deep learning voor beeldreconstructie en beeldnabewerking.

Timingmogelijkheden van monolithische detectoren

Terwijl de isotoop vervalt, worden paren van gammafotonen in tegenovergestelde richting geproduceerd en vervolgens opgevangen door de PET-detectoren, waaruit de verdeling van de tracer kan worden gereconstrueerd. In dit proefschrift richten we ons op nieuw ontwikkelde monolithische detectortechnologie. Monolithische detectoren bestaan uit een enkel monolithisch kristal gekoppeld aan een array van meerdere silicon photomultipliers (SiPMs), en bieden een hoge ruimtelijke resolutie naast diepte-van-interactie (DOI) decoderingsmogelijkheden. Ze bieden superieure systeemresoluties vergeleken met de huidige klinische PET-systemen, die gebaseerd zijn op gepixelde detectoren (waarbij het kristal is onderverdeeld in kleinere pixels).

Voor het scintillatiemateriaal zijn er voornamelijk twee keuzes voor PET-detectoren: L(Y)SO (lutetium(-yttrium) oxyorthosilicaat) of BGO (bismutgermanaat). LYSO heeft superieure scintillatiekarakteristieken, met een snellere scintillatiepuls en meer optische fotonen gegenereerd per gamma-interactie, wat zich over het algemeen vertaalt in betere ruimtelijke en tijdsresoluties voor de detector. Time-of-flight (TOF) informatie over het verschil in detectietijd van de twee gammafotonen kan worden opgenomen in de beeldreconstructie, waardoor beter gelokaliseerd kan worden waar langs de line-of-response (LOR) de annihilatie plaatsvond, waardoor de signaal-ruisverhouding van het beeld verbetert. Daarom is de detectortijdresolutie een belangrijke parameter voor de prestaties van de scanner, waardoor L(Y)SO het materiaal bij uitstek is voor TOF-PET. BGO daarentegen is aanzienlijk goedkoper en biedt, ondanks de trage scintillatiekarakteristieken, het potentieel voor een verbeterde tijdresolutie door de bijna onmiddellijke generatie van Cerenkovfotonen. Deze Cerenkovfotonen zijn echter erg klein in aantal (ongeveer 18 fotonen per 511 keV gamma-absorptie) en zijn daarom moeilijk te detecteren.

In hoofdstuk 4 voeren we een simulatiestudie uit om de tijdresolutie in monolithische detectoren beter te begrijpen en te optimaliseren. We onderzoeken het effect van de geometrie van de detector (monolithisch/gepixeld), het scintillatiemateriaal (LYSO/BGO) en de afwerking van het kristaloppervlak op de fotonverzamelingsefficiëntie (aantal optische fotonen gedetecteerd per gamma-interactie) en de voorspelde tijdsresoluties. Zowel scintillatie als Cerenkov fotonproductie worden beschouwd en we onderzoeken het potentieel van Cerenkov-gebaseerde tijdschatting in monolithisch BGO. Zoals verwacht zien we dat reflecterende zijoppervlakken de fotonverzamelingsefficiëntie verhogen. Inte-

ressanter is dat de oppervlakteafwerking aan de zijkant van de SiPMs ook een grote invloed heeft op de fotonverzamelingsefficiëntie, waarbij een ruw oppervlak een betere verzamelingsefficiëntie oplevert dan een gepolijst oppervlak. We merken op dat voor monolithisch BGO, de toename in verzamelingsefficiëntie vertaalt naar een betere tijdresolutie (winsten van 10 tot 15%), terwijl we voor monolithisch LYSO weinig verschil zien. We concluderen ook dat het erg moeilijk is om gebruik te maken van de Cerenkov fotonen om de tijdresolutie verder te verbeteren in monolithische BGO-detectoren. Dit komt door de verspreiding van licht in monolithische kristallen, waardoor bijna geen enkele SiPM meer dan één Cerenkov foton per SiPM detecteert. Het consistent meten van deze Cerenkov-fotonen vereist daarom triggeren onder het niveau van één foto-elektron, wat moeilijk is vanwege de SiPM dark counts (thermische in plaats van optische excitaties).

In hoofdstuk 5 proberen we de tijdresolutie in monolithische (LYSO) detectoren verder te verbeteren door gebruik te maken van deep learning gebaseerd op tijdschatting. Conventioneel kan de gamma-interactietijd in een monolithische detector worden geschat door leading edge discriminatie op de SiPM-signalen, wat een enkele tijdstempel per SiPM oplevert. De eerste paar opgenomen tijdstempels kunnen dan worden gemiddeld om een geschatte gamma-interactietijd te verkrijgen. Hierbij verliezen we echter zowel waardevolle informatie die is opgeslagen in de SiPM-signalen, en alle ruimtelijke informatie over de verdeling van de tijdstempels. Deep learning tijdvoorspelling op basis van convolutionele neurale netwerken (CNN's) kan deze informatie echter behouden en gebruiken voor een betere tijdsresolutie. We onderzoeken twee opties: (i) gebruik van de matrix van SiPM-tijdstempels, verkregen door leading edge discriminatie, als invoer voor een 2D CNN voor gamma-tijdvoorspelling, en (ii) gebruik van de matrix van het leading edge gedeelte van gedigitaliseerde SiPM-signalen als invoer voor een 3D CNN voor gamma-tijdvoorspelling. De 2D CNN bevat dus ruimtelijke informatie over de SiPM-tijdstempels en de 3D CNN bevat extra temporele informatie. Op basis van onze simulaties verbeterde de 2D CNN de tijdresolutie met 17%, van een TOF-resolutie van 177 ps full width at half maximum (FWHM) voor tijdstempelgemiddelden tot 151 ps FWHM voor de 2D CNN. De 3D CNN liet een verbetering zien van 26%, met een TOF-resolutie van 141 ps FWHM. Bovendien lieten de neurale netwerkmethoden een betere ruimtelijke uniformiteit zien voor de tijdresolutie, met minder variantie op basis van de positie van de gamma-interactie. Hoewel de 3D CNN

het beste presteerde, vereist het verkrijgen van gedigitaliseerde SiPM-signalen duurdere elektronica, waardoor de 2D CNN in de praktijk een haalbaardere aanpak is.

Beeldreconstructie voor de Walk-Through PET

In het tweede deel van het proefschrift richten we ons op het ontwikkelen van een beeldreconstructie raamwerk voor de Walk-Through PET (WT-PET): een vlakpaneel, lang (106 cm) AFOV PET-systeem gebaseerd op monolithische detector technologie. De WT-PET wordt in onze onderzoeksgroep ontwikkeld met als doel een hoge resolutie, hoge doorvoer en kosteneffectief alternatief te bieden voor bestaande lange AFOV PET-systemen. PET-systemen met lange AFOV bieden een hogere sensitiviteit (waarschijnlijkheid van het detecteren van de gammafotonen), wat zich vertaalt in een betere beeldkwaliteit en dus betere diagnostische mogelijkheden. Deze systemen zijn echter zeer duur vanwege het grote aantal detectoren dat nodig is. De WT-PET verlaagt deze kosten door een geometrie met twee panelen te gebruiken in plaats van een cilindrische geometrie. Hierdoor kunnen detectoren dichter bij de patiënt worden geplaatst, waardoor er minder detectoren nodig zijn om een bepaalde sensitiviteit te bereiken. Het systeem bereikt een hoge ruimtelijke resolutie door het gebruik van monolithische detectoren en patiënten worden rechtop gescand om de doorvoer te verhogen door de tijd te verminderen die nodig is om de patiënt op het bed te positioneren.

De unieke configuratie van de WT-PET, met betrekking tot de vlakke paneelgeometrie en het gebruik van monolithische detectoren, vereist aangepaste beeldreconstructiesoftware. In hoofdstuk 7 ontwikkelen en introduceren we ons iteratieve beeldreconstructieraamwerk. We evalueren de ruimtelijke resolutie van het systeem op basis van puntbronreconstructies en verkrijgen een ruimtelijke resolutie van minder dan 2 mm FWHM over het grootste deel van de FOV. De reconstructies worden verder geëvalueerd met behulp van het fantoom voor beeldkwaliteit (image quality, IQ) en het gedigitaliseerde humanoïde XCAT-fantoom. We stellen vast dat een goede beeldkwaliteit kan worden verkregen in korte, 30 seconden durende acquisities, hoewel we enkele beperkte projectiehoek artefacten waarnemen als gevolg van de vlakke paneelgeometrie van het systeem.

Tijdens een PET-acquisitie kunnen de gammafotonen in de patiënt verstrooid raken, wat leidt tot verkeerd gepositioneerde LOR's die zich manifesteren als additieve ruis in de reconstructie. In hoofdstuk 7 hebben

we deze verstrooide coincidenties kunstmatig verwijderd op basis van de ground-truth labels verkregen uit de Monte Carlo simulatie, maar in de praktijk moeten deze verstrooiingen worden geschat en gecorrigeerd tijdens de beeldreconstructieprocedure. Conventioneel wordt hiervoor het SSS-algoritme (Single Scatter Simulation) gebruikt, dat de verstrooiingsbijdragen schat op basis van een combinatie van de activiteitsverdeling en de attenuatiemap, de fysische modellen van fotonverstrooiing en de scannergeometrie en detectorparameters. Een van de nadelen van SSS is de noodzaak van een attenuatiemap, die meestal wordt afgeleid van een bijbehorende computer tomografie (CT) scan. Voor dosisreductiedoeleinden kan het echter gunstig zijn om de CT over te slaan, in welk geval SSS ongeschikt wordt. Een alternatief voor SSS is een energiegebaseerde verstrooiingscorrectie, waarbij de verstrooiingsbijdragen worden geschat op basis van de dubbele energieverdeling van verstrooide en niet-verstrooide coincidenties, en een model van de energierespons van de detector. De attenuatiemap is niet langer nodig, wat ons een CT-loos verstrooiingscorrectiealgoritme geeft. We implementeren dit algoritme in ons reconstructieraamwerk in hoofdstuk 8 en evalueren het opnieuw op de IQ en XCAT-fantomen. We zien dat we nauwkeurig kunnen corrigeren voor verstrooiing met energiegebaseerde verstrooiingscorrectie, met minimale beelddegradatie in vergelijking met het ideale geval van het kunstmatig verwijderen van verstrooide coincidenties uit de simulatiedata.

De attenuatiemap is ook nodig voor attenuatiecorrectie bij PET: als gammafotonen in het lichaam worden verzwakt (geabsorbeerd of verstrooid), zullen er minder samenvallende annihilaties dieper in de patiënt worden gedetecteerd. Dit effect moet worden gecorrigeerd tijdens de beeldreconstructie om een nauwkeurige schatting van de tracerdistributie te verkrijgen. In hoofdstuk 7 werd dit gedaan met behulp van de ground-truth attenuatiemap gebruikt voor de simulatie, die zoals vermeld in de praktijk van een CT zou worden afgeleid. Als we de CT willen overslaan voor dosisvermindering, hebben we dus ook een CT-loze attenuatiecorrectiemethode nodig. We onderzoeken dit in hoofdstuk 9, waar we twee specifieke methoden bestuderen: (i) het gebruik van externe gammatransmissiebronnen om een attenuatiemap af te leiden, en (ii) gelijktijdige schatting van de activiteit en attenuatie uit de TOF-emissiedata zelf. Visueel zijn de verschillen tussen reconstructies op basis van de ground-truth attenuatiemap en reconstructies op basis van CT-loze attenuatiecorrectiemethoden minimaal. Kwantitatief zien we wel enkele systematische verschillen voor de gelijktijdige schatting van

activiteit en attenuatie, met een overschatting van activiteit in de longen en armen, en een onderschatting van activiteit in het abdominale gebied. We verwachten dat het opnemen van een regularisatieterm in de reconstructie om onrealistische attenuatiecoëfficiënten te bestraffen de resultaten zou kunnen verbeteren, maar dit is niet onderzocht in dit werk.

Ten slotte implementeren we in hoofdstuk 10 een deep learning gebaseerde methode om de beperkte projectiehoek artefacten te corrigeren die ontstaan door de vlakke-paneelgeometrie van het systeem. We doen dit door een U-Net (een CNN-architectuur die vaak wordt gebruikt voor beeld-naar-beeld translatie) te trainen om artefactvrije beelden te voorspellen uit de originele reconstructies die beperkte projectiehoek artefacten bevatten. In plaats van het neurale netwerk te implementeren als een post-processing stap, integreren we het netwerk in de iteratieve reconstructieprocedure. Dit maakt het proces robuuster doordat de gemeten LOR's bij elke iteratie worden meegenomen, waardoor de consistentie van de gegevens wordt gewaarborgd. Training en testen worden uitgevoerd met de humanoïde XCAT-fantomen en laten goede resultaten zien. De integratie van het netwerk in de reconstructieprocedure leverde scherpere beelden op met beter contrastherstel ten koste van meer ruis, vergeleken met de implementatie als een post-processing stap. We merken wel op dat verdere tests (bijvoorbeeld op experimentele gegevens) nodig zouden zijn voor een betere evaluatie van het algoritme, aangezien de gesimuleerde data mogelijks te simplistisch van aard zijn en niet geheel representatief voor experimentele data.

Conclusie

De WT-PET biedt een veelbelovend alternatief voor bestaande lange AFOV PET-systemen, met een aantal voordelen door het gebruik van monolithische detectortechnologie en het vlakke paneelontwerp. De unieke configuratie heeft echter ook enkele beperkingen, die we in dit proefschrift hebben aangepakt. We hebben de tijdresolutie van de monolithische detectoren verbeterd, CT-loze beeldreconstructiemethoden ontwikkeld voor de WT-PET en een raamwerk geleverd voor beperkte projectiehoekartefactcorrectie in het systeem.

List of Abbreviations

ADC	analog-to-digital converter
AFOV	axial field of view
AI	artificial intelligence
ASIC	application-specific integrated circuit
BGO	bismuth germanate
BMI	body mass index
BV	background variability
CNN	convolutional neural network
CNR	contrast-to-noise ratio
CPU	central processing unit
CRC	contrast recovery coefficient
CT	computed tomography
CTR	coincidence time resolution
CTW	coincidence time window
CUDA	compute unified device architecture
DL	deep learning
DOI	depth of interaction
EM	expectation-maximization
ESR	enhanced specular reflector
FBP	filtered backprojection
FDG	fluorodeoxyglucose
FOV	field of view
FWHM	full width at half maximum

FWTM	full width at tenth maximum
GAN	generative adversarial network
GATE	GEANT4 Application for Tomographic Emission
GPU	graphics processing unit
IQ	image quality
JIT	just-in-time
LM	list-mode
LOR	line of response
LSO	lutetium oxyorthosilicate
LUT	lookup table
LYSO	lutetium–yttrium oxyorthosilicate
MAP-EM	maximum a posteriori expectation-maximization
MC	Monte Carlo
ML	maximum likelihood / machine learning
MLAA	maximum likelihood activity and attenuation
MLEM	maximum likelihood expectation-maximization
MLTR	maximum likelihood transmission
MRD	maximum ring difference
MRI	magnetic resonance imaging
MRP	median root prior
MSE	mean squared error
NAC	non attenuation corrected
NECR	noise equivalent count rate
NEMA	National Electrical Manufacturers Association
NUV	near ultraviolet
OSEM	ordered subset expectation-maximization
PDE	photon detection efficiency
PET	positron emission tomography
PMT	photomultiplier tube
PSMA	prostate-specific membrane antigen

RC	resistor–capacitor
ReLU	rectified linear unit
ROI	region of interest
SDF	signed distance function
SF	scatter fraction
SiPM	silicon photomultiplier
SNR	signal-to-noise ratio
SPAD	single-photon avalanche diode
SPECT	single-photon emission computed tomography
SPTR	single photon time resolution
SRM	system response matrix
SSRB	single-slice rebinning
SSS	single scatter simulation
TOF	time-of-flight
TV	total variation
WT-PET	Walk-Through PET
XCAT	Extended Cardiac-Torso

Contents

Chapter 1

Introduction

1.1 Context

The topic of this dissertation deals with positron emission tomography (PET), a functional imaging technique used to study metabolic processes in living organisms. In the medical field, PET offers complementary information to structural (anatomical) imaging techniques such as computed tomography (CT) or magnetic resonance imaging (MRI). PET is for example widely used in the field of oncology, for cancer diagnosis and tumor detection. These metabolic processes are studied using a radiotracer, administered to the patient, where its distribution can then be visualized by the PET scanner. This is done using detectors containing a scintillation crystal to absorb gamma photons generated during the PET acquisition process, from which ultimately the tracer distribution can be reconstructed. Current clinical systems make use of pixelated scintillation crystals, where the crystal is subdivided in smaller pixels. More recently, monolithic crystals have shown considerable advancements in terms of spatial resolution, and are therefore being investigated for use in newer systems. Besides spatial resolution, the time resolution of PET detectors is also an important parameter for time-of-flight (TOF) PET, where more accurate timing of the gamma detections helps to improve the signal-to-noise ratio of scans. One of the recent focuses of the PET community has been the pursuit to build longer axial field-of-view (AFOV) scanners, to increase the probability of detecting the gamma photons. This translates to better imaging quality and diagnostic capabilities, while also offering the potential for dose reduction. These long AFOV systems are, however, very expensive due to the large number of detectors required. Finding

ways to reduce the system cost can therefore help to make these systems more widely available to a larger population.

1.2 Outline

This dissertation is focused on simulation studies and software development for a new, cost-efficient, flat-panel, long AFOV PET design based on monolithic detector technology: the Walk-Through PET (WT-PET). The thesis deals with two main components: time resolution in monolithic detectors, and the development of an image reconstruction framework for the new PET system. First, some necessary background knowledge on the principles behind PET and image reconstruction are given in **Chapter 2**. **Chapter 3** provides some further introduction by discussing two new evolutions in PET: long AFOV PET and the use of deep learning for signal processing and image enhancement. The section on deep learning (Section 3.2) has been published as part of a review paper [1]. **Part I** then deals with the monolithic detector design. Here, **Chapter 4** involves a simulation study to better understand the effect of various detector parameters (scintillation material, size and surface finishes) on timing performance in monolithic detectors, with special attention given to the impact of Cerenkov photons. This chapter concerns published work [2]. **Chapter 5** involves the development of a deep learning based approach to time estimation in monolithic detectors, and offers a comparison with more conventional methods. This study has also been published [3]. Next, **part II** involves the development of image reconstruction algorithms for the WT-PET. First, the WT-PET system is properly introduced in **Chapter 6**, explaining the design philosophy behind it, and describing what sets it apart from existing (long AFOV) PET scanners. We also have a look at some key performance parameters based on simulation results. Then, in **Chapter 7** the reconstruction framework is introduced, focusing on the calculation of the system response matrix, and an evaluation of the image quality for the WT-PET. Some of these results in Chapters 6 and 7 have also been published as part of a more extensive simulation study [4]. The remaining chapters deal with alternative or additional data corrections for the image reconstruction. An energy based scatter correction method for the WT-PET is developed and evaluated in **Chapter 8**. In **Chapter 9**, two methods for CT-less attenuation correction are investigated, one based on the use of a transmission source, and the other using the emission data itself. In

Chapter 10, we take a closer look at the limited angle artifacts arising due to the flat-panel geometry of the system. A deep learning based approach to limited angle artifact correction is explored, and incorporated within the iterative image reconstruction procedure. Finally, we finish with some concluding remarks and future perspectives in **Chapter 11**.

Chapter 2

Positron emission tomography

Positron emission tomography (PET) is a functional imaging technique used to study metabolic processes in living organisms. It makes use of a radiotracer: a chemical compound in which one of the atoms has been replaced with a radioactive isotope. This radiotracer is administered to the patient, and the PET scanner is used to visualize its distribution within the body. Nowadays, PET is almost always combined with computed tomography (CT) into a dual imaging modality appropriately named PET-CT. The CT provides additional anatomical information for diagnostic and localization purposes.

Different radiotracers are used to study a variety of metabolic processes, each associated with different chemical reactions. For example, fluorodeoxyglucose (FDG), the most frequently used radiotracer in PET, is a glucose analogue labeled with the isotope ^{18}F. It is used to identify cancerous tissue, since malignant cells consume large amounts of energy and therefore exhibit increased glucose uptake. As such, PET plays a major role in oncology where it is used for lesion detection, diagnosis and disease prognosis in cancer patients. It is also used in post-treatment follow-up studies, where the aim is to examine how the patient is responding to treatment, to check for any residual tumor after therapy, and to monitor the possibility of cancer recurrence in the longer term. An example FDG scan is shown in Figure 2.1 [5]. Another common family of radiotracers are prostate-specific membrane antigen (PSMA) targeting ligands. PSMA is an enzyme which is highly expressed in the prostate, with cancerous prostate cells showing increased levels compared

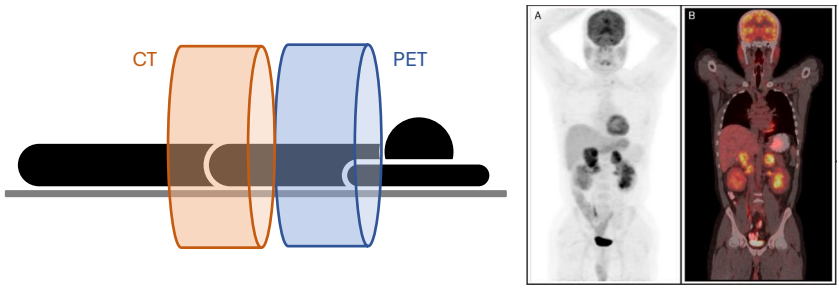


Figure 2.1: Schematic of a PET-CT scanner, and an example FDG PET image (**A**), and the PET image overlaid on top of the CT (**B**). PET-CT image from [5].

to noncancerous prostate cells. There exists a variety of radiotracers, using different isotopes such as ^{18}F, ^{68}Ga, ^{99m}Tc and ^{111}In, which bind to PSMA and thereby allow for diagnosis and staging of prostate cancer [6]. PET is also used in neurology, e.g., in the diagnosis of Alzheimer disease, which is associated with the progressive deposition of amyloid-ß peptides in the brain. Here, tracers which specifically bind to these peptides, such as ^{11}C-labeled Pittsburgh compound B ([^{11}C]PiB), are often employed [7]. PET is further used in studies involving epilepsy [8], Parkinson's disease [9], cardiology [10], etc., enabled by the wide diversity of available tracers.

Throughout this chapter, we will delve deeper into the physical principles behind PET, both on the system and detector level, providing some necessary background knowledge for this dissertation. We will give special attention to the different aspects that influence system performance and image quality in PET.

2.1 PET physics

As the name suggests, PET makes use of isotopes that emit a positron upon decay (β^+ decay). These isotopes are produced in a cyclotron, a particle accelerator where non-radioactive materials are bombarded by protons to produce radioactive isotopes. They are short-lived, with a typical half-life in the range of a few tens of minutes to a few days. During a PET scan, shortly after radionuclide decay, the emitted positron

annihilates with a nearby electron, in turn generating two 511 keV gamma photons in almost opposite directions. Detection of such a gamma photon by the scanner is commonly referred to as a single event. If two of these gamma photons (singles) are detected within a certain coincidence time window (CTW), they are assigned to the same annihilation event, and a line of response (LOR) can be derived along which the positron annihilation took place. Detection of the gamma photons is done by scintillation crystals, which absorb the gamma ray by means of the photoelectric effect and re-emit the energy as visible light. These optical scintillation photons are then detected by photomultiplier tubes (PMTs), or nowadays, silicon photomultipliers (SiPMs) attached to the crystal, converting the optical signals into an electrical current, which can be read out and processed by the system's electronics. By registering many of these coincident gamma photons, line projections of the emission distribution can be obtained at different angles. Tomographic image reconstruction techniques are then used to reconstruct the tracer distribution from this projection data.

This is an idealized version of what is happening during the imaging procedure in PET. In reality, there are a variety of different effects that introduce noise or uncertainties in the process, contributing to image degradation. We will discuss them now in the order of the image formation process.

Uptake time

The tracer distribution post-injection is dynamic, and a different image is observed at different time points. At the start, most of the radiotracer is still in the bloodstream, resulting in more activity for the background tissue. As time goes on, the radiotracer accumulates in the relevant organs and potential lesions, giving them a higher contrast compared to the background. Eventually, the radiotracer makes its way to the bladder for excretion. Meanwhile, the isotopes are decaying, with for example a half-life of 110 minutes for ^{18}F, resulting in fewer emissions as time goes on. That means that there is usually an optimal time point for scanning, where the radiotracer has been given sufficient time to properly distribute through the body, but there still remains sufficient available activity. In the case of FDG studies, the waiting time between injection and scanning is usually around 1 hour.

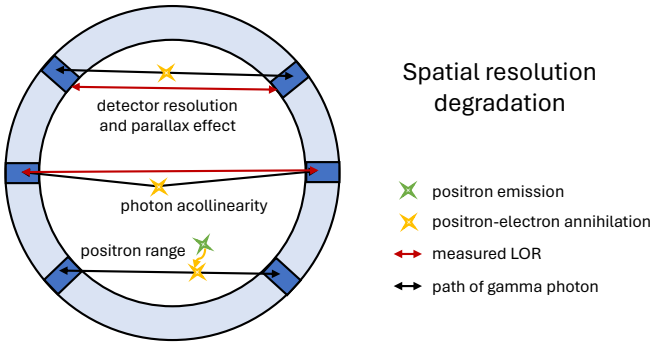


Figure 2.2: Overview of the spatial resolution degrading effects in PET.

Positron range

After positron emission, the positron collides with atoms and slows down, eventually annihilating with one of the electrons in the material. The positron range is the median distance traversed by the positron between emission and annihilation, and therefore forms an intrinsic limitation to the resolution in PET (see Figure 2.2). While we are interested in the emission position (tracer location), we instead measure the annihilation position along the LOR, resulting in a small degradation of spatial resolution. The positron range depends on its initial kinetic energy and the electron density of the surrounding medium. Isotopes that emit more kinetic positrons exhibit a longer positron range, while denser tissue results in a shorter positron range. For ^{18}F, the (median) range is relatively small and in the order of 0.4 mm for soft tissue, increasing to 1.4 mm for lung tissue [11]. In clinical FDG studies, positron range is therefore a negligible contribution to spatial resolution degradation, as typical spatial resolutions of PET systems are in the range of 3 - 4 mm. Some other isotopes such as gallium-68 (^{68}Ga) have considerably longer positron range (2.4 mm in soft tissue and 8 mm in the lung [11]), to the point that its effect can become visible in clinical PET scans.

Electron-positron annihilation

Upon collision, there is a chance that the electron and positron do not annihilate instantly, but instead form a short-lived bound electron-positron pair, called positronium. The probability of such a system forming is about 38% in water, with a lifetime in the order of 125 ps to

140 ns, depending on the exact positronium state [12]. Most of these states also decay into two gamma photons and are therefore measured just the same in PET, but a small fraction ($\sim 0.5\%$) results in three-gamma decay. Although a negligible contribution in conventional PET usage, measurement of three-gamma decay has been proposed as a method for positronium lifetime imaging [13].

Photon acollinearity

When the two gamma photons are produced by the electron-positron annihilation, the non-zero kinetic energy of the positron at annihilation results in gamma photons being emitted not entirely in opposite directions, but instead with a mean angular deviation of 0.2° full width at half maximum (FWHM) [14]. This effect is named photon acollinearity, and causes the LOR to not entirely coincide with the annihilation position, posing another intrinsic resolution limitation of the imaging process (see Figure 2.2). Its effect on system spatial resolution increases with the diameter of the system, and is around 1.8 mm FWHM for the standard PET diameter of around 80 cm [15].

Attenuation

Once the gamma photons are emitted, there is a possibility that one or both of the photons are attenuated (absorbed or scattered) within the patient, and do not reach the detectors (see Figure 2.3). The amount of attenuation depends on the travel distance and the density of the tissue, and needs to be taken into account during the image reconstruction process. Otherwise, the tracer distribution deeper within the patient would be underestimated, as a smaller fraction of the emitted coincidences are detected from these annihilations. As attenuation reduces the number of detected coincidences, it has an adverse effect on image quality and can make detection of lesions in larger patients more challenging. Larger patients do receive a higher dose (often normalized to weight, with a typical scan dose in Europe being ~ 3 MBq/kg), but this only partially offsets the loss due to attenuation.

Sensitivity

Not all photons escaping the patient will be detected by the PET scanner. A certain fraction will miss the scanner entirely, and yet another fraction will pass through the detectors without being (fully) absorbed (see Figure

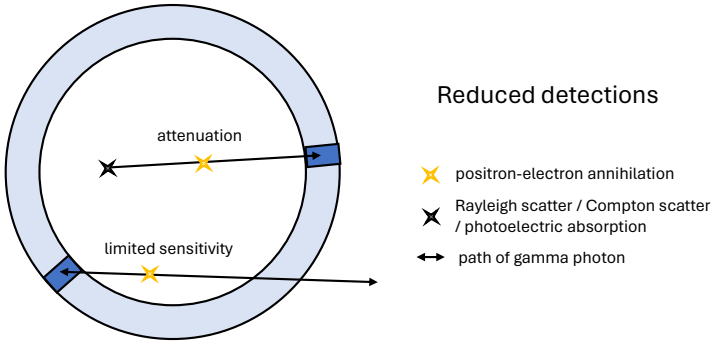


Figure 2.3: Effects resulting in reduced count statistics in PET.

2.3). A PET scanner's sensitivity indicates what fraction of emitted gamma photons, in the absence of attenuation, will be detected as coincidences. The higher the sensitivity, the better the noise statistics of the reconstructed images. We will delve into more detail on sensitivity, and how it can be improved, in Section 2.4.

True coincidences

A measured coincidence event is considered to be a true coincidence when both gamma photons came from the same positron annihilation, and both photons reached the detectors unperturbed, without scattering. True coincidences contain the necessary information to reconstruct the tracer distribution. In addition, there are also scattered and random coincidences, which do not contain any useful information on the tracer location, and manifest as additive noise for the measurement.

Scattered coincidences

Attenuation of gamma photons can happen through either the photoelectric effect (complete absorption), Compton scattering (inelastic scattering with loss of energy) or Rayleigh scattering (elastic scattering without loss of energy). At the energies and typical tissue densities involved in PET, Compton scattering is predominant. Sometimes, these scattered photons miss the scanner entirely and remain undetected, but in other occasions they may be detected as a scattered coincidence (as opposed to a true coincidence), see Figure 2.4. A scattered coincidence is a coincidence event where both gamma photons came from the same

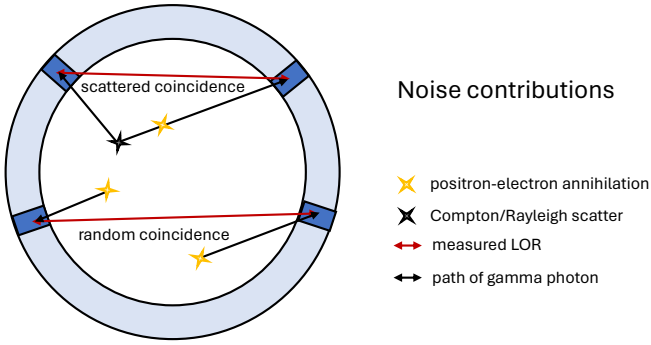


Figure 2.4: Additive noise contributions in PET.

annihilation event, but one or both of the photons were scattered in the patient, resulting in LOR mispositioning. As the photon loses energy when undergoing Compton scattering, a large portion of scattered events can be removed by using a minimum energy threshold on the detectors. If the energy resolution of detectors were perfect, this would allow to completely remove all Compton scattered coincidences. Real detectors however have a limited energy resolution, which largely depends on the scintillation material and light detectors. Better energy resolutions allow for an increased rejection of scattered coincidences.

Scattered coincidences are highly dependent on the activity and attenuating medium distributions. As a result, they not only increase noise, but also lead to a spatially varying change of intensity in reconstructed PET images. Therefore, they should be specifically corrected for during image reconstruction.

Random coincidences

Besides true and scattered coincidences, there is a third type of event: a random coincidence. A random coincidence may occur when two gamma photons are detected within the same CTW, but were actually a result of two separate annihilation events (see Figure 2.4). This results in completely random LOR positions, resulting in increased background noise. The rate of random coincidences depends on the injected activity and the duration of the CTW. A higher injection dose leads to more detections overall, thereby increasing the random coincidence rate. At very high activities, the randoms rate may become so large that the true coincidence rate saturates, or even starts to drop. This is because,

what would otherwise be a true event, is missed when one or both of its detections are part of a random coincidence. The random rate can be reduced by decreasing the CTW, although the window should remain sufficiently large to measure emissions from anywhere within the desired FOV. Furthermore, when coincidence processing is done in real-time on the electronics, clock speeds and other limitations may put a lower bound on the CTW. In practice, CTWs are usually in the range of a few ns.

As was the case with attenuation and scattered coincidences, random coincidences should be corrected for during the image reconstruction procedure in order to obtain correct activity values and reduce background noise.

Detector spatial resolution

PET detectors do not have perfect spatial resolution, resulting in a mispositioning of the LOR endpoints, see Figure 2.2. Not only the 2D resolution, but also the depth-of-interaction (DOI) decoding capabilities of detectors are important. That is, the ability to determine how deep within the detector a gamma interaction took place. In PET systems without DOI capability, the spatial resolution is not uniform over the FOV due to the parallax effect. Off-center LORs are increasingly incorrectly positioned, resulting in a degradation of spatial resolution for increasing radial distances in the transverse plane. Current clinical PET scanners have detector resolutions in the range of 3-4 mm and no DOI capability, making the detector resolution the primary contributing factor to system-level spatial resolution degradation.

Time-of-flight (TOF)

As PET detector technology improved, it became possible to measure the time difference between the detection of two coincident gamma photons to an accuracy of a few hundred ps, leading to the implementation of time-of-flight (TOF) PET. TOF information provides an estimate of where along the LOR the annihilation took place, and the spatial accuracy along the LOR is given by $\Delta x = c\Delta t/2$, with c the speed of light and Δt the TOF resolution, see Figure 2.5. The TOF kernel (probability distribution for the error on the TOF difference measurement) is usually modelled as a Gaussian distribution, and the TOF resolution of a PET system refers to the FWHM of this kernel. The newest clinical PET scanners have TOF resolutions close to 200 ps FWHM, which corresponds to a spatial

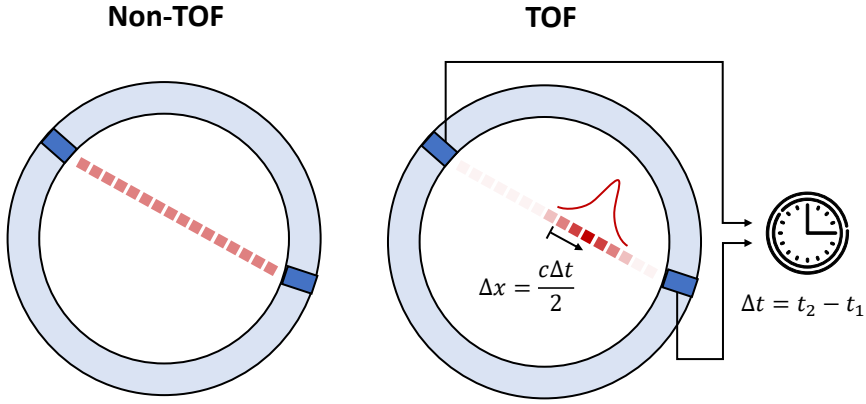


Figure 2.5: Illustration of time-of-flight (TOF) in PET.

resolution along the LOR of around 3 cm FWHM. This is insufficient to determine the annihilation position directly, but the information can nonetheless be incorporated into the image reconstruction procedure, improving the signal-to-noise ratio.

2.2 PET detector technology

2.2.1 Components

A typical PET detector consists of three main components: the scintillation crystal, an array of SiPMs, and readout electronics.

Scintillator

The role of the scintillation crystal is to absorb the high-energy (511 keV) gamma photons and re-emit their energy as multiple optical photons. This requires specific materials, with two common ones for PET being L(Y)SO (lutetium(-yttrium) oxyorthosilicate) and BGO (bismuth germanate). Whereas BGO was more prominent in the past, most PET scanners nowadays make use of L(Y)SO. This is due to the development of TOF-PET, where the superior timing properties of L(Y)SO are most desirable.

Scintillation happens through a three-step process consisting of (i) conversion, (ii) transport and (iii) luminescence. In the **conversion** step,

the interaction of the gamma photon with an atom in the crystal causes an electron to be emitted from its shell through either the photoelectric effect or Compton scattering, creating an electron-hole pair. In the case of the photoelectric effect, the gamma photon is completely absorbed, and its energy is fully transferred to one of the core electrons of the atom. On the other hand, in Compton scattering the gamma photon only transfers part of its energy, emitting an electron from one of the outer shells. Compton scattered gamma photons may then further interact with the crystal through another Compton interaction or the photoelectric effect. In either case, the generated electron-hole pair is highly energetic, causing it to interact with other particles in the material, leading to an avalanche event where many secondary electrons and holes are created.

In the **transport** step, these electrons and holes migrate towards luminescence centers in the material, where they are recombined, and the excess energy is released as visible light through the process of **luminescence**. While the conversion process happens very quickly (in the order of 1 ps [16]), considerable delay can occur in the transport and luminescence steps. This results in a scintillation pulse that can be (approximately) characterized by a bi-exponential relationship with a fast rise time (e.g., in the order of 70 ps for LSO [17]) and a slower decay time ($\sim$ 40 ns for LYSO and $\sim$ 300 ns for BGO [18], [19]). The transport step also has a considerable impact on the scintillation efficiency (number of optical photons generated per MeV of absorbed energy), as electrons and holes can become trapped, or recombine without luminescence. For LYSO, in the order of 30 000 photons are generated per MeV (15 000 per 511 keV event), whereas for BGO, it is in the order of 8 500 photons per MeV (4 250 per 511 keV event) [18], [19]. This also results in a better energy resolution for L(Y)SO, as the higher number of generated photons allows for a better approximation of the total energy deposited in the crystal. Some typical energy resolutions are around 11% FWHM for L(Y)SO and 15% FWHM for BGO, as measured around the 511 keV peak [18], [19].

Silicon photomultiplier (SiPM)

The scintillation photons can then be detected by silicon photomultipliers (SiPMs) coupled to the crystal. These SiPMs are attached to the crystal using, for example, optical glue, with a refraction index that matches that of the scintillation material as close as possible, to ensure efficient photon transport from crystal to SiPM. An SiPM contains a large number

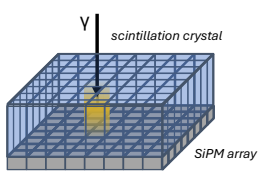
(hundreds or thousands) of microcells connected in parallel, with each microcell consisting of a single-photon avalanche diode (SPAD) with a quenching resistor. A SPAD is based on a semiconductor p-n junction, which creates a free charge carrier (electron and/or hole) upon excitation by an optical photon. As the SPAD is operated at a reverse bias above the breakdown voltage, the charge carrier is rapidly accelerated by the electric field, sufficiently increasing the carrier's kinetic energy to overcome the ionization energy of the bulk material. This results in electrons being knocked out of their atoms, creating an avalanche of more charge carriers, and causing the current to grow exponentially from as little as a single absorbed photon.

The SiPM response to a single photon is roughly a bi-exponential function with $\sim$ 70 ps rise time and $\sim$ 40 ns decay time, arising from the RC-circuits (resistor- and capacitor-circuits) connected to the SPAD [20]. As the SPADs are connected in parallel, the overall SiPM signal arising from multiple photons hitting the SiPM within a short time-frame will be a sum of individual SPAD signals. As such, SiPMs have a wide dynamic range, being able to detect single photons, up to thousands of photons, due to the arrangement of multiple microcells in a single circuit.

While SiPMs can detect single photons, this is complicated by the presence of dark counts. A dark count is the result of an electron initiating an avalanche due to thermal, rather than optical, excitation. As the ensuing SiPM signals are identical regardless of excitation method, dark counts form a source of noise at the single photon level. The dark count rate is proportional to the number of microcells (active area of the SiPM), but can be reduced by operating the SiPM at a lower temperature.

Electronic readout

The role of the electronics is to read out the SiPM signals and extract relevant information in real-time, usually a timestamp and energy (charge integration) per SiPM. The readout system may also contain a pre-amplification circuit to further increase the SiPM signal prior to processing. The energy is commonly measured using an analog-to-digital converter (ADC), and the timestamp can be obtained using a thresholding technique, measuring the time at which the SiPM signal crosses a specific voltage threshold. For good time resolution in PET, this threshold should be set sufficiently low in order to measure the first few arriving photons, although the dark count rate may complicate measuring at the single photon level, where dark counts can result in



pixelated		monolithic
3 – 4 mm	**spatial resolution**	~ 1 mm
no	**DOI capable**	yes
good	**time resolution**	decent
limited	**calibration required**	yes
higher	**cost**	cheaper

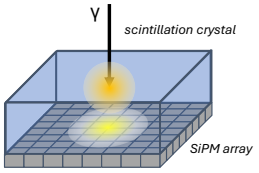


Figure 2.6: Summary of the main differences between monolithic and pixelated PET detectors.

false triggers. Dedicated PET electronics may also contain the logic for event processing, such as rejection of events based on energy thresholds and coincidence sorting.

2.2.2 Monolithic and pixelated detectors

We can classify (the majority of) PET detectors in two groups: pixelated and monolithic detectors. A summary of the main differences is given in Figure 2.6.

Pixelated detectors

Current clinical scanners make use of pixelated detectors. Here, the scintillation crystal is subdivided into an array of smaller crystals (pixels), so that inter-pixel optical photon transfer is limited. This is achieved by inserting reflecting walls between the pixels. Pixelated detectors have the advantage of straightforward positioning and timing estimation of gamma events, as the event can easily be assigned to a single pixel, and only one SiPM waveform is needed to measure the gamma interaction time. They can achieve good time resolutions, as all the scintillation light is focused onto a single SiPM, resulting in the waveform quickly rising above the noise level for early timestamp thresholding. The downside is that the spatial resolution is limited by the pixel size, which is usually around 3 - 4 mm wide and 10 to 30 mm thick. Depth-of-interaction (DOI) information is also not available, unless some, usually expensive, modifications are made, such as dual-sided readout systems. In theory, it is possible to further reduce the pixel size, in width and/or thickness, to obtain better spatial resolutions. Reduced crystal thickness can also improve the time resolution by reducing the spread in optical photon travel times within the crystal. This would, however, negatively impact the detector's ability

to absorb gamma photons, as either more non-scintillating material is used for the reflecting walls, or the crystal thickness is reduced. The crystal cutting procedure also becomes more difficult, and therefore expensive.

Monolithic detectors

Monolithic detectors on the other hand consists of one large continuous crystal, usually a few centimeters wide, coupled to an array of multiple SiPMs. They have up until now only been used in research settings or preclinical systems. Since monolithic detectors do position estimation by analyzing the distribution of scintillation light, they are capable of achieving much better detector resolutions, in the range of 1 mm FWHM. Furthermore, by analyzing how spread out the scintillation light is, it is possible to obtain DOI information without the need for additional readouts. The lack of crystal subdivision further results in reduced crystal cutting costs and somewhat higher sensitivities for the same detector volume. A large downside of monolithic detectors is, however, the rather complex and time-consuming calibration procedure. Time estimation can also be more problematic, as the scintillation light is spread out over multiple SiPMs, resulting in fewer photons per SiPM and therefore an increased relative contribution of dark counts.

In part I, we investigate the timing capabilities of monolithic detectors, and look for potential ways to optimize them. Chapter 4 studies the potential use of Cerenkov emission for improved time resolution in monolithic BGO, with special attention given to the role of crystal surface finish on photon collection efficiency and timing capabilities. In Chapter 5, we investigate the use of deep learning to improve time resolution in monolithic LYSO detectors beyond what is possible with conventional methods.

2.3 Image reconstruction

The raw data produced by a PET scanner (the measured LORs) can not be interpreted directly by a physician. Instead, tomographic image reconstruction techniques are required to convert this projection space data back into image space, producing tomographic images. The PET

acquisition procedure can be represented mathematically as

$$\begin{aligned} \mathbf{y} &= \mathbf{H}\lambda + \mathbf{e} \\ y_i &= H_{ij}\lambda_j + e_i \end{aligned} \tag{2.1}$$

where $\mathbf{y}$ is the measured data (the number of detections along each LOR), λ is the activity (tracer) distribution, $\mathbf{H}$ is the system response matrix (SRM) and $\mathbf{e}$ is an additive noise contribution to the measured data due to random and scattered coincidences. The projection space data $\mathbf{y}$ is also said to be sinogram data, due to the fact that projection data of an off-center point source, visualized in 2D in function of a radial distance and a projection angle, is a sinusoid. The system response matrix describes the physics and geometry of the imaging process, where the elements H_{ij} of the SRM represent the probability that an emission in voxel j would result in a detection along LOR i. The purpose of image reconstruction is then to retrieve the tracer distribution λ from the measured data $\mathbf{y}$.

2.3.1 Evolution of reconstruction techniques

Analytic reconstruction

The inverse problem of equation 2.1 is ill-posed and lacks an exact solution, but an analytical closed-form approximation can be found by the filtered back-projection (FBP) algorithm. In this algorithm, the measured data is projected back into image space directly, by smearing out the contribution of each LOR over the image. A convolution filter is applied during the process (usually a ramp filter with frequency response $|\omega|$), to counteract blur that would otherwise occur due to an oversampling of small spatial frequencies. Such an analytical solution has the advantage of computational simplicity, but its failure to model scanner non-idealities or noise statistics leads to inaccurate image estimations. While analytical reconstruction techniques are still very common in, for example, CT, they are no longer the favored method in PET.

Iterative reconstruction

It is instead preferred to find a solution to equation 2.1 by minimizing an objective function f, which provides a measure for the reconstruction accuracy:

$$\hat{\lambda} = \arg\min_{\lambda}[f(\mathbf{H}\lambda + \mathbf{e}, \mathbf{y})] \tag{2.2}$$

A common choice for f is the (negative) likelihood function, where for PET, due to the Poisson distribution of single photon detections, the Poisson log-likelihood is used. The extremum of this function can be found using the maximum likelihood expectation maximization (MLEM) algorithm, where a single update, from iteration k to $k+1$, is given by:

$$\lambda^{(\mathbf{k+1})} = \frac{\lambda^{(\mathbf{k})}}{\mathbf{H^T 1}} \mathbf{H^T} \left(\frac{\mathbf{y}}{\mathbf{H}\lambda^{(\mathbf{k})} + \mathbf{e}} \right) \tag{2.3}$$

where $\mathbf{1} \in \mathbb{R}^{\mathbf{I}}$ (with I the number of LORs) and $\lambda^{(\mathbf{0})}$ is initialized by uniform values.

The update equation can intuitively be understood as follows: the denominator on the right is a forward projection of the current estimate of the activity distribution, which is compared to the measured data $\mathbf{y}$ in the numerator. This mismatch in projection space is then projected back to image space by $\mathbf{H^T}$, providing a correction factor to be multiplied with the current image estimate in order to converge towards a solution matching the measured data. The denominator on the left ($\mathbf{H^T 1}$) can be seen as a sensitivity correction factor in image space, taking into account the sensitivity of each voxel (probability of an emission resulting in a detection) in the presence of attenuation. It is therefore commonly referred to as the sensitivity image.

Image quality is improved compared to analytical methods by making explicit use of the forward model during each update, in which various physical limitations and non-idealities of the imaging process can be incorporated. In addition, knowledge of additive noise $\mathbf{e}$ can be included during the forward projection. We note that in FBP, the only way to correct for random and scattered coincidences is to subtract them from the measured data, prior to back projection. This is however only correct when the exact number of random and scattered coincidences along each LOR is known, which is never the case in practice. Incorporation in the forward model, as is the case in equation 2.3, however relies on the expectation value of contributions to each LOR, which can be approximated much more reliably.

Image reconstruction in PET largely deals with methods to calculate or approximate the elements of the system response matrix H_{ij}. Importantly, the full matrix $\mathbf{H}$ is too large to be kept in memory, as its rows represent the LORs and the columns represent the voxels, both of which can be in the order of hundreds of millions or more. For example, the Siemens Biograph Vision Quadra [21], a long axial FOV PET scanner

(i.e., larger than most other scanners), has a total of 243 200 detector elements, resulting in $\sim$ 30 billion LORs ($n(n-1)/2$ with n detector elements). Admittedly, many of these LORs can safely be ignored as they do not sufficiently intersect the FOV, but we can still assume that at least $\sim$ 10 billion LORs are required. Reconstructing over the entire FOV of the scanner ($\sim$ 78 x 78 x 106 cm^3) using 1 mm^3 voxel sizes requires 780 x 780 x 1060 $\simeq$ 645 million voxels. That means that the system matrix $\mathbf{H}$ would contain approximately 10 billion x 645 million = 6.45 x 10^{18} elements. Assuming these are stored in half-precision (16 bit) floating-point format, this would require 12.9 million TB of memory. For reference, the total amount of memory available on the (at the time of writing) default cluster of the Ghent University HPC (high-performance computing) system is "only" $\sim$ 34 TB.

To a good approximation, most elements can however be assumed to be zero, namely those for which the LOR i does not intersect with the voxel j. Computing which voxels a LOR intersects with, is done on the fly during each iteration, usually by making use of some ray tracing algorithm. This forms the backbone of most iterative image reconstruction algorithms in PET. In practice, this means that reconstruction can be done on a system with available memory in the range of a few tens of GB to a few TB, depending on the exact system, reconstruction FOV, voxel size, etc...

To further speed up the iterative reconstruction procedure, the projection data is sometimes split up into subsets. Every iteration is then computed using only the data of a single subset, alternating between the subsets. This gives rise to the ordered subset expectation maximization (OSEM) algorithm, which improves reconstruction speed, at the cost of increased noise and some potential convergence issues.

From 2D to 3D reconstruction

In the earlier days of PET, scanners would make use of septa between detector rings in order to limit coincidences between different rings. This would allow for 2D reconstruction algorithms to be used in a time when the available memory and computational power was insufficient to reconstruct 3D PET data. Furthermore, it would lower the count rates so that, the then slower, electronics could keep up. Each slice of the image could be reconstructed separately, and afterwards stitched together into a 3D image. The use of septa also had the advantage of limiting random and scattered coincidences, as random coincidences between

different rings were disallowed, and scatter would mostly be limited to angles in the 2D imaging plane. The septa however also blocked many true coincidences, decreasing count statistics and resulting in a loss of sensitivity.

After removal of the septa, LORs could be acquired in 3D, but still be rebinned into 2D projection data. This would allow increasing count statistics while still making use of more computationally efficient 2D reconstruction techniques. More accurate, however, is to both acquire and reconstruct data fully in 3D. This vastly increases memory requirements, as 3D PET data produces 4D projection data (represented in sinogram space by the transverse radius r, the azimuthal angle ϕ, the axial offset z, and the polar angle θ). Inclusion of TOF information further increases this to 5D. Nonetheless, due to the rapid advancement of computational hardware, 3D reconstruction algorithms are nowadays the standard practice in PET.

Returning to the example of the Quadra with $\sim$ 10 billion LORs, using a one-to-one mapping between LOR and sinogram bin, non-TOF projection data stored in sinogram format would require 20 GB of memory when using 16 bit representations. Including TOF (Quadra TOF resolution = 228 ps) with for example 200 TOF bins (a 5 ns coincidence time window, in total a 10 ns range when accounting for both negative and positive time differences, subdivided into 50 ps bins) increases this to 4 TB. To decrease memory requirements, especially in systems with many detector elements, it is a common approach to instead store data in so-called list-mode format, where only the measured LORs are stored, listing each measured LOR separately. This is in contrast with sinogram format, where the number of detections are stored for each potential LOR, resulting in many zeros (unmeasured LORs) for systems with many detector elements, and therefore redundant memory usage.

In Chapter 7, we develop our own 3D PET image reconstruction framework. It is developed to help design, test, and ultimately be used on the Walk-Through PET. The reconstruction framework is therefore adapted for systems based on monolithic detector technology with non-cylindrical (e.g., flat-panel) geometries.

2.3.2 Data corrections

As has been alluded to in Section 2.1, certain aspects of the PET image formation process should be taken into account during the image reconstruction process in order to obtain correct activity values. These

are included in equation 2.3, either in the system response matrix **H** (attenuation correction), or in the additive noise term **e** (random and scatter correction).

Attenuation correction

Attenuation correction in PET requires a map of the attenuation coefficients of the imaging object or patient, in order to estimate and correct for the loss of coincidence detections due to attenuation. In the early days of PET, this map would be derived using a Germanium-68 (^{68}Ge) rod source, rotating along the inner diameter of the system, producing a transmission scan at 511 keV. This transmission scan could then be compared to a blank scan (a scan without patient), in order to derive the attenuation coefficients. Nowadays, as PET is practically always combined with CT, attenuation correction is normally done using the accompanying CT-scan. The CT provides an estimate of the attenuation map at a lower energy (20 - 150 keV), which can then be rescaled to the energy in PET (511 keV).

However, the use of a CT can sometimes be undesired due to, e.g., radiation dose considerations. Therefore, we investigate some alternative methods for (CT-less) attenuation correction in Chapter 9, and implement them within our reconstruction framework.

Random correction

Estimating the contribution of random coincidences is fairly straightforward, and is typically done by implementing a **delayed coincidence time window** [22]. As the detections in random coincidences, by definition, come from a different positron annihilation, the rate of random coincidences for a certain CTW is independent of the presence of a delay. By choosing a sufficiently large delay, we can guarantee that no true or scattered coincidences remain present, thereby giving us a direct approximation of the random coincidence rate.

Another method for randoms correction is to estimate the random coincidence rate from the **singles count rate** as [22]:

$$R_{12} = 2\tau S_1 S_2 \tag{2.4}$$

where R_{12} is the randoms rate between two detectors with singles count rates S_1 and S_2, for a certain CTW τ. This method has the advantage of improved statistics, as there are far fewer detectors than there are

possible LORs, and the singles rate is often an order of magnitude larger than the coincidence rate. However, it requires that the singles rate either be measured directly at the detector level, or that the singles data be transmitted and stored somewhere for further processing. This is therefore not always done, either due to storage constraints, or due to data transfer rate limitations.

Scatter correction

Scatter estimation is, however, more complex. One approach is to estimate the contribution of scattered coincidences by making use of Monte Carlo simulation. While conceptually simple and providing accurate results, it has the major drawback of being incredibly computationally intensive, severely increasing reconstruction times.

Instead, a common approach is the **single scatter simulation** (SSS) algorithm, which approximates scatter contributions numerically by making use of the following information [23]:

- The estimated activity distribution
- The estimated attenuation distribution (scatter medium)
- The physical model of photon scattering (relationships between cross-section, scattering angle and energy)
- The scanner geometry and detector parameters

While SSS can provide good approximations of the scatter contributions, it does have a few drawbacks as well. As the name suggests, the algorithm operates under the (incorrect) assumption that photons are at most only scattered once. Luckily, this only has a small impact, as only a smaller fraction of all scattered coincidences are multiple scatter. Furthermore, most implementations of the SSS algorithm only give relative contributions of scatter estimations, meaning that they need to be normalized by comparison with the measured data. In practice, this is often done through a procedure called tail fitting, by fitting the tails of the scatter sinogram (where no true coincidences should be present) to the measured data. This does introduce an additional source of error, especially when the available tails are either noisy or small. Finally, SSS requires an estimate of the true activity distribution, that is, one without the contribution of randoms or scatters, for which the scatter

estimation is required in the first place. This means that scatter correction and image reconstruction depend on each other, and should ideally be optimized iteratively in an alternating fashion, which adds to the computational complexity. Nonetheless, the SSS algorithm is currently the most common method for scatter correction in PET.

There are however alternative options, such as energy-based scatter correction, which do not include many of these drawbacks. In Chapter 8, we give an overview of the methodology, and implement and evaluate energy-based scatter correction for the Walk-Through PET.

2.3.3 Regularization

Due to the ill-posed nature of image reconstruction, low statistics or small perturbances in the measured data can lead to large changes in the image estimate, easily resulting in an over-fit on the measured data, and causing noise to be amplified through further iterations. To counteract this, a regularization term R can be included in equation 2.2, penalizing unlikely solutions based on a priori assumptions about the image properties, such as demanding smooth or low-noise solutions:

$$\hat{\lambda} = \arg\min_{\lambda}[f(\mathbf{H}\lambda + \mathbf{e}, \mathbf{y}) + \beta\mathbf{R}(\lambda)] \tag{2.5}$$

The optimization problem becomes a trade-off between the data consistency term f (how well the image estimate λ produces the measured data $\mathbf{y}$) and the regularization term R (e.g., the overall noise level). The scalar regularization parameter β controls the relative strength between both terms. A common choice for R is for example total variation (TV), which promotes piece wise uniform solutions for λ. The ability to include regularization in the reconstruction is another advantage of iterative, over analytic, reconstruction techniques.

In Chapter 10, we use a deep learning based regularization approach to eliminate image artifacts during reconstructions for the Walk-Through PET, formed by missing projection angles in a flat-panel scanner geometry.

2.4 System performance

The system performance of a PET scanner can be quantified using a few benchmarks, which are meant to give an idea of the attainable imaging quality. There exist a number of standardized tests and metrics for

this purpose, governed by NEMA (National Electrical Manufacturers Association) [24]. These metrics are the sensitivity, noise equivalent count rate (NECR), scatter fraction (SF), image quality (IQ) and spatial resolution. Below, we give a short overview of what scanner parameters impact these metrics, and how they are measured in practice.

Sensitivity

The sensitivity of a scanner determines the ratio of detected to emitted gamma photon pairs, in the absence of attenuation. It depends both on the geometrical coverage of the system, and the sensitivity of the individual detectors. Conventional clinical PET scanners have relatively poor **geometrical coverage** due to their limited axial field-of-view (AFOV) of around 20 - 30 cm, resulting in the majority of emissions to remain undetected. If a portion of the patient larger than the AFOV needs to imaged, this is done by moving the bed into different positions. Sensitivity can be increased by enlarging the AFOV, although this incurs an increase in cost relative to the number of detectors. The individual **detector sensitivity** depends on the stopping power of the specific scintillation material in use (the attenuation coefficient for 511 keV gamma photons), and the crystal thickness. For example, BGO has a slightly higher stopping power than L(Y)SO. Increasing the detector thickness improves detector sensitivity, but increases material costs. It also negatively impacts the time resolution, as there is greater variability in photon transfer times from gamma interaction position to SiPM. Furthermore, increasing detector thickness has diminishing returns in terms of sensitivity, due to the exponential nature of attenuation resulting in the majority of gamma interactions happening closer to the front surface of the detector. Crystal thickness is therefore usually a compromise between sensitivity, time resolution and cost.

By the NEMA standards, the sensitivity is measured by making use of a 70 cm long ^{18}F line source of known activity, surrounded by five concentric aluminum sleeves. It is positioned along the axial length of the scanner, once in the center, and once at an offset of 10 cm in the radial direction, see Figure 2.7 (A). The sensitivity is usually reported in cps/kBq, and the sensitivity profile along the axial length of the scanner can be visualized.

Another effect closely related to sensitivity, is the TOF resolution. Disregarding the effects of random and scattered coincidences, the signal-to-noise ratio (SNR) in PET increases proportional to the square root

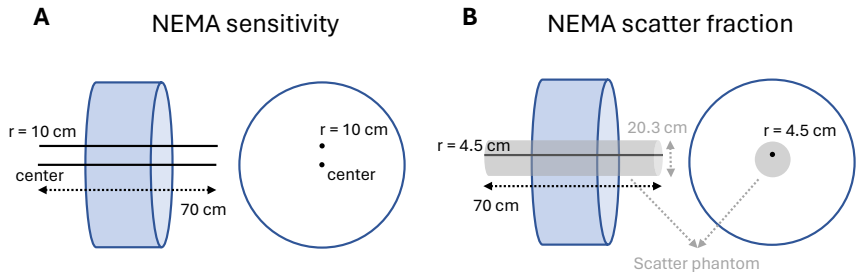


Figure 2.7: NEMA source and phantom used for sensitivity (**A**) and scatter fraction (**B**) measurements.

of the sensitivity. The SNR is further improved with the inclusion of TOF information, due to the localization of the annihilation position along the LOR. The gain in SNR from non-TOF to TOF is a factor of $\sqrt{(1/1.6) \cdot (D/\Delta x)}$, for an object of diameter D and a system with TOF resolution $\Delta t = 2\Delta x/c$ [25]. It is therefore said that TOF improves the effective sensitivity of a PET scanner, making the TOF resolution an important parameter in modern PET scanners.

Scatter fraction (SF)

The scatter fraction provides a measure for the relative contribution of scattered coincidences in the measured data. Scattered coincidences, even when corrected for in the image reconstruction, negatively impact the SNR. The scatter fraction, which is ideally as low as possible, is computed as:

$$\mathrm{SF} = \frac{S}{S+T} \tag{2.6}$$

where S and T are the number of scattered and true coincidences, respectively.

By the NEMA standards, it is measured using a 70 cm long ^{18}F line source, positioned along the axial length of the scanner, at a 4.5 cm radial offset. The source is placed in a polyethylene cylinder of 20.3 cm diameter, centered radially and of equal length as the source, which acts as a scattering medium, see Figure 2.7 (B). The acquisition is done at a low enough count rate, for which random rates are below 1% of the true rate. The values for S and T can be obtained from the projection data, using the single slice rebinning (SSRB) algorithm.

Noise equivalent count rate (NECR)

As the activity concentration within the FOV is increased, the rate of random coincidences increases more rapidly than that of true coincidences. At some point, the rate of true coincidences will even saturate and begin to decline, due to dead time effects on the detector or system level, and losses to random coincidences. The total (prompt) coincidence rate is therefore not a good measure for the noise statistics in a PET scan. Instead, a better measure is the NECR, representing an effective true rate. It has a direct impact on the SNR ($\text{SNR} \propto \sqrt{\text{NECR}}$) and is computed as:

$$\text{NECR} = \frac{T^2}{T + S + R} \tag{2.7}$$

where T is the true rate, S the scatter rate and R the random rate. In practice, the NECR will peak at a certain activity concentration, at which point injecting a higher radiotracer dose would only be detrimental. Ideally, the NECR peak should therefore be located at an activity higher than what would normally be administered to patients.

The NECR is measured according to NEMA standards using the same phantom as used for the scatter fraction, but this time with varying source concentrations to plot the NECR in function of the activity concentration. The peak NECR value, as well as the peak's location (activity concentration), are the two most important count rate performance parameters of a scanner.

Spatial resolution

The spatial resolution determines how accurately the annihilation position can be determined, and provides an estimate for the minimum distance required between two objects for them to be distinguishable from one another. Good system spatial resolution is especially important when imaging smaller lesions or structures. It is a combined effect of positron range, photon acollinearity, and limited detector resolution. In current clinical PET systems, the detector resolution is usually the primary limitation, resulting in overall system spatial resolutions in the range of 3 - 4 mm. It is also not uniform over the FOV due to the lack of DOI capabilities and the parallax effect, resulting in a degradation of spatial resolution for increasing radial distances in the transverse plane. In preclinical systems, the positron range may play a larger role, as their detectors often have higher resolutions, and the effects of photon acollinearity are reduced by the smaller scanner diameter.

In NEMA, spatial resolution is measured by placing a point-like source at six different positions in the scanner, three centered axially at a radial distance of 1 cm, 10 cm and 20 cm, and three at 1/8 of the AFOV, again at the same radial distances. NEMA recommends performing image reconstruction using the FBP algorithm, as iterative reconstruction algorithms may overestimate the spatial resolution for point sources in a cold background. Line profiles through the reconstructed images can then be used to measure the spatial resolution in the x, y and z directions for the different point sources, usually reported in FWHM. The FWHM serves as a good estimate for the minimum distance between two points at which they can be resolved as separate points.

Image quality (IQ)

Finally, overall image quality can be assessed by making use of a dedicated phantom: the NEMA IQ phantom. It is a (very) simplified model of the human torso, with a central lung insert and six spheres of varying diameter (10 mm - 37 mm) that represent lesions or other hotspots. The spheres and background can be filled with a radiotracer solution (e.g., FDG in water). The spheres are filled with a higher activity concentration, with a known sphere-to-background concentration ratio. An acquisition of the phantom can then be performed to evaluate image quality visually, as well as compute certain image quality metrics. Usually, these metrics are the contrast recovery coefficient (CRC) and background variability (BV). The CRC provides a measure for what percentage of contrast (between sphere and background) was recovered in the reconstructed image, by comparison with the ground-truth concentration ratio. Its value is normally between 0 and 1, where 1 indicates perfect recovery. The BV on the other hand provides a measure for the noise in the background. Together, these metrics provide an estimation of the SNR in reconstructed images for specific hotspot sizes. While the spatial resolution has a more limited impact on the larger sphere sizes, the CRC of the smaller spheres may be affected greatly in systems with poor spatial resolution. This is referred to as the partial volume effect, where the limited spatial resolution results in a loss of apparent activity, generally accompanying an increase in activity in adjacent regions.

Throughout part II, we will use the IQ phantom to evaluate the performance of the Walk-Through PET system, to compare to other systems, and to compare different reconstruction and/or data correction methods. The aforementioned metrics allow for a more quantitative

evaluation, in addition to a qualitative visual evaluation.

2.5 Monte Carlo simulation

NEMA has designed the aforementioned tests with an experimental setting in mind, with the purpose to evaluate system performance of existing PET systems. When it comes to designing new PET scanners however, we wish to know how certain design choices will impact performance prior to building the system. To a certain extent, general parameters such as spatial resolution and sensitivity can be approximated through analytical methods. But if high accuracy is required, or more complex variables need to be studied, Monte Carlo (MC) methods are the better, or sometimes, only choice. Monte Carlo simulation relies on the repeated sampling of random variables from known probability distributions in order to obtain numerical results. Even when experimental data is available, Monte Carlo simulation may still be useful as it provides access to ground truth data. It is therefore an invaluable tool in the field of nuclear medicine when it comes to designing new scanners, evaluating new algorithms, and improving the understanding of complex phenomena emerging from the various interactions between particles and matter.

There exist a number of Monte Carlo simulation software packages for nuclear medicine, but perhaps the most well known is GATE (Geant4 Application for Tomographic Emission) [26]. It makes use of Geant4 [27], a simulation toolkit for high energy physics, widely used at for example CERN. GATE, by making use of the extensive, well-validated, Geant4 libraries for physics models, is capable of providing very accurate simulation results, closely matching experiment. In PET, GATE is primarily used to model the radioactive decay of sources, the positron physics and annihilation with electrons, production of gamma photons and their interaction with matter, such as photoelectric absorption and Compton scattering. GATE provides tools specific for PET scanners, such as coincidence processing and aggregation of sinogram data. When performing studies on the system level, the GATE simulation would typically not include the detector physics itself, and only gamma photons and their interactions would be tracked and recorded. If we are specifically interested in investigating the detector physics, it is possible to include optical photon generation in GATE, so that we can simulate the scintillation process and optical photon transport through the crystal. Optical photon simulations are however slow, as thousands of optical

photons are generated per gamma photon. These types of simulations are therefore usually only done on the detector level itself. The SiPMs and electronics, however, can not be (accurately) simulated in GATE, and must be modeled separately if so desired.

A system-level GATE simulation for PET usually requires at least the following information, provided through a series of commands in macro (.mac) files:

- **Physics**: The physics model to be used in the simulation. Usually, a default model which encompasses many processes (e.g., the standard model) would be used, but it is also possible to enable and disable physics processes one by one. This may be done either to speed up the simulation, or to investigate a particular effect without the influence of others.

- **Geometry**: The scanner geometry, detector sizes and scintillation material. Ideally, other scanner components such as the casing would also be included, but normally these only have a minor effect, given that the majority of gamma photon interactions happen with the high density scintillation material.

- **Source**: The activity distribution, particle type and energy. In PET, the choice is often between a positron source and a back-to-back 511 keV gamma source. The back-to-back gamma source is computationally more efficient, but has the disadvantage that positron range will not be modelled. Photon acollinearity can however be included in both source types.

- **Phantom**: The attenuating medium in which the source is located. When a positron source is used, the phantom is also essential to provide the material in which the positrons annihilate. GATE also makes it possible to record the gamma interactions with the attenuating medium, which gives us the ability to easily distinguish between true, random, and scattered coincidences.

- **Digitizer**: A definition of the detector parameters, such as energy and time resolution, and the coincidence processing pipeline. The digitizer is meant to provide a simplified model of the detector response to gamma interactions, without the need to simulate scintillation photons, SiPMs or electronics.

For detector-level (optical photon) simulations in GATE, we also require the following:

- **Material information**: The material properties related to the generation of scintillation light and optical photon interactions, such as the scintillation yield, the scintillation pulse shape and refractive index of the material.
- **Surface information**: The surface properties of the interface between two materials, related to the reflection and absorption of optical photons. These properties can be both energy and incidence angle dependent, and have a large impact on the photon collection efficiency in PET detectors.

Throughout this thesis, we will make frequent use of GATE Monte Carlo simulation for our studies, both on the detector level (part I) and system level (part II). More specific simulation parameters will be discussed in their relevant chapters.

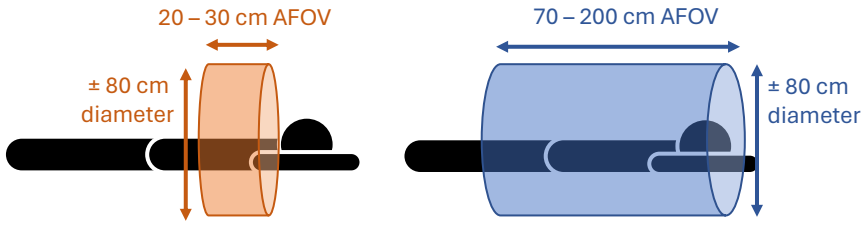
Chapter 3

New evolutions in PET

3.1 Long axial field-of-view PET

There is an ongoing effort to build more and more sensitive PET scanners, given the importance of sensitivity for image quality, and therefore diagnostic capabilities. Doing so by increasing detector thickness, however, offers diminishing returns due to the exponential nature of attenuation, and negatively impacts timing performance. Therefore, people have instead looked at increasing the AFOV, in order to improve the geometric coverage of PET scanners. This has lead to the development of long AFOV, or total-body, PET scanners, see Figure 3.1. Compared to the conventional AFOV of 20 - 30 cm, these scanners are characterized by an AFOV of 70 - 200 cm. The first long AFOV PET, the uEXPLORER, was built in 2018 at the University of California, Davis, and offers whole-body imaging with an AFOV of 194 cm [28]. It makes use of pixelated LYSO detectors with 2.76 $\times$ 2.76 $\times$ 18.1 mm^3 pixel sizes, and has a TOF resolution of 412 ps. Other long AFOV PET systems have since been developed, including the PennPET Explorer at the University of Pennsylvania (142 cm AFOV) [29] and the commercial Siemens Biograph Vision Quadra (106 cm AFOV) [21], both again based on pixelated L(Y)SO detector technology.

The gain in sensitivity for increasing the AFOV is highly dependent on the length of the object under investigation. For example, when the object of interest is the whole patient, increasing the AFOV from 20 cm to 200 cm increases the sensitivity by a factor 20 - 40. When the object of interest is substantially smaller however, such as for example a single organ, the gain in sensitivity is more modest with a factor ~ 3,



Short AFOV PET		Long AFOV PET
€ 2 – 3 million	**cost**	€ 8 – 12 million
4 – 8 min	**scan time**	< 1 min possible
~ 3 MBq/kg	**patient dose**	< 1 MBq/kg possible
3 – 4 patients/h	**throughput**	6 – 8 patients/h

Figure 3.1: Comparison between short and long axial field-of-view (AFOV) PET systems.

and negligible gain is observed for increasing the AFOV over 100 cm [30]. Depending on the patient/study, we may be interested in only a single organ, or the whole body. In general though, the legs of a patient rarely need to be scanned, due to their lack of organs. This makes an AFOV in the range of 100 cm, sufficient to allow hips, torso and head to be within the FOV at the same time, an interesting option in terms of price for performance.

The improved sensitivity of long AFOV scanners not only lends itself to better image quality, but can also be used to reduce injection doses or shorten scan durations. Reduced injection doses are especially attractive for imaging in at-risk populations, such as pediatric patients, or for post-treatment follow-up studies, where multiple scans need to be acquired at different time points. Shorter scan durations allow for increased patient throughput, making PET more accessible and reducing wait times. An enlarged AFOV also reduces or eliminates the need for scanning at different bed positions, further reducing scan durations.

The ability to scan the whole body, or at least head and torso, simultaneously, has also enabled dynamic whole-body PET imaging. The tracer distribution is not fixed in time, but instead it is injected through the bloodstream, reaches and accumulates in the relevant organs, until it makes its way to the bladder for excretion. Analysis of this time-dependent nature of the activity distribution can be leveraged to make kinematic models of the tracer, allowing us to study organ-organ

interactions and providing additional information not visible in fixed time point PET images. Dynamic PET imaging is still very much an active area of research, made possible by long AFOV PET systems, in which the relevant organs can remain within the FOV throughout the whole scan duration.

Whilst long AFOV PET offers substantial benefits over conventional PET, its widespread adoption remains somewhat limited, primarily due to economic reasons. The bulk of the system cost for a PET scanner comes from the detectors (scintillation material, SiPMs and electronics), with fixed costs such as patient bed, cooling system and computational hardware only making up a smaller fraction of the total cost. This means that the price of a PET system scales approximately linearly with the AFOV, almost quadrupling the cost when increasing the AFOV of a conventional 25 cm system to 100 cm, all other aspects being equal. This has lead to efforts in investigating alternative, more affordable long AFOV PET systems. For example, one option is to introduce a sparse detector design to reduce the number of detectors while keeping the AFOV. This of course does lower the sensitivity, however other benefits of long AFOV PET remain, such as simultaneous head and torso imaging, which enables single bed position scanning and dynamic PET imaging. Another option is to use cheaper scintillation materials, as is for example done in the J-PET system [31], which makes use of cheap, plastic scintillators. The low density of these scintillators does again result in a substantial loss of sensitivity.

Besides cost considerations, long AFOV PET also comes with some technical obstacles to overcome. The high sensitivity and AFOV results in a substantial increase of acquisition data, requiring better electronics to keep up with the data transfer rates, increased storage capacity for projection data, and better hardware to enable these reconstructions. Data corrections can be more challenging and time-consuming due to the increased number of potential LORs. The increased number of oblique coincidences due to the extended AFOV also negatively impacts spatial resolution when no DOI information is available. Despite these drawbacks, long AFOV PET remains a very powerful tool in nuclear medicine, and the number of installed long AFOV PET scanners is rapidly increasing.

In an aim to address these high acquisition costs for long AFOV PET systems, our research group has proposed the Walk-Through PET: a dual-panel, long AFOV PET system based on monolithic detector

technology. Its design reduces costs compared to normal cylindrical systems, while maintaining high sensitivity with a long AFOV. The system will be properly introduced in Chapter 6.

3.2 Deep learning in nuclear medicine

Another new development in PET has been the increased use of artificial intelligence (AI), and more specifically, deep learning (DL), in a wide variety of tasks related to medical imaging. Deep learning, as a subset of machine learning (ML), deals with systems or algorithms capable of automatically improving from experience, without being explicitly programmed. This is usually achieved by providing numerous examples (training data), to which the parameters of a mathematical model are fit. Once trained, the model can then be used to make predictions on new, unseen data. In the case of deep learning, this mathematical model is represented by a neural network, and the trainable parameters are the weights and biases of the neurons that make up the network.

Based on the type of example data and available information, we can define different types of machine and deep learning. In **supervised learning**, the most common type of machine learning, example data consists of known input-output pairs. Labelled data is available, and the model is trained such that its output is as close as possible to the desired label for every input. After training, the model can be applied to new unlabeled input data. The second type of machine learning is **unsupervised learning**, where no output labels are available. The aim is to find hidden structure in the input data, for example clustering algorithms that divide the data into groups of similar inputs. The final type of learning is often used in game playing or robot control, and is called **reinforcement learning** [32]. Here, an artificial agent learns a policy on which actions to take in an environment in order to reach a certain goal or maximize a cumulative reward. There is not one sequence of best actions, but an action is good if it is part of a good policy that in the end leads to a maximal reward. The agent explores the environment and possible actions using trial and error. Based on past good action sequences, the agent can learn a good policy.

Deep learning has found use in many aspects of the imaging pipeline in PET, such as detector signal processing, image reconstruction, image post-processing, image registration, image translation, image segmentation and disease diagnosis. We will give a short overview of deep learning in

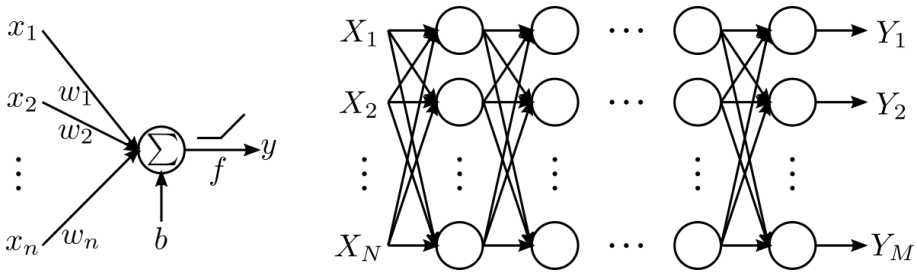


Figure 3.2: Schematic of a fully connected neural network.

topics relevant to this thesis, namely detector signal processing (Section 3.2.2), image reconstruction (Section 3.2.3) and image post-processing (Section 3.2.4). For a more extensive overview, which also includes other imaging modalities such as single photon emission computed tomography (SPECT), CT and magnetic resonance imaging (MRI), we refer the reader to our review paper on deep learning in nuclear medicine [1].

3.2.1 Relevant network architectures

Fully connected neural network

Deep learning is inspired by the biological functioning of the brain, in which networks of simple interconnected processing units called neurons are used to model complex functions [33], [34]. These artificial neurons or perceptrons take an input $\mathbf{x} = [x_1, x_2, ..., x_N]$, multiply it with weights $\mathbf{w} = [w_1, w_2, ..., w_n]$ and sum these weighted inputs with a possible bias b. This result is then passed through an activation function f to produce an output y [35]:

$$y = f(\sum_{i=1}^{N} w_i x_i + b) \tag{3.1}$$

Multiple neurons can be connected in layers to form a fully connected neural network, where the outputs of one layer serve as the inputs to the following layer, producing a mapping from input to output, see Figure 3.2. The role of activation functions is to introduce non-linearity in the network, required to model nonlinear relationships between input and output. A common activation function is the rectified linear unit (ReLU) [36], which simply sets negative output values to zero. Other popular activation functions are the sigmoid, hyperbolic tangent and leaky ReLU [37].

Convolutional neural network

The network shown in Figure 3.2 is a fully connected network, where all neurons of one layer are connected to all neurons of the following layer. When dealing with structured input data, however, such as 2D or 3D images in nuclear medicine, convolutional neural networks (CNNs) are commonly used instead [38]. These networks make use of convolutional layers, which consist of several kernels, containing the trainable weights or parameters, that are convolved with the input. They have the same number of dimensions as the input, but are usually much smaller in the spatial dimensions. The kernel size determines their receptive field. The kernel slides over the entire input with a predefined step size or stride, and at every position, a dot product is performed between the kernel and the current input patch. This way, a feature map is created containing the output responses of the kernel at every spatial position. Every convolutional layer consists of several kernels and produces an equal amount of feature maps. The motivation behind using convolutional layers is twofold: sparse connectivity and parameter sharing.

Sparse connectivity means that, in contrast to fully connected layers, the output neurons are not connected to all input units. Input images can contain millions of pixels, so instead of connecting a neuron with every input pixel, relevant features such as edges can be detected using kernels that are much smaller than the input. Although the receptive field of each kernel is small, deeper layers that interact with multiple outputs of earlier layers have an increasingly large receptive field with respect to the input. This allows the network to model complex interactions between simple building blocks across the input.

Parameter sharing denotes that the same kernel is used multiple times across the entire input, while in a fully connected network each weight is only used once. Consequently, a feature only needs to be learned once instead of multiple times for every location. Parameter sharing also causes a convolutional layer to be translational equivariant. This means that, if the input is translated, the output translates in the same way. This is especially useful when features, that detect edges for example, are relevant across the entire input. Moreover, because of parameter sharing, the input size does not have to be fixed, which allows processing inputs with varying sizes.

Sparse connectivity and parameter sharing results in a large reduction in number of parameters, which improves statistical efficiency and reduces memory requirements and amount of computations [39]. In addition

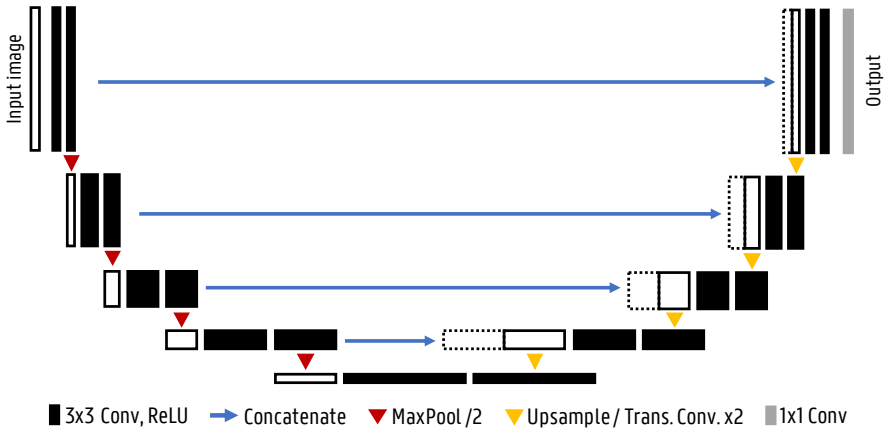


Figure 3.3: The U-Net architecture.

to convolutional layers, a CNN may also contain other layers such as fully connected layers, activation layers, pooling (subsampling) layers, normalization layers, etc.

U-Net

In 2015, U-Net (see Figure 3.3) was proposed as a biomedical image segmentation architecture [40]. The authors employed the architecture in several segmentation challenges such as segmenting neuronal structures in electron microscopy stacks or cell segmentation in light microscopy images and won with a large margin.

The typical use of CNNs was to classify an entire image into a single class label. In many computer vision tasks, however, localization is required where every pixel is labeled with the class of the object it belongs to. These so-called semantic segmentation tasks were usually tackled using standard classification CNN architectures. Each pixel is separately classified by providing a local region (also called patch) around the pixel to the classification network. Using a sliding-window approach, all pixels of an image are classified. There are however two drawbacks to this strategy. First, segmentation of an image is inefficient as many overlapping patches need to be propagated through the network. Second, finding the optimal patch size is difficult due to the trade-off between larger patches containing more context, and smaller patches for better localization.

To combine both context and good localization accuracy, the fully

convolutional network was introduced [41]. The idea is to add upsampling layers after the usual contracting classification network to increase the resolution of the output back to the input image resolution. No fully connected layers are used to preserve spatial information. To increase the output resolution, simple bilinear upsampling can be employed. Another approach is to use transposed convolutions, also called up- or deconvolutions, where the upsampling parameters are learned. The output size of the transposed convolution layer depends on the chosen kernel size and stride.

In the U-Net architecture this upsampling path is further extended with convolutional layers, allowing to propagate context information to the higher resolution layers [40]. This results in a more or less symmetric U-shaped architecture with a contracting and expansive path. This type of architecture is also called an encoder-decoder network. To improve localization, skip connections are added between the high resolution features of the encoder path and the upsampled feature maps in the decoder path. U-Nets efficiently use semantic and spatial information for accurate segmentation and are still the state-of-the-art for many segmentation tasks. They have since also found use in other image-to-image translation tasks, such as for example image denoising.

3.2.2 Deep learning for PET detectors

Positioning

In pixelated detectors, the interaction pixel is easily determined by centroid weighing methods such as Anger logic. More advanced techniques such as deep learning offer little advantage as the 2D resolution ultimately remains constrained by the pixel size. Improvements can however be made when it comes to obtaining DOI information normally not available in these detectors. A possible solution is the addition of a front- or lateral-sided readout, but the additional electronics increase costs and create additional dead-space between detector blocks. As an alternative, a linear method was developed for continuous DOI estimation based on scintillation light sharing through a common light guide on the front surface of the crystal [42]. This was later improved upon by replacing the linear method with a neural network estimator. The energies measured by the SiPM array are used as input features to predict a continuous DOI position. Both a dense neural network and a convolutional neural network (CNN) were tested, showing performance similar to each other

but improved by 12 - 26 % compared to the linear method. Uniformity was also significantly better throughout the crystal array [43].

Monolithic detectors on the other hand are attractive as these are not constrained by pixel size for spatial resolution and offer easier access to DOI information. Unfortunately, the aforementioned Anger logic no longer provides optimal spatial resolution. It particularly fails at the edges of the crystal due to non-linear light distributions and leads to incorrectly predicted interaction locations for non-perpendicular incidences. Early works have shown that neural networks could offer superior spatial resolution with good uniformity, and by providing training data at different incidence angles, could predict the interaction location for non-perpendicular incidences without the need to correct for DOI [44]–[46]. Later works included the DOI as an additional output, allowing for 3D positioning [47]–[49]. Training data can be obtained by the pencil-beam method or Monte Carlo simulation. The charge collected by the SiPM, possibly obtained in a row-column summing configuration, is used as the input to a dense neural network for predicting a 2D or 3D position. Some later studies replace the dense neural network with a CNN [50]. Performance is generally improved compared to other conventional methods, with better uniformity owing to higher spatial accuracy at the crystal edges.

Detector scattering

As a gamma photon passes through a scintillation crystal, it may undergo Compton or Rayleigh scattering before photoelectric absorption in another pixel or detector block. Rayleigh scattering, an inelastic process without energy transfer, is practically undetectable as no optical photons are generated. The elastic Compton interactions on the other hand convert a part of the gamma energy, proportional to the scattering angle, to scintillation light and reduce the energy available for subsequent photoelectric absorption. Such Compton scattered events are easily observed for interactions in different crystals or pixels, but identification of the first gamma interaction remains complicated, leading to image degradation due to incorrectly assigned LOR in PET. They are therefore often discarded, resulting in a loss of sensitivity.

In [51], a neural network approach was developed for assigning the LOR in PET for triple coincidences, where one 511 keV photoelectric event P coincides with two more singles S_1 and S_2, whose energy sum also equals 511 keV. In ideal circumstances, it is often possible to analytically

derive which single lies on the LOR by considering the relation between scattering angle and energy deposit. The limited energy resolution and positioning accuracy however degrade these analytical methods considerably, which the deep learning approach seeks to solve by inherently taking such limitations into account with realistic training data. The interaction coordinates of S_1 and S_2 are first re-defined in a plane w.r.t. P, so that a dense neural network with only 6 inputs (2D coordinates of S_1 and S_2, and their measured energies) can be trained to predict which of the two scattered singles lies on the LOR. Ground truth data is provided by means of Monte Carlo simulation. This method showed a LOR recovery rate of 75%, yielding a 55% sensitivity increase when including these triple coincidences on real data from the LabPET scanner [52]. It showed acceptable resolution degradation similar to other sensitivity increasing methods with little to no contrast loss.

Monolithic detectors additionally suffer from intra-crystal scatter degradation, but as scintillation light is not confined to pixels, it becomes difficult to discern scattered from non-scattered events, let alone assign separate positions and energies to subsequent interactions within the same crystal. Nonetheless, by training on realistic data which includes these scatters, it is possible to inherently take them into account in the positioning algorithm.

Timing

Timing estimation is traditionally done by recording the moment the SiPM signal crosses a pre-defined threshold. This however condenses all the potentially useful signal information into a single linear estimator. In [53], it was shown that convolutional networks could be used to predict the TOF difference directly from the detector signals themselves. The study used the outputs of two opposing detector pixels, digitized using 100 ps binning and then stacked side-by-side, as a single CNN input for predicting the TOF difference between both detectors. Only the short rising edges of the signals were used, as most of the important timing information is contained within the first few arriving scintillation photons. This method showed promising results, improving the coincidence time resolution (CTR) by 20% compared to leading edge detection and 23% compared to constant fraction discrimination.

Building upon this, we investigate the use and potential benefits of deep learning for time estimation in monolithic, rather than pixelated, detectors in Chapter 5.

3.2.3 Deep learning for image reconstruction

Iterative reconstruction methods, while certainly an improvement over analytical ones, are not without their own drawbacks. They are computationally expensive, may still include modeling errors in the system matrix $\mathbf{H}$, and the (potential) regularization term and its strength β ultimately involve user-specified assumptions about what are considered acceptable image properties. Deep learning based approaches seek to solve these limitations by replacing the uncertain user-defined variables in traditional methods with parameters learned from data.

Data-driven approaches

One option for deep learning image reconstruction is to replace equation 2.5 with a neural network F that takes into account all system properties and noise statistics so that:

$$\hat{\boldsymbol{\lambda}} = F(\boldsymbol{y}) \tag{3.2}$$

The network learns to directly reconstruct the image from projection data by training on known data pairs $\boldsymbol{\lambda}$ and $\boldsymbol{y}$. Convolutional encoder-decoder networks are typically used, having proven capable in various other image-to-image translation tasks [54]–[56]. These networks contain a contractive path, extracting (encoding) features from the input data, and an expansive path that constructs (decodes) the output from these features.

Once training is finished, reconstruction of new images is fast as it only requires a single forward pass through the network. These direct reconstruction methods are entirely data-driven, meaning the full inverse mapping is learned from training pairs without making any underlying assumptions about the imaging process itself. This limits modeling errors and allows the noise characteristics to be learned from data rather than being predefined by the regularization term. Learning such a complex relationship does require large amounts of training data, which can be difficult to obtain since the true image $\boldsymbol{x}$ is generally unknown to us. Simulated data with known $\boldsymbol{\lambda}$, or traditionally reconstructed images with low noise levels (e.g., high-dose images) for which $\hat{\boldsymbol{\lambda}} \sim \boldsymbol{\lambda}$ may be used instead.

One of the first algorithms developed for direct deep learning reconstruction in PET is DeepPET [57]. It uses a convolutional encoder-decoder architecture to reconstruct PET images from 2D sinograms

by training on simulated PET data, obtained from humanoid XCAT digital phantoms [58]. The network was later also used as the generator in a Wasserstein generative adversarial network (GAN) for improved reconstruction quality [59]. One common drawback of these algorithms is that fully 3D reconstruction is not possible with current GPU memory sizes, therefore remaining limited to 2D slice by slice reconstruction.

Model-driven approaches

Besides the large data requirements, the aforementioned approaches lack in interpretability given their black-box nature and concerns remain about the generalization capability for out-of-domain cases. Such limitations have lead to an increasing interest in physics-aware deep learning, where the neural network incorporates existing domain-knowledge prior to training.

As a concrete example from CT, it was shown that the FBP algorithm could be translated into a neural network by mapping each mathematical operation to a network layer [60]. For parallel beam geometry, the FBP algorithm can be written as:

$$\hat{\boldsymbol{x}} = \boldsymbol{H}^T \boldsymbol{C} \boldsymbol{y} \tag{3.3}$$

with $\hat{\boldsymbol{x}}$ the predicted image, $\boldsymbol{H}^T$ the back-projection operator and $\boldsymbol{C}$ the convolution of the projection data with a ramp-filter. When transforming this into a neural network with input $\boldsymbol{y}$ and output $\hat{\boldsymbol{x}}$, the first layer implements the operator $\boldsymbol{C}$, which is readily achieved by a convolutional layer with a single one-dimensional filter of size equal to the projection size. The following layer implements the operator $\boldsymbol{H}^T$ as a fully connected layer, but its weights are kept fixed due to memory constraints. Lastly, a ReLU activation function imposes the non-negativity constraint on the image data. This approach can be extended to fan-beam and cone-beam geometries by implementing additional element-wise weighting layers before the convolutional layer, and by translating the back-projection operator $\boldsymbol{H}^T$ to the appropriate geometry. The network weights are initialized to the values known from the analytical approach, so that prior to any training, a forward pass through the network is identical to the FBP algorithm. By training on known data pairs $\boldsymbol{x}$ and $\boldsymbol{y}$, the weights are then updated to include processes previously not accounted for in FBP. Only a small amount of training data can already provide reconstruction improvements due to the solid starting point offered by the weight initialization. Moreover, such a network offers easy interpretation

given the one-to-one mapping between analytical operations and network layers, and is less likely to give incorrect results for edge-cases due to the constraints imposed by the network architecture and fixed back-projection weights. The primary downside is that the network architecture and its number of learnable parameters may be too limiting to correctly model all imperfections and noise characteristics.

Similar to how it is done for FBP, iterative approaches can also be translated into a neural network through a process commonly referred to as algorithm unrolling or unfolding. This methodology was first proposed to improve the computational efficiency of sparse coding algorithms [61], but can be extended to the iterative methods used in medical imaging. The core idea of algorithm unrolling is to fix the number of iterations, map each update $\boldsymbol{\lambda}^{(k)} \rightarrow \boldsymbol{\lambda}^{(k+1)}$ to a block of network layers $F^{(k)}$, and stack these together to form an end-to-end mapping $\boldsymbol{y} \rightarrow \hat{\boldsymbol{\lambda}}$. Network parameters can then be optimized using data pairs $\boldsymbol{y}$ and $\boldsymbol{\lambda}$. The mathematical formulation and therefore network architecture of the iteration blocks $F^{(k)}$ depend on the iterative framework, but will contain terms relating to the data-consistency f and the regularization R. Parameters we are fairly confident in can be kept fixed (those relating to the data-consistency), while others we are less knowledgeable about should be learned (the regularization parameters). In contrast with the original iterative algorithm, each block $F^{(k)}$ and its corresponding step size can be different, and optimized with their own unique weights.

It should be noted that there is a fair amount of flexibility when it comes to how the regularization steps are implemented in the neural network. Certain studies opt to keep the original regularization update and simply make its parameters learnable [62], whereas others replace the entire regularization update with a more generic denoising CNN [63]. The latter can be seen as a middle ground between data-driven and model-driven approaches, combining aspects of both. Several studies from CT [64], [65], MRI [62], [63], [66]–[71] and PET [72], [73] have shown that unrolled algorithms can improve both computation speed and reconstruction quality compared to traditional iterative methods, while offering a robust and interpretable reconstruction procedure. For a more detailed overview, we refer to some review studies specific on the topic [74]–[76].

In Chapter 10, we implement a deep learning regularization technique to correct for limited angle artifacts: image artifacts arising during reconstruction due to limited projection angles for the dual-panel geometry of

the Walk-Through PET. We use a more traditional approach, where the regularization update is maintained, but the regularization component itself is replaced by a learned neural network (a 2D U-Net).

3.2.4 Deep learning for image restoration

One of the primary image degrading factors in medical imaging is noise arising from physical process randomness and scanner limitations, with possible artifacts produced by non-uniformity or incompleteness in the measurement data further reducing image quality. While the deep learning reconstruction methods discussed in Section 3.2.3 learn to correct for these effects through training data, no such corrections are included in analytical approaches. Even iterative algorithms that include noise suppression via the regularization term may still exhibit artifacts, or result in poor images when presented with limited measurement data. In these cases, deep learning can be used as a post-processing tool for restoring noisy or corrupted images. Common examples would be low-dose or limited-angle tomography scans, from which the matching high-dose or full angle acquisition is to be restored.

Supervised methods

Supervised image restoration requires training pairs of low-quality images $\hat{\boldsymbol{\lambda}}_L$ (containing artifacts or high noise levels), and high-quality images $\hat{\boldsymbol{\lambda}}_H$ (artifact free or low noise levels). A neural network F is then trained to map the low-quality image to its corresponding high-quality version.

$$\hat{\boldsymbol{\lambda}}_H = F(\hat{\boldsymbol{\lambda}}_L) \tag{3.4}$$

The procedure shares many similarities with the data-driven reconstruction methods in Section 3.2.3, but rather than the measurement data, the already reconstructed images are used as the input. This facilitates training as the network no longer has to learn the entire imaging process, and leads to reduced data needs for good network performance. Simulations or experiments can provide the training targets $\hat{\boldsymbol{\lambda}}_H$, from which the corresponding inputs $\hat{\boldsymbol{\lambda}}_L$ are easily obtained by removing a subset of measurement data or by introducing artificial noise prior to reconstruction. A variety of network architectures can be used for F, of which a few examples will be discussed.

One of the simplest architectures conceivable for this task are the 3-layer deep CNN used for limited angle CT artifact removal [77] or for

low-dose CT denoising [78]. The limited angle CT network uses a full image obtained by FBP as input, and removes the directional artifacts arising from the removed angles. The low-dose CT denoising network instead opts to work on patches of the image. One advantage of using patches is that many can be extracted from a single image, leading to a much larger training dataset. Additionally, if the patches are small enough, 3D convolutional networks become viable due to the reduced memory requirements, although this particular network opted for 2D convolutions. A disadvantage of using patches is the loss of long-range spatial information, which could play an important role depending on the specific noise generation procedure. Streak artifacts produced by limited angle tomography propagate throughout the whole image, whereas the noise present in low-dose scans remains more local. In both networks, all three layers are implemented as a convolution, with the first two using a ReLU activation for non-linearity. Each layer corresponds to a specific mathematical operation: the first performs feature extraction, the second applies a non-linear mapping suppressing those features corresponding to artifacts or noise, and the final layer re-combines them into a new image. These networks have the advantage of being interpretable, but may be too constraining compared to more general, deeper networks.

The encoder-decoder design used for direct image reconstruction can again be used for image restoration, although in this case skip-connections are usually added between the layers, resulting in the well-known U-Net [40] architecture seen in Figure 3.3. The skip-connections are essentially a copy-paste-concatenate operation where the output of early layers in the network are added to the later layers. They allow high-level features to be re-used later on by bypassing other layers, thereby improving training convergence and performance. While the U-Net architecture was originally used and continues to be used for image segmentation tasks, it is nowadays also one of the more prominent network architectures in image restoration. An additional modification that is often added to the U-Net for image restoration is a residual connection between input and output. Given the structural similarity between $\hat{\boldsymbol{\lambda}}_L$ and $\hat{\boldsymbol{\lambda}}_H$, the network essentially needs to learn the identity mapping as a part of the image restoration procedure. Therefore, a residual connection is often employed (which simply adds the input to the output), so that the network only has to learn the residual noise $\hat{\boldsymbol{\lambda}}_{noise} = \hat{\boldsymbol{\lambda}}_L - \hat{\boldsymbol{\lambda}}_H$, rather than directly generating $\hat{\boldsymbol{\lambda}}_H$. This methodology was first proposed as a general image denoising method [79], and quickly found its way to

medical imaging. Now the network only needs to find the perturbations with reference to the identity transform, a generally easier task. Such a small change can lead to large improvements in convergence and training data needs. These U-Net based networks have been used to great success in sparse-view CT [80], low-dose CT [81] and converting low-count to high-count PET images [82], [83].

Another possible network architecture is based on ResNet [84], where rather than employing an encoder-decoder style network with symmetrical skip-connections, many residual blocks, where the output of each block is summed with its input, are appended one after another. A variant of ResNet has for example been used for denoising PET images in [85].

The aforementioned networks can also be adapted for other types of inputs and outputs. Some studies on limited-angle tomography for example choose to perform image restoration in sinogram space ($\hat{\boldsymbol{y}}_L \rightarrow \hat{\boldsymbol{y}}_H$) prior to image reconstruction [86], [87], although both options are compared for partial-ring PET in [88], showing better results using image space data. Alternatively, dual imaging modalities such as PET/MRI may use the MRI scan as an additional input to provide anatomical information, helping with the denoising of the PET scan [89]. The relative weight that should be given to both inputs is automatically derived during the training procedure, without any need for manual tuning. It is also possible to use multiple sequential image slices as input, where each slice is a different channel, in order to incorporate some spatial information along the third dimension without resorting to 3D CNNs.

Unsupervised methods

Most unsupervised image restoration methods are derived from the deep image prior proposed in [90], which can be used for common tasks such as denoising, super-resolution, inpainting, etc. The authors showed that a randomly initialized CNN can itself serve as a prior for image restoration, by treating the low-quality images as training labels. In this framework, a convolutional network F is trained to produce the noisy scan data $\hat{\boldsymbol{\lambda}}_L$ from a random input vector or image $\boldsymbol{z}$:

$$\hat{\boldsymbol{x}}_L = F(\boldsymbol{z}) \tag{3.5}$$

As the number of training iterations increase, the output approaches the noisy image $\hat{\boldsymbol{\lambda}}_L$. It is however observed that the optimization procedure leads us through a path for which, prior to reaching final convergence,

the network outputs a restored version of $\hat{\boldsymbol{\lambda}}_L$, so that $F(\boldsymbol{z}) \sim \hat{\boldsymbol{\lambda}}_H$. The authors suggested that this phenomenon likely emerges due to convolutional operations imposing self-similarity on the generated images, making it easier for the networks to learn meaningful signals rather than noise. In other words, it is possible to stop training at a point where the network has more or less learned the signal but has yet to learn the noise present in $\hat{\boldsymbol{\lambda}}_L$. The U-Net like architectures are a particularly good choice for F, since the skip-connections allow to impose this self-similarity at various feature scales. We emphasize that this method requires a separate network F to be trained for each distinct image. In practice, the random input $\boldsymbol{z}$ is usually replaced with a prior image containing additional information, such as the CT or MRI image for hybrid PET/CT or PET/MRI denoising [91]. A similar approach is taken in dynamic PET imaging, where the time-aggregated scan can be used as the input for denoising separate dynamic slices [92].

Besides its use as a post-processing tool, the deep image prior can also be incorporated into the iterative image reconstruction procedure as a replacement to traditional regularization schemes [93], [94]. During each update step, the network is trained to generate the current image estimate $\boldsymbol{\lambda}^{(k)}$ from a prior image $\boldsymbol{z}$, thereby performing a denoising step between each update. This methodology is different from the unrolled algorithms discussed in Section 3.2.3, as it still makes use of traditional iterative optimization steps rather than providing a single network used for end-to-end reconstruction. But, compared to image restoration as a post-processing step, such an integrated approach has the advantage of ensuring data-consistency on the final denoised image.

Not requiring any training data naturally offers a significant benefit, as it essentially solves one of the main difficulties in constructing good machine learning models. A downside of the deep image prior is however its need to be separately trained for each image, making the process rather slow in comparison to supervised approaches, which can use a single pre-trained network for all images. Performance is also unlikely to match that of supervised algorithms trained for a specific noise level, but the flexibility of unsupervised algorithms nonetheless makes them an attractive option when supervised algorithms are unfeasible.

Part I

Timing capabilities of monolithic detectors

Chapter 4

Cerenkov time estimation in monolithic BGO

4.1 Introduction

Since the advent of TOF-PET, BGO based PET detectors are no longer frequently used due to their slow scintillation decay time, leading to prohibitively low TOF resolutions. Most current PET scanners are therefore based on L(Y)SO scintillation material given its superior scintillation characteristics compared to BGO, both in terms of light yield and decay time. BGO however recently had a resurgence in popularity due its lower cost and Cerenkov based time estimation [95], made possible by improvements in silicon photomultipliers (SiPMs) and readout electronics. Cerenkov emission is nearly instant, but has a very low light yield. Only 17 photons are emitted on average per 511 keV event in pixelated BGO detectors [96]. Therefore, improvements to photon detection efficiency, especially at shorter wavelengths where the majority of Cerenkov light is emitted, combined with the capability to detect individual photons at good time resolutions, played a big role in enabling time-of-flight capabilities for BGO.

Coincidence time resolution spectra of Cerenkov based PET detectors are however non-Gaussian in nature, with long tails produced by the statistical fluctuations in the number of detected Cerenkov photons. Image reconstruction may be improved by the use of multiple time-of-flight kernels, where each event is assigned a specific kernel based on its estimated time resolution (or estimated number of Cerenkov detections), significantly improving the image signal-to-noise ratio [97].

Characterizing events based on their time resolution can for example be achieved by sorting events by their signal rise time [98]. Studies on Cerenkov based time estimation in BGO have up until now however primarily been focussed on pixelated detectors.

There has also been a push towards monolithic PET detectors as they provide good spatial resolution using only a limited amount of electronic channels (compared to pixelated detectors of similar spatial resolution) and they intrinsically offer depth-of-interaction (DOI) information [99].

The spatial resolution of monolithic detectors has been optimized quite well, providing especially good positioning performance when combined with deep learning, down to the 1 mm range in full width at half maximum (FWHM) [100], [101]. Further improvements are unlikely to result in significant enhancement of image quality, as at this point the photon acollinearity and positron range start to become the dominant factors in LOR positioning. Enhancements to time resolution on the other hand lead to an increase in effective sensitivity and therefore signal-to-noise ratio of the reconstructed scans, providing more room for improvement.

In order to optimize the timing resolution, we should pay attention to the choice of surface finishes for the scintillator, as it can have a large impact on the photon transport characteristics and therefore affect the photon transfer times in the crystal [102]. In monolithic detectors, which have until now mostly been used in preclinical PET where TOF only becomes relevant below 100 ps, these surfaces are often chosen to maximize spatial resolution. E.g. by using absorbing side surfaces to minimize reflections and thereby reduce edge effects. Optimizing for spatial resolution may however have an adverse effect on the time and energy resolution, by reducing the overall light collection efficiency. This is especially important for Cerenkov based time estimation, where we have to rely on the detection of a very small amount of photons for timing.

In this study, we perform a series of Monte Carlo simulations to better understand and evaluate the effect of detector geometry and scintillator surface finish on the photon collection efficiency and time resolution in monolithic BGO detectors. We additionally compare to monolithic LYSO, as well as pixelated BGO and LYSO detectors.

4.2 Materials and Methods

We use GATE v9.2 [103], based on the Geant4 toolkit [27], to model gamma interactions in PET detectors and the subsequent production and transport of optical scintillation and Cerenkov light. We investigate both BGO and LYSO as scintillation material, with a variety of detector geometries (both monolithic and pixelated) and surface finishes.

The detector sizes under consideration are 50x50x16 mm^3, 50x50x12 mm^3, 25x25x16 mm^3 and 25x25x12 mm^3 for the monoliths. We choose the 50x50x16 mm^3 size as a starting point since it has the same aspect ratio as the smaller 25x25x8 mm^3 monolithic detectors already proven and currently in use in certain preclinical systems [104]. It is also used in other monolithic detector studies [100]. A 12 mm thick version is additionally examined since making the switch from LYSO to BGO should allow for thinner detectors given the higher stopping power of BGO. The 25x25 mm^2 detectors are considered as an alternative option, as these smaller detectors would help with the count rate, especially important in total-body PET systems.

For each distinct detector configuration, we simulate 10 000 gamma photons impinging perpendicularly on the front surface of the detector, with entry points spread uniformly over the surface. Any generated optical photons are subsequently tracked, and those that are transmitted through the back surface (photodetector side) are recorded for further analysis.

The optical and scintillation properties of LYSO and BGO (Materials.xml file in GATE) are based on the data sheets provided by Epic Crystal [18], [19], see Table 4.1. The wavelength dependency of the scintillation spectrum is taken into account for both scintillators, as is the wavelength dependency of the refractive index for BGO. The refractive index for LYSO was chosen constant as it shows little change over optical energy ranges (1 eV - 5 eV). Surface reflections of optical photons are modeled by the LUT Davis Model [105], using custom look-up tables to represent different surface finishes.

4.2.1 Cerenkov production

Cerenkov light is produced when a charged particle, in this case a recoil electron ejected from its atom due to the energy deposited by the gamma photon, moves faster than the speed of light in a specific medium, see Figure 4.1. The emission of Cerenkov light is directional, with photons

	unit	LYSO	BGO
scintillation yield	MeV^{-1}	29 000	8 500
energy resolution	%	10.9	11.9
rise time	ps	70	70[a]
decay time	ns	42	317
emission peak wavelength	nm	420	480
refractive index		1.82	2.15[b]

Table 4.1: Material properties of LYSO and BGO used in the GATE simulations. The energy dependency of the scintillation spectra is taken into account for both materials, as is the wavelength dependency of the refractive index for BGO.
[a]Value was not provided in the datasheet, so it was chosen equal to LYSO for the comparison. It is nonetheless in line with experimental results [95].
[b]At the emission peak wavelength.

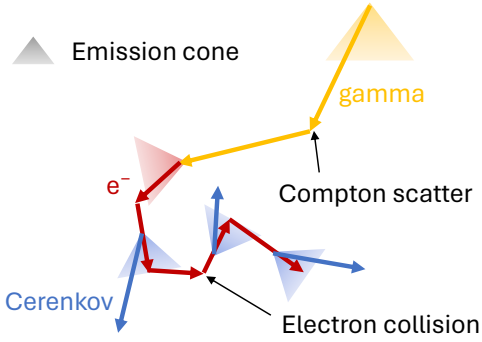


Figure 4.1: Cerenkov production process in scintillators.

emitted in a cone along the electron path. Despite this, Cerenkov photons emitted following a gamma interaction show relatively little correlation with the incoming gamma velocity vector. The electron itself is emitted in a cone along the direction of the incoming gamma photon, effectively increasing the maximum angle of Cerenkov emission. Furthermore, as the electron moves through the material, it constantly loses energy and changes direction due to collisions with other particles. This results in all but the first few Cerenkov photons to be emitted in fairly random directions relative to the gamma photon.

How quickly the electron loses "directionality" is affected by the mean distance between collisions, also called the mean free path. Correct

modeling of the electron path in Geant4 is therefore crucial for simulations regarding Cerenkov production. The default settings of GATE/Geant4 however lead to an overestimation of the mean free path, thereby resulting in overly focused Cerenkov photon emission angles [106]. The mean step length of particles in Geant4 can be controlled by the $\Delta\beta$ parameter, which puts a limit on how much the kinetic energy of a particle is allowed to change during each step. We use $\Delta\beta$=0.02, resulting in a mean step length of ≈ 0.150 μm for 450 keV electrons in LYSO. This is close to the values of the mean free path of electrons in elemental solids with similar effective atomic number as LYSO [107], and is also an appropriate value to use for BGO [106]. Note that we are checking 450 keV, rather than 511 keV, since the ionization energy ($\sim$ 63 keV for lutetium) should be subtracted to obtain the initial energy of the recoil electron.

Additionally, it should be taken into account that Geant4 does not automatically limit the Cerenkov emission spectrum to reasonable energy ranges, instead creating photons over the full energy range for which the refractive index n is specified [108]. We therefore restrict the refractive index to transparent and physical energy regions for Cerenkov emission, where we use 310 - 850 nm for BGO and 390 - 750 nm for LYSO [96].

4.2.2 LUT Davis Model

The LUT Davis Model [105] allows to more accurately simulate reflection and transmission of optical photons at the interface between two materials by making use of measured 3D surface topographies. Given a surface sample, the model generates look-up tables (LUTs) containing the angular distribution of reflectance/transmittance, as well as the angular distribution of reflected and transmitted photons as a function of incidence angle. This is especially important when considering rough surfaces, where the reflectance/transmittance deviates greatly from the Fresnel equations, which are based on perfectly flat surfaces. The LUT Davis Model app [109] allows users to generate their own LUTs for use in GATE, by specifying the scintillation material (LYSO or BGO in our case), the surface finish (both a pre-existing polished and rough finish are available), the index of refraction of the coupling material (e.g. 1 for air or 1.5 for optical grease) and an optional reflector such as ESR. It should be noted here that the wavelength dependency of ESR is not taken into account, which may be non-negligible for Cerenkov emission due to the rapid drop-off in reflectivity below 370 nm. This is a limitation in the LUT Davis model at the time of performing this study.

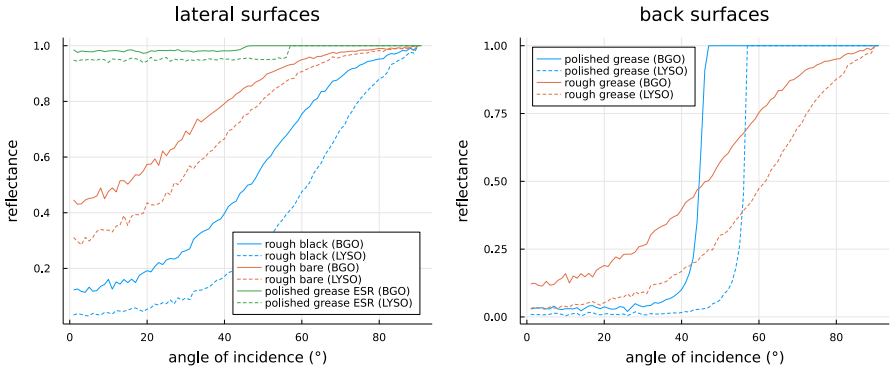


Figure 4.2: Angular distribution of reflectance of the surface look-up tables used throughout our simulations.

The different detectors considered in this study all use a polished front surface (opposite of the photodetector) coupled with optical grease to a specular reflector. We test three different lateral surfaces. The first is the exact same as the front surface (polished grease ESR), which is often found in pixelated detectors to maximize internal reflection. The second is a rough surface coated with black paint (rough black) to minimize reflections. It is modeled as a rough surface coupled with a material of refractive index 1.5 to a perfectly absorbing "reflector". This type of surface is often used in monolithic detectors to minimize side reflections and therefore reduce edge effects. The third lateral surface is a bare rough surface, so a rough surface coupled to air (rough bare). Finally, for the back surface we test both a rough and polished surface, both of them coupled with optical grease to the photodetector. Figure 4.2 shows the angular distribution of reflectance for the obtained LUTs.

4.2.3 Photon detection efficiency

As shown in Figure 4.2, not all photons reaching the back surface will be transmitted through to the actual SiPM surface. This results in the "rejection" of a certain percentage of optical photons prior to any non-idealities unique to the photodetector itself. Of those transmitted photons, only a certain percentage will actually trigger an avalanche and therefore be detected, which is determined by the photon detection efficiency (PDE) of the SiPM in question.

In our simulations, we additionally take into account this (energy dependent) PDE of the SiPMs, for which we base ourselves on the

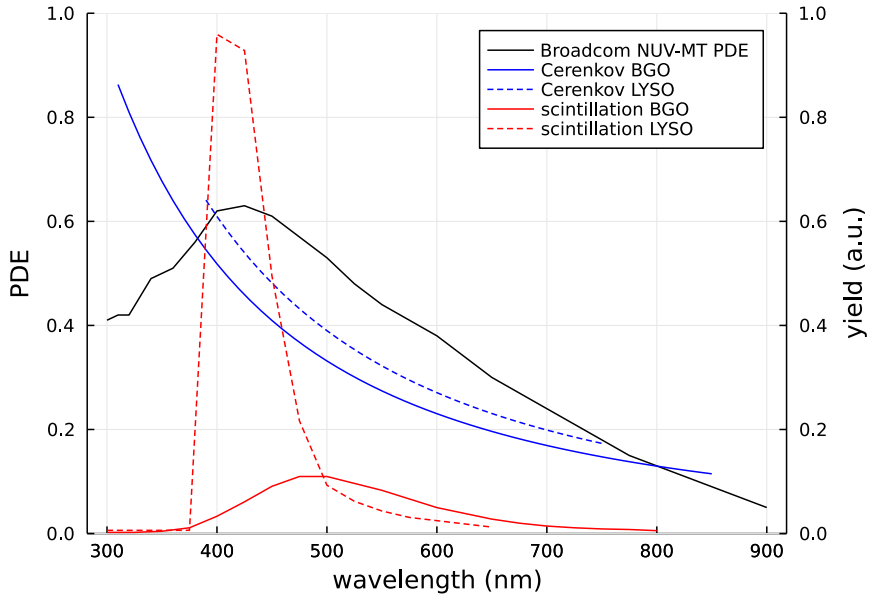


Figure 4.3: Photon detection efficiency (PDE) wavelength dependency of Broadcom NUV-MT SiPMs and Cerenkov/scintillation spectra of BGO and LYSO.

Broadcom NUV-MT SiPMs [110]. These have a maximum PDE of 63% at a wavelength of 420 nm, and maintain over 40% PDE down to 300 nm. This makes them a good choice for Cerenkov photon detection, given that the Cerenkov emission spectrum scales with $1/\lambda^2$ and is therefore concentrated at lower wavelengths. Figure 4.3 shows the PDE overlaid with the Cerenkov and scintillation spectra of BGO and LYSO.

4.2.4 SiPM waveform simulation

As a final comparison, we will also estimate the coincidence time resolution of the different detector configurations by simulating the SiPM signals and predicting the gamma arrival time based on leading edge discrimination. The methodology for simulating the SiPM signals is inspired by [20], with simulation parameters again based on the Broadcom NUV-MT SiPMs [110].

The exact photon timestamps (both Scintillation and Cerenkov) as obtained from GATE are first assigned to their corresponding SiPM, randomly removing photons to account for the limited photon detection

efficiency (wavelength dependent as before). In addition, dark counts are generated by sampling from a Poisson process at a rate of 4.4 MHz per SiPM. A photodetector transit time spread is then modeled by convolving the given timestamps with a Gaussian, where the standard deviation corresponds to the intrinsic single photon time resolution (SPTR) of the SiPM, calculated prior to the inclusion of electronic noise. A value of σ = 30 ps was used, estimated based on intrinsic SPTR measurements of similar SiPMs [96].

Next, prompt optical crosstalk with a certain probability (p = 23%) gives rise to duplicate counts, each of which can again result in crosstalk with the same probability. The i-th crosstalk event receives a time smearing $\sigma = SPTR\sqrt{i+1}$ with an additional time delay $\mu = SPTR\sqrt{i}$ [20]. The SiPM signal $s(t)$ is then generated as a sum of tri-exponential functions centered around the photon detection times t_i, with rise time τ_{rise} = 100 ps, fast decay time τ_{fast} = 5 ns and slow decay time τ_{slow} = 55 ns:

$$s(t) = \sum_i \left(C_{fast} \exp\left[\frac{t_i - t}{\tau_{fast}}\right] + (1 - C_{fast}) \exp\left[\frac{t_i - t}{\tau_{slow}}\right] - \exp\left[\frac{t_i - t}{\tau_{rise}}\right] \right) H(t - t_i) \tag{4.1}$$

The Heaviside function H sets the signal to zero prior to photon detection and C_{fast} ($\in$ [0, 1], here 1/3) denotes the relative strength of the fast component of the decay time. The signal is then passed through a first-order low-pass Butterworth filter to simulate the limited bandwidth of the readout electronics. The low-pass Butterworth filter has a near-flat frequency response in the passband (below the cut-off frequency), whereas signals in the stopband (above the cut-off frequency) are attenuated, with increased attenuation for increasing frequency and higher orders of the filter. A cut-off frequency of 500 MHz was used, as is for example the case for the HRFlexToT ASIC, a readout ASIC compatible with monolithic detectors [111]. Finally, electronic noise is added to the signal as zero-mean white Gaussian noise with σ = 5% of a single photoelectron pulse amplitude.

4.3 Results

4.3.1 Photon collection efficiencies

Figure 4.4 shows the photon collection efficiency (number of optical photons detected per event) of BGO and LYSO for two detector geometries, 50x50x12 mm^3 monolithic and 3x3x20 mm^3 pixelated, with different lateral and back surface finishes. Only events with full (511 keV) energy deposit in the crystal are considered, including both scatter + photoelectric and purely photoelectric events. Note that the PDE of the SiPMs is already included in these results.

As was observed in Figure 4.2, a rough back surface allows detection of more high incidence angle photons, at the cost of reduced transmission at lower incidence angles. Since scintillation photons are emitted isotropically in 3 dimensions, the angular distribution relative to the photodetector surface normal is proportional to the radius of the corresponding circle on the unit sphere. This results in more emission of high incidence angle photons. For monolithic detectors, or when using reflective side surfaces, the rough back surface results in a net gain of photon collection efficiency since many of these high angle photons do in fact make it all the way to the photodetector surface. For the monolithic BGO detector with reflective side surfaces, the rough back surface leads to a 34% increase in photon collection efficiency compared to a polished back surface. We also observe that the differences are larger in BGO than in LYSO, since the higher index of refraction of BGO (2.15 compared to 1.82 for LYSO) leads to a larger mismatch with the optical grease.

The same trends can be seen for the Cerenkov photons specifically as well. The combination of the short mean free path of the electron, the fact that the recoil electron itself will not be emitted in the exact same direction as the incoming gamma photon, and the possibility of the gamma photon itself having been scattered, results in many high incidence angle Cerenkov photons. A median of only 2 Cerenkov photons is detected per event for a rough black lateral surface with a polished back surface, whereas using a polished reflective lateral surface with a rough back surface increases this to a median of 5 Cerenkov photons.

Figure 4.5 shows the same results but for different monolithic BGO detector sizes. The overall trends remain the same, with the larger and thinner detectors generally showing higher photon collection efficiencies. For these detectors, there is a smaller probability of losing photons to the lateral surfaces.

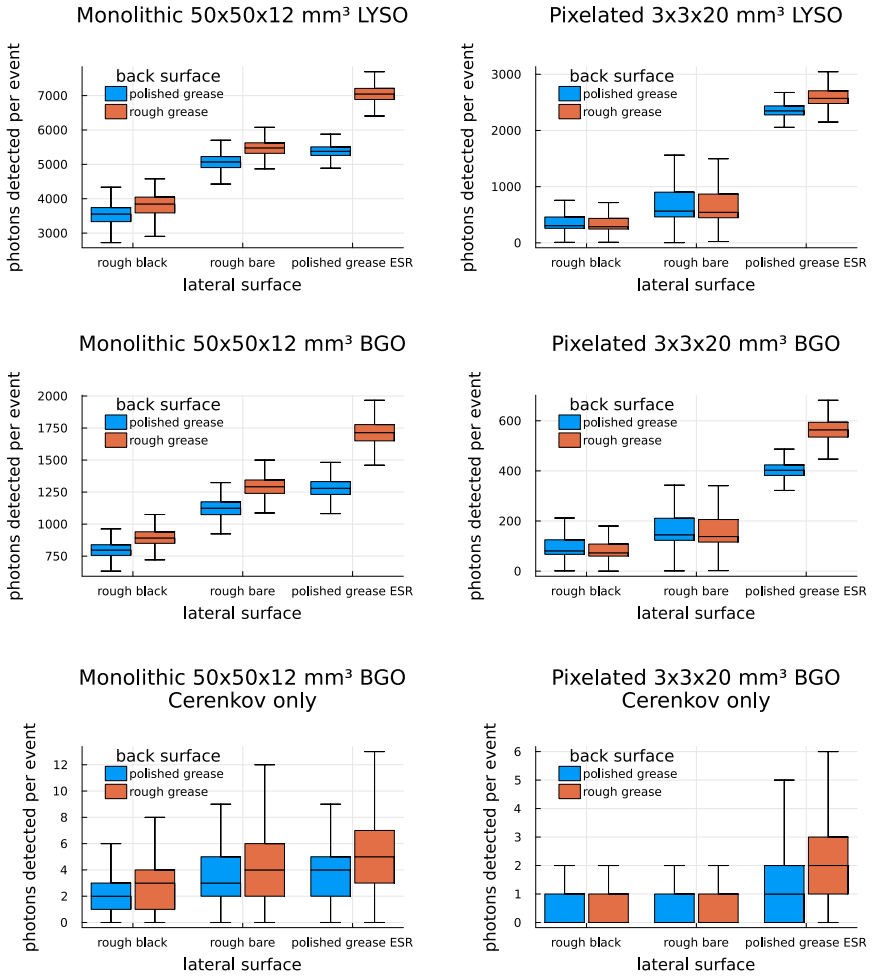


Figure 4.4: Comparison of photon collection efficiency for different surface finishes in monolithic and pixelated BGO and LYSO detectors. The top two rows consider both scintillation and Cerenkov photons, whereas the bottom row looks only at Cerenkov photons in BGO.

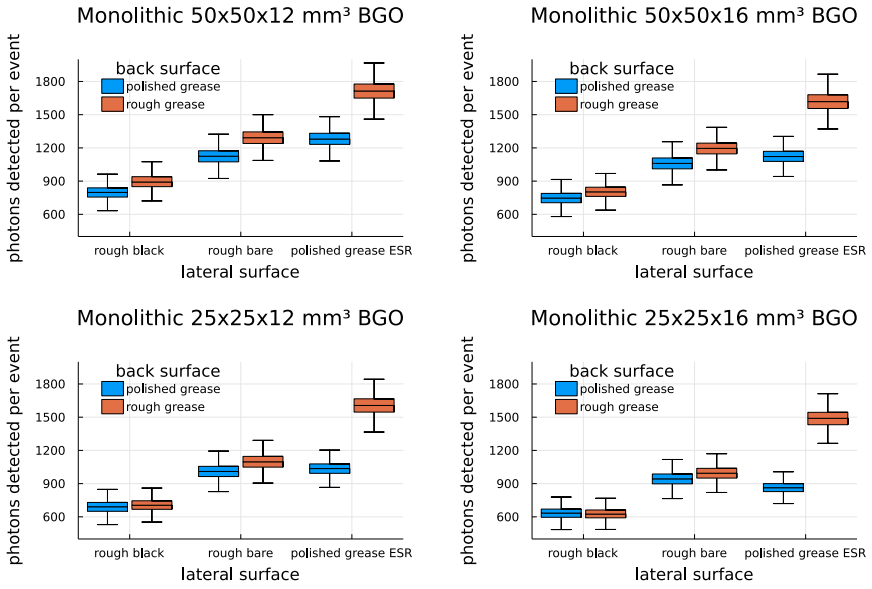


Figure 4.5: Effect of surface finish on photon collection efficiency in different monolithic BGO detector sizes.

4.3.2 Photon detection delays

In Figure 4.6 we take a look at the photon detection delays, which is the time between the gamma photon passing through the front surface of the detector and the detection of the n-th optical photon. A higher photon collection efficiency generally results in earlier photon detections, most obvious for the pixelated detectors, where otherwise many early photons are lost through the sides of the detector. On the other hand, when there were many direct photons initially detected, as is the case for monolithic LYSO, further enhancing the photon collection efficiency has little to no effect. For monolithic BGO, we do observe reductions in photon detection delays when improving the collection efficiency of reflected and/or high angle-of-incidence photons. While these photons possess longer transit times, if they were an early scintillation or Cerenkov photon, they will still be detected relatively fast compared to later emitted direct scintillation photons, due to the slow scintillation pulse of BGO.

The absolute values of the photon detection delays however do not have any direct impact on the coincidence time resolution (CTR), since these cancel out for two identical detectors in coincidence. More impor-

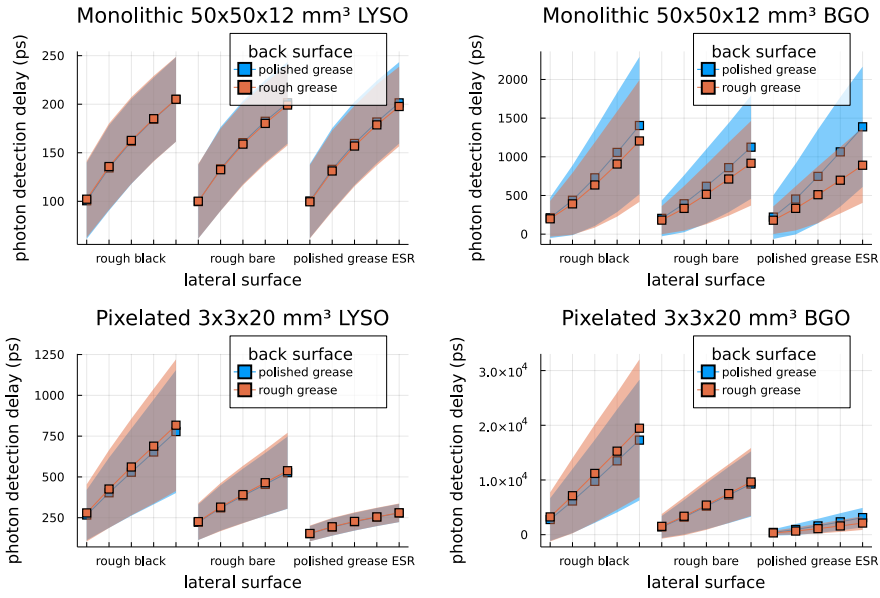


Figure 4.6: Time delay between the gamma photon passing through the front surface of the detector and the detection of the n-th photon (increasing left to right from n=1 to n=5). Values are averaged over all events of a specific configuration, with confidence intervals showing the standard deviation ($\pm\sigma$).

tant is the spread on the detection delays, especially in those cases where we rely on single photon detection per SiPM, as is the case for monolithic BGO. We see here that the rough back surface results in earlier photon detections with less variation, with the best results being obtained when combined with reflective lateral surfaces.

How fast the photons are coming in one after another (i.e. the slope in Figure 4.6) may also have an effect on the CTR. This is primarily relevant when there are sufficient detections per SiPM for the signals to quickly pile up (e.g. for pixelated detectors), increasing the slope dV/dt of the SiPM signal and therefore reducing noise on the leading edge discrimination [112].

4.3.3 CTR estimations

We estimate the CTR of the monolithic detectors by simulating the SiPM signals and predicting the gamma arrival time based on leading edge

discrimination. For the 50x50 mm^2 detectors we use an 8x8 readout of 6x6 mm^2 SiPMs, and for the 25x25 mm^2 detectors a 4x4 readout of the same 6x6 mm^2 SiPMs. Leading edge discrimination is performed on each individual SiPM waveform, resulting in a matrix of 8x8 or 4x4 timestamps. We use a leading edge discrimination threshold of 0.5 photoelectron pulse amplitudes in order to detect individual Cerenkov photons. Due to the presence of dark counts, the signal amplitude prior to gamma detection is not necessarily centered around 0 and may even be above the 0.5 photoelectron threshold level. This leads to a fraction of the SiPMs recently having triggered prior to the actual gamma event, resulting in 'dead' SiPMs incapable of triggering again for a certain time. We therefore lose the timing information of those SiPMs. In our simulations this limitation was modelled by considering those SiPMs with a signal amplitude > 0.25 photoelectrons prior to the gamma event as dead, and therefore not being present in the timestamp matrix.

The gamma arrival time itself was then predicted as the first SiPM timestamp. Averaging of the first few timestamps was also tested, but consistently underperformed for monolithic BGO. The TOF kernels are constructed by randomly subtracting different events from one another (to obtain events in coincidence), after which a kernel density estimation fit is performed to obtain a distribution from the TOF histogram. Figure 4.7 shows such a TOF kernel for a case with long tails (monolithic 50x50x16 mm^3 BGO with rough black sides and a polished detector surface). The kernel density estimation fit is compared to a Gaussian with the same FWHM, showing the non-Gaussian nature of the TOF kernel.

Figure 4.8 shows the coincidence time resolutions for the different monolithic detector configurations. Here we have used a leading edge discrimination threshold of 0.5 photoelectron pulse amplitudes. Since the distributions are not perfectly Gaussian, we report the full width at half maximum (FWHM), full width at tenth maximum (FWTM) and the full width at twentieth maximum (FW20M). While the FWHM remains in the same range for the different configurations, the FWTM and FW20M show larger differences. That is, the different surface finishes mostly affect the tails of the distribution. The thinner (12 mm) crystals provide a better time resolution, with the rough back surface resulting in lower FWTMs and FW20Ms (shorter tails), especially when using non-absorbing lateral surfaces. E.g. for monolithic 50x50x12 mm^3 BGO with reflective side surfaces, we see a reduction of 18% in FWTM going

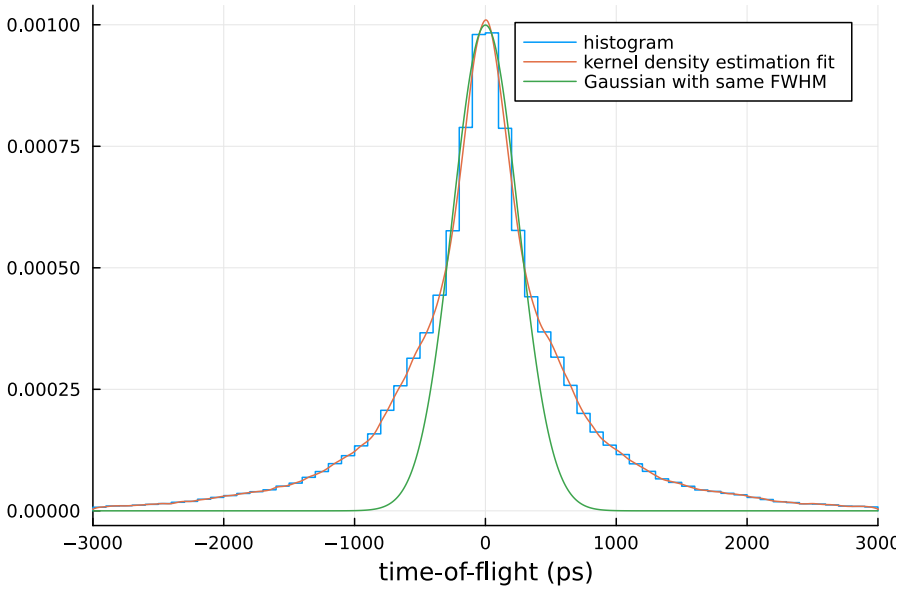


Figure 4.7: Time-of-flight kernel of monolithic 50x50x16 mm^3 BGO with rough black sides and a polished detector surface. The kernel density estimation fit, which accurately describes the time-of-flight histogram, is compared to a Gaussian kernel with the same FWHM, showing the extent of the tails.

from a polished back surface to a rough back surface. Again we also see that for monolithic LYSO, the surface finish has negligible impact on the coincidence time resolution.

In addition, we perform the same analysis as before, but now with a leading edge discrimination threshold of 1.5 photoelectron pulse amplitudes. As this is above the single photoelectron level, it is a more realistic triggering setting to properly handle rejection of dark counts. The results are shown in Figure 4.9. As expected, the CTR values are increased, especially for the monolithic BGO configurations, which are now all above 1000 ps FWHM (essentially, non-TOF capable). The relative trends between the different configurations remain largely unchanged, with again the largest differences notable in the tails of the distributions.

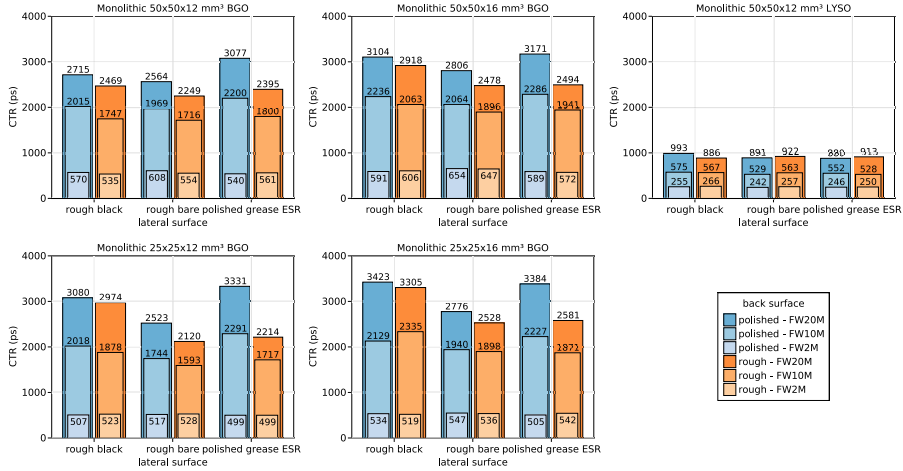


Figure 4.8: Coincidence time resolution obtained by leading edge discrimination at the 0.5 photoelectron level for the different detector configurations, reported in FWHM (FW2M), FWTM (FW10M) and FW20M.

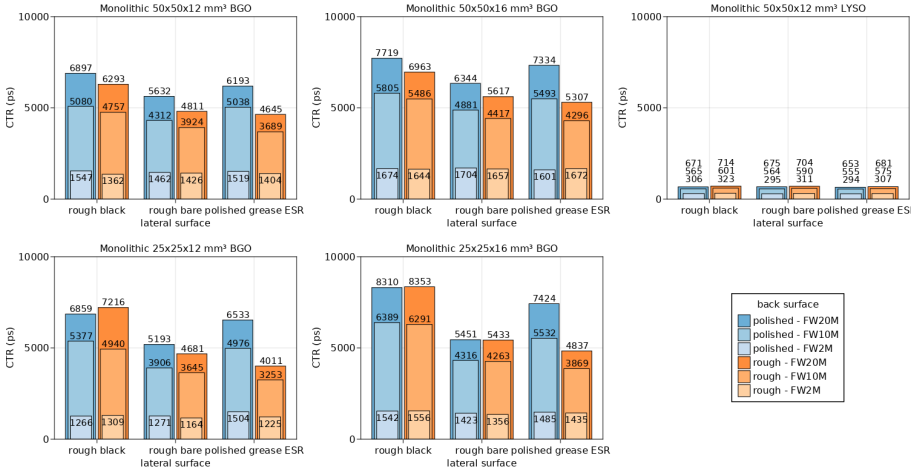


Figure 4.9: Coincidence time resolution obtained by leading edge discrimination at the 1.5 photoelectron level for the different detector configurations, reported in FWHM (FW2M), FWTM (FW10M) and FW20M.

4.4 Discussion

In order to benefit from Cerenkov based time estimation, it is important to use SiPMs and readout electronics that are in fact capable of resolving individual photon detections. Note that only a few Cerenkov photons are detected per event for an entire 50x50x12 mm^3 monolithic BGO detector (Figure 4.4), over a total of 64 SiPMs. Given the large spread in Cerenkov emission angles, it is highly improbable that more than one Cerenkov photon will be detected by the same SiPM, and we can therefore not rely on fast signal pile-up. This is reaffirmed by the results in Figure 4.9, showing that triggering above the single photoelectron level leads to a severe degradation of TOF. Modern SiPMs can easily detect individual photons, and readout electronics have no problem detecting the leading edge of an SiPM signal below a single photoelectron level. The difficulty lies in the generation of dark counts, which can lead to missing Cerenkov photon detections or dark counts being misinterpreted as a Cerenkov photon. This contaminates the timestamp matrix with noise or missing data, making accurate gamma interaction time estimation more difficult.

4.4.1 Effect of dark counts

In order to detect (sufficient) Cerenkov photons in monolithic detectors, it is required to set the leading edge threshold below a single photoelectron level. The SiPMs will rarely trigger on Cerenkov photons at higher thresholds, since most SiPMs detect no more than a single Cerenkov photon. While the dark count rate of a single SiPM is not that high (4.4 MHz), we are dealing with 64 SiPMs for the 50x50mm^2 detector configurations. This means that on average, we detect a dark count every $\sim$3.6 ns over the whole detector, which is only an order of magnitude larger than the typical coincidence time resolutions.

An SiPM that is triggered by a dark count will be incapable of triggering again for a short time while the signal amplitude drops off. If a Cerenkov photon happens to be absorbed during this time frame, it will still contribute to the SiPM signal, but no leading edge detection will trigger. These 'dead' SiPMs essentially equate to a loss of photon collection efficiency when it comes to timing information. In addition, while most dark count triggers can be rejected based on energy integration, measuring a dark count every $\sim$3.6 ns means that there is a non-negligible probability that a dark count occurs just prior to the gamma event. It would therefore be indistinguishable from a Cerenkov photon, leading to

a false data point in the timestamp matrix.

Dark counts are likely less problematic in pixelated detectors, since the overall dark count rate is lower (only one SiPM per pixel) and all Cerenkov photons are detected by the same SiPM, so that a higher leading edge threshold would still be capable of detecting Cerenkov photons. This is still partially true of monolithic LYSO, where many scintillation photons reach the SiPMs sufficiently quick one after another to still obtain valuable time information using thresholds above the single photoelectron level. Therefore, further reductions in dark count generation of SiPMs, and readout electronics better capable of dealing with dark count rejection are especially important for timing in monolithic BGO.

4.4.2 Effect of photon collection efficiency

The photon collection efficiency of the detectors plays a significant role in timing performance when relying on Cerenkov photons for time-of-flight estimation. We can appreciate this in Figure 4.8, consistently showing better time resolutions for the higher photon collection efficiencies. The detectors with higher photon collection efficiencies do detect more high incidence angle (for the rough back surface) or reflected (for the reflective lateral surfaces) photons. These photons have longer (and higher variation in) photon transfer times from the gamma interaction position to the photodetector. Therefore, accurate prediction of the transfer time of an individual photon is more difficult. Nonetheless, these detectors show better overall timing statistics due to the larger amount of photon detections.

4.4.3 Time-of-flight kernels

As mentioned, the photon collection efficiency in monolithic BGO primarily has an effect on the tails of the time-of-flight kernel, showing shorter tails for higher collection efficiencies. Figure 4.10 shows the time-of-flight kernel (as was obtained in Figure 4.8) for a low and high photon collection efficiency configuration of monolithic 50x50x16 mm^3 BGO: absorbing lateral sides with a polished back surface, and reflective lateral sides with a rough back surface. The longer tails can be attributed to more events with fewer Cerenkov detections.

A similar effect can be seen when comparing purely photoelectric with scattered (but still 511 keV) events, showing considerable time resolution

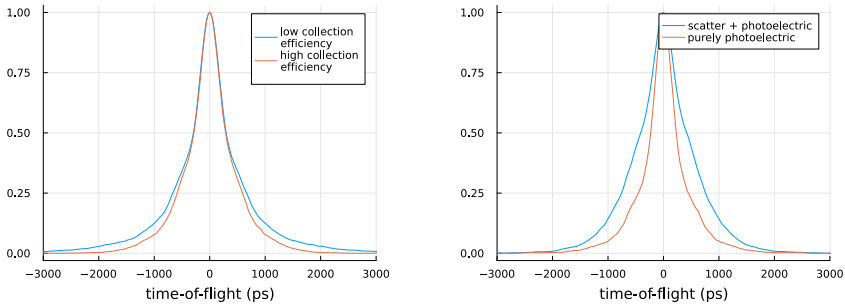


Figure 4.10: Time-of-flight kernels for monolithic 50x50x16 mm^3 BGO. **Left**: two different surface finish configurations, one with low photon collection efficiency (rough black lateral and polished back surface) and one with high photon collection efficiency (polished reflective lateral and rough back surface). **Right**: comparison between purely photoelectric and scatter + photoelectric events for the high photon collection efficiency configuration.

degradation for scattered events. The effect is now no longer confined to the tails, increasing the FWHM from 465 ps for purely photoelectric events to 817 ps for scatter + photoelectric events. This is due to a reduced emission of Cerenkov light. Simulations of the monolithic 50x50x16 mm^3 BGO detector show only 10 Cerenkov photons emitted on average per scatter + photoelectric event, compared to 18 photons for purely photoelectric events. Our simulations showed that for monolithic BGO (50x50x16 mm^2), 41% of 511 keV events were purely photoelectric, and the remaining 59% were scatter + photoelectric events. Note that this poses less of a problem for pixelated detectors, where the majority of 511 keV events are purely photoelectric since a scattered gamma photon often exits the crystal before depositing the rest of its energy. For monolithic detectors though, scattered events will not only contribute to a degradation of spatial resolution, but also time resolution.

Identification of scattered events could however allow for image reconstruction with multiple time-of-flight kernels, improving overall image signal-to-noise ratio [97]. Scatter identification was for example done for LYSO with a deep learning based approach using simulated data [113]. The difficulty here lies in using a network trained on simulated data for experimental data. Additionally, BGO has shorter range scatters com-

pared to LYSO, which makes them more difficult to identify. Therefore, the feasibility of such an approach would require further investigation.

Another approach for time-of-flight kernel separation would be identifying events with good timing based on the SiPM signal rise time, as was previously done in pixelated BGO detectors [98]. This would likely require summing of the SiPM signals in monolithic detectors, since most Cerenkov photons are absorbed by different SiPMs and the signal rise time of individual SiPMs is therefore unlikely to vary much. It should nonetheless be possible to implement for monolithic detectors, providing an additional method to improve image signal-to-noise ratio for a given detector configuration.

4.4.4 Accuracy of simulation results

While we ourselves did not perform any experimental validation of the simulation pipeline, it should be noted that GATE, the LUT Davis Model, and the SiPM waveform simulation have previously been validated experimentally [20], [103], [105]. In addition, the CTR value we obtain with our simulations for 3x3x20 mm^2 pixelated LYSO (with polished ESR grease side surfaces and a polished back surface) was 166/169 ps FWHM, when triggering at 0.5/1.5 photoelectrons respectively. This matches well with experimental results from various groups:

- 157 ps FWHM for 3x3x19 mm^3 LYSO:Ce,Ca with Broadcom NUV-MT SiPM and TOFPET2 ASIC [114].
- 148 ps FWHM for 3x3x20 mm^3 LYSO:Ce,Ca with Broadcom AFBR-S4N33C013 SiPM and Radioroc ASIC [115].
- 170 ps FWHM for 3x3x15 mm^3 LYSO with Hamamatsu LCT MPPC SiPM and NINO ASIC [116].
- 183 ps FWHM for 3x3x20 mm^3 LYSO with HamamatsuS13361-3050AE-04 MPPC SiPM and TOFPET2 ASIC [117].

4.5 Conclusion

This study shows the importance of optimizing the photon collection efficiency to improve Cerenkov based time estimation in monolithic BGO detectors. Multiple detector configurations were simulated for comparison, including different geometries (pixelated and monolithic),

surface finishes (for the lateral and back sides) and scintillation materials (LYSO and BGO). While changing the surface finish had little effect on the time resolution in monolithic LYSO due to already high photon collection efficiencies, it plays an important role in Cerenkov based time estimation for BGO due to much lower photon statistics. Surface finishes that improve photon collection efficiency result in time-of-flight kernels with shorter tails. Commonly used surface finishes in monolithic detectors (normally optimized for spatial resolution) result in inferior time resolution, showing potential benefit in making a switch to other surface finishes that increase photon collection efficiency such as reflective sides and a rough back surface.

In the next chapter, we will investigate how the time resolution of monolithic detectors can further be improved with the use of deep learning to extract more relevant information from the SiPM signals. We will focus on the specific case of monolithic LYSO, although the general methodology should be applicable to BGO as well.

Chapter 5

Deep learning time estimation in monolithic LYSO

5.1 Introduction

Most modern positron emission tomography (PET) systems include time-of-flight (TOF) capabilities. It allows for additional information on the position of the positron annihilation to be used during the reconstruction process, improving the signal-to-noise ratio of scans and resulting in better lesion detectability [118]. Therefore, an important parameter of these TOF-PET systems is the coincidence time resolution (CTR), the precision to which the time-of-flight difference between two coincident gamma photons can be measured.

Clinical PET scanners currently make use of pixelated detectors, where a scintillation crystal is subdivided into a matrix of so-called pixels, each a few millimeters wide. When a gamma photon is absorbed by the crystal, its energy is converted into optical scintillation light, which remains confined to the pixel in which the gamma absorption took place. Silicon photomultipliers (SiPMs) coupled to the crystal then convert this optical light into electronic signals, from which positioning and timing information on the gamma photon can be derived. State-of-the-art clinical TOF-PET scanners offer CTRs just over 200 ps on the system level [119]–[121]. Meanwhile, isolated setups making use of thin pixelated scintillation crystals and dedicated readout electronics have achieved CTRs below 100 ps, accomplished by the use of digitizers to analyze the

detector waveforms [96], [122].

Although such pixelated detectors remain popular due to their readout simplicity and good timing performance, monolithic detectors have been proposed as an alternative for use in PET. They have already been introduced in certain pre-clinical systems as well [104]. These detectors, consisting of a large monolithic scintillation crystal tens of millimeters wide, offer high spatial resolution and depth-of-interaction decoding capabilities at often lower costs [99]. A substantial amount of work has already been carried out toward improving the spatial resolution of monolithic detectors, using both conventional [45], [100] and deep learning [47]–[49] based methodologies. Timing estimation however remains difficult. Whereas in pixelated detectors the scintillation light is usually focused onto a single SiPM due to internal reflections, in monolithic detectors the scintillation light is spread out and diluted over multiple SiPMs, resulting in inferior signal-to-noise ratios, negatively impacting the electronic signal pickoff for timestamp calculation. The inclusion of intra-detector scattered events also splits the scintillation pulses, adding a layer of complexity. On the other hand, monolithic detectors provide access to multiple SiPM waveforms per gamma event, intrinsically containing spatial information that could help with time-of-flight prediction.

Traditionally, the gamma arrival time in a monolithic crystal would be calculated by an averaging of the first few recorded SiPM timestamps [123]. These can be obtained directly in digital SiPMs or generated by e.g. leading edge discrimination on the electronic signal in analog SiPMs. In this work, we focus on the latter type of photodetector. A big advantage of leading edge discrimination is the ability to be implemented on an analog readout without the need for signal digitization. This is however rather susceptible to noise from dark counts and other statistical fluctuations. Digitizing the SiPM signals allows us to perform an accurate baseline correction, helping to alleviate these problems. Nonetheless, the use of a simple predictor such as leading edge discrimination to compress a large amount of possibly relevant signal information into a single value per SiPM remains a limitation of this methodology. We expect that more advanced methods making direct use of the detector waveforms could extract more relevant information while filtering out noise influences, ultimately improving the timing resolution.

Inspired by the success of deep learning approaches for time-of-flight estimation between two pixelated detectors [53], [124], we perform a

simulation study to investigate the potential of deep learning for gamma arrival time estimation in monolithic scintillation detectors, based on digitized detector waveforms. We propose the use of a convolutional neural network (CNN) using 3D convolutional kernels, so we can simultaneously include temporal and spatial information contained in the array of detector waveforms.

5.2 Materials and Methods

5.2.1 Data generation

We simulate 511 keV gamma photons impinging on a monolithic 50x50x16 mm^3 LYSO crystal, coupled to an 8x8 readout array of 6.07x6.07 mm^2 analog SiPMs. This specific configuration is based on an existing monolithic detector design, for which a spatial resolution of 1.14 mm full width at half maximum (FWHM) has been achieved [100]. GATE v8.2 [26] is used to model the gamma interactions and production of scintillation light. The optical photons hitting the photodetector surface are then further used to simulate the electronic SiPM signals. These processes are described in more detail in the following sections.

We only take into account events that deposit their full energy in the crystal. In this way, a dataset of 10 million gamma photons uniformly and perpendicularly incident on the front surface of the crystal is obtained. It is further subdivided into a training, validation and test set of 6 million, 2 million and 2 million events respectively. The validation set is used for hyperparameter tuning, and for early stopping during training. Making use of the 8-fold symmetry in the crystal, we only simulate 12.5% of each dataset and generate the rest from rotations. In doing so, we make sure that all 8 rotations of a particular event belong to the same dataset, in order to avoid any cross-contamination. Additional test datasets of 1 million events each are obtained for oblique incidence angles at 10°, 20°, 30°, 40° and 50° to the surface normal.

GATE simulation

511 keV gamma photons are generated uniformly over the crystal top surface, traveling downward to the photodetector. Along the way, the gamma photons may interact with the crystal by the photoelectric effect, Compton or Rayleigh scattering, thereby producing scintillation light which is further tracked by GATE. For the LYSO material, we

use a density of 7.36 g/cm^3, an expected scintillation yield of 29 000 photons/MeV, rise time of 70 ps, decay time of 42 ns, refractive index of 1.82 and optical absorption length of 1 m [18].

Surface reflections of scintillation photons are simulated using the LUT Davis model [105], [125], [126]. We define the lateral sides of our detector as rough surfaces with black paint, the top surface as polished and coupled with optical grease to a specular reflector, and the bottom surface as polished and coupled with optical grease to a photodetector. Any optical photons that are transmitted through this bottom surface reach the photodetector and can be further used in the simulation of the electronic SiPM signals.

SiPM simulation

To a good approximation, the SiPM response can be modelled as a sum of bi-exponential functions centered around the optical photon detection times. Prior to doing so, we include additional sources of noise to account for non-idealities, largely inspired by the work in [20]. The parameters used in our simulation are based on the SensL J-Series 60035 SiPMs [127], and are given in Table 5.1. We use two different sets of parameters, one for the SiPMs operating at an overvoltage of +2.5 V and one for an overvoltage of +6 V. The maximum photon detection efficiency of 50% is reached at +6 V. This overvoltage is optimal for spatial resolution, since the gamma positioning improves with increasing photon detection efficiencies in these monolithic PET detectors [128]. Note however that the rise time, dark count rate and optical crosstalk probability are all lower for +2.5 V. Therefore, we would like to investigate if better time resolution could be achieved in this case, despite the reduced photon detection efficiency.

The photon timestamps obtained from GATE are first assigned to their corresponding SiPM, with photons being randomly removed to account for the limited photon detection efficiency. Additional dark counts are then generated by sampling timestamps from a Poisson process. Next, a photodetector transit time spread is modeled by convolving the given timestamps with a Gaussian, with standard deviation corresponding to the single photon time resolution (SPTR) of the SiPM. Note that this is the SPTR prior to the effects of electronic noise and is therefore not often reported, but we use a realistic estimate of 50 ps [96]. Finally, prompt optical crosstalk with a certain probability gives rise to duplicate counts, each of which can again result in crosstalk with the same probability.

overvoltage	+2.5 V	+6 V
photon detection efficiency	38 %	50 %
dark count rate	50 kHz/mm^2	150 kHz/mm^2
single photon time resolution σ	50 ps	50 ps
prompt optical crosstalk probability	8 %	25 %
rise time τ_r	82 ps	114 ps
fall time τ_f	50 ns	50 ns
electronic noise σ	0.1 p.e.	0.1 p.e.

Table 5.1: The SiPM parameters for the SensL J-Series 60035 SiPMs at two different overvoltage values [127]. Note that the single photon time resolution and the electronic noise contribution were not given by the supplier, so we had to resort to an estimation. The electronic noise is expressed in terms of a single photoelectron (p.e.) pulse amplitude.

The i-th crosstalk event receives a time smearing $\sigma = \text{SPTR} \cdot \sqrt{i+1}$ with an additional time delay $\mu = \text{SPTR} \cdot \sqrt{i}$ [20].

The SiPM signal $s(t)$ is then generated as a sum of bi-exponential functions centered around the photon detection times t_i, with rise-time τ_r and fall-time τ_f:

$$s(t) = \sum_i \left(\exp\left[\frac{t_i - t}{\tau_f}\right] - \exp\left[\frac{t_i - t}{\tau_r}\right] \right) H(t - t_i) \tag{5.1}$$

The Heaviside function H sets the microcell signal to zero prior to photon detection. The signal is generated over a time frame of roughly 1 µs, using a varying time step size. This allows for both a fine time resolution around the gamma event and a large enough time window to include the full signal, while keeping computational demands reasonable. The smallest time step size of 50 ps is used for a range of 14 ns, including about 4 ns of baseline prior to the gamma event. Finally, electronic noise is added as zero-mean white Gaussian noise with standard deviation σ_e equal to 10% of the amplitude of a single photoelectron (p.e.) pulse. The electronic noise contribution could vary wildly depending on the electronics used for signal digitization and is therefore chosen somewhat arbitrarily. A standard deviation of 0.1 p.e. would correspond to an electronic noise contribution of roughly 50 ps FWHM to the SPTR of a single SiPM, according to $\sigma_e = \sigma_t \frac{dV}{dt}$, measuring the signal slew rate dV/dt at a threshold level of 0.5 p.e. and using +6 V overvoltage.

5.2.2 Neural network training

There are two ways for developing neural networks for time-of-flight prediction between PET detectors. The first would be to consider two detectors in coincidence, and use the SiPM signals from both to directly predict the time-of-flight difference. Such an approach was used in [53] for two detector pixels in coincidence, using a 2D CNN for prediction. The ground truth labels were obtained by moving a source to different known positions between the two detectors, from which the time-of-flight difference is easily calculated.

The second is to consider only a single detector, and use its detector waveforms to predict the time at which the gamma photon arrived at the crystal. This can then be done separately for two detectors in coincidence, after which the time-of-flight difference is calculated by subtracting both gamma arrival times. This methodology has the advantage that we can perform detector calibration on a per-detector basis, rather than requiring a separate calibration for each possible pair of detectors. The ground truth gamma arrival times can be obtained by using a small pixelated reference detector with good timing performance in coincidence. In addition, there is no need to acquire data at different source distances, since we are only interested in the symmetric case where the gamma photons reach both the monolithic and reference detector at the same time.

Using the second approach, our goal is to train a neural network that, given an 8x8 array of SiPM signals as input, will be able to predict the gamma arrival time with respect to the time axis used in the digitized detector waveforms. Utilizing the full SiPM signals over the whole 1 µs range is however not feasible as this would either result in too high memory usage, or would require us to use too large time steps. Instead, we truncate the signals to a time window of 3 ns using time steps of 100 ps. This corresponds to a sample rate of 10 GS/s, and is well within the limits capable of modern digitizers. The time window is the same for all 64 SiPMs, but is chosen separately for each individual gamma event. The truncation is performed without any prior knowledge of the gamma arrival time and is solely based on the signals themselves. It is done by calculating the onset of the sum of the 64 signals, where the time window runs from 500 ps prior to the signal onset until 2500 ps after. Finally, the truncated waveforms are baseline corrected using the first 500 ps as reference. This results in 8x8x31 input arrays for the neural network. As target labels for training, we use the true gamma arrival times obtained

from GATE, convolved with a 100 ps FWHM Gaussian, representing the reference detector that would be used for calibration.

Note that the network requires no information about the gamma interaction positions, in contrast to other advanced time-of-flight prediction methods such as maximum likelihood interaction time estimation [129]. This means that the detector can be calibrated using only a single source position (placed sufficiently far away to irradiate the whole detector), greatly simplifying and speeding up the detector calibration procedure.

Given the nature of our data, we have opted to use a convolutional neural network architecture. It makes use of trainable convolutional kernels to extract feature maps from structured input data, resulting in a more regularized network due to a combination of parameter sharing and sparse connectivity. In our specific case, we are dealing with two spatial dimensions and one time dimension, so we make use of 3D kernels to extract features that combine the information from all three dimensions. The proposed network architecture is shown in Table 5.2. The architecture was chosen empirically to be as small as possible without showing performance degradation, verified using the validation set. While network size is not a bottleneck in these research studies, it needs to be considered for implementation on a physical scanner, where inference speed (and potentially, memory footprint if implemented on the detector itself) are more important. This is another potential benefit of using a CNN rather than a fully connected neural network: weight sharing through the use of convolutional kernels results in a reduced memory footprint. Our chosen network weights are optimized using the ADAM algorithm with a learning rate of 0.001 and mini-batch size of 256. The mean squared error between predicted and reference gamma arrival time is used as loss function, and early stopping was employed by training until the validation loss did not improve for 5 epochs.

5.2.3 Reference method 1: timestamp averaging

In order to provide a frame of reference for the neural network performance, we also predict the gamma arrival time using the traditional methodology of leading edge discrimination followed by an averaging of the first few recorded timestamps. As a preprocessing step, the full, non-truncated SiPM waveforms are again baseline corrected using the same 500 ps time window prior to the signal onset of the sum of the 64 SiPM waveforms. We test a range of different numbers of timestamps (1 - 10) used for the averaging as well as different threshold levels (0.1 p.e. -

	conv 1	conv 2	conv 3	dense 1	dense 2
channels	$1 \rightarrow 16$	$16 \rightarrow 24$	$24 \rightarrow 32$	$128 \rightarrow 32$	$32 \rightarrow 1$
kernel size	3x3x5	3x3x5	3x3x5		
stride	1x1x2	1x1x2	1x1x2		

Table 5.2: The network architecture for the 3D CNN. Each block consists of its respective operation (a 3D convolution or a fully connected layer) followed by a batch normalization and ReLU activation. The final dense block does not contain any batch normalization or activation function.

5 p.e.) for the leading edge discrimination, and use the best performing combination going further.

Note that this method still requires the use of digitized SiPM waveforms to perform the baseline correction, especially when using low threshold levels. The comparison between this and the neural network should therefor be seen as a measure for how much additional information could be obtained from digitized SiPM waveforms by making the switch to more advanced time prediction methods, rather than as a comparison between a fully analog and a digitized readout.

5.2.4 Reference method 2: 2D CNN

Finally, we also compare the results to another deep learning based approach, which uses the SiPM timestamps generated in reference method 1 as an input to a 2D CNN to predict the gamma arrival time. With this method, we again lose valuable temporal information stored in the rising edge of the SiPM waveforms. However, compared to the timestamp averaging method, the 2D CNN should be better at making use of other available information (e.g., spatial correlations) in the timestamp matrix, resulting in better performance. The training procedure is the same as for the 3D CNN, only we now use 2D (8x8) input data. The timestamps are obtained using the best performing threshold level from reference method 1. The network architecture is shown in Table 5.3. We only trained and tested the 2D CNN for the data at +6 V overvoltage.

	conv 1	conv 2	conv 3	dense 1	dense 2
channels	$1 \rightarrow 32$	$32 \rightarrow 48$	$48 \rightarrow 64$	$256 \rightarrow 32$	$32 \rightarrow 1$
kernel size	3x3	3x3	3x3		
stride	1x1	1x1	1x1		

Table 5.3: The network architecture for the 2D CNN. Each block consists of its respective operation (a 2D convolution or a fully connected layer) followed by a batch normalization and ReLU activation. The final dense block does not contain any batch normalization or activation function.

5.3 Results

5.3.1 Overall detector performance

We report the detector performance in terms of the CTR that would be obtained from two identical detectors in coincidence, see Figure 5.1. To this end, the single detector time prediction errors are first calculated with respect to the true gamma arrival time, as obtained from GATE prior to convolution with the reference detector. These are then fit to a distribution by means of kernel density estimation, from which the CTR can be obtained by convolving the single detector distribution with its own reflection around t=0. Using the 3D-CNN on the test set of perpendicular incidences, we achieve a CTR of 141 ps FWHM for the SiPMs operating at +6 V overvoltage and 162 ps FWHM at +2.5 V.

We can compare these results to the traditional methodology of leading edge discrimination followed by an averaging of the first few recorded timestamps. Figure 5.2 shows the obtained CTR FWHM for different threshold levels and numbers of timestamps at both overvoltage values. The best CTR for +6 V is obtained by an averaging of the first 4 timestamps at a threshold of 0.6 p.e, resulting in 177 ps FWHM. The best CTR for +2.5V is obtained by an averaging of the first 3 timestamps at a threshold of 0.5 p.e, resulting in 187 ps FWHM. That is, the 3D-CNN improves performance by 26% for +6 V and by 15% for +2.5 V compared to this traditional approach. For the 2D CNN at +6 V, we obtain a CTR of 151 ps, an improvement of 17% compared to the timestamp averaging method. A more detailed overview of the timing performance is given in Figure 5.3. For the remainder of the results section, we will focus only on the results obtained at an overvoltage of +6 V.

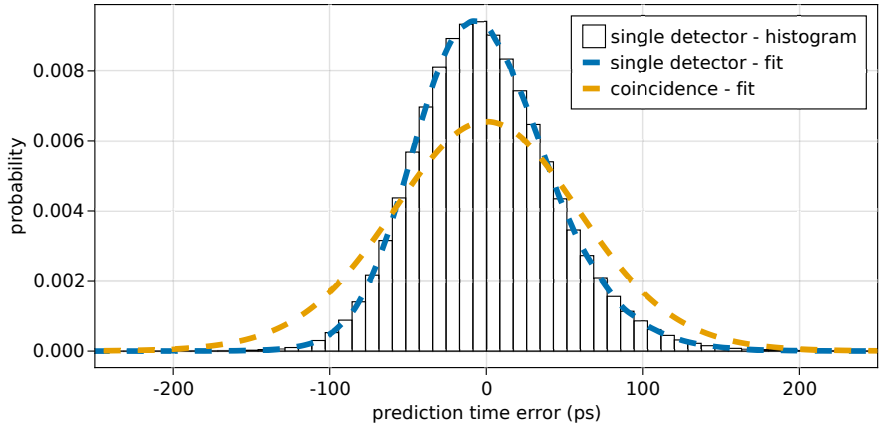


Figure 5.1: Single and coincidence time distributions obtained for the 3D-CNN, as evaluated on the waveforms obtained at +6 V overvoltage from the test set of perpendicular incidences. From this, we obtain a CTR of 141 ps FWHM.

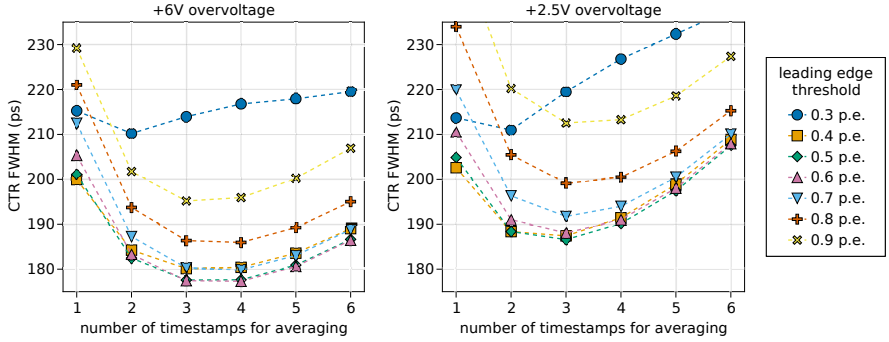


Figure 5.2: Coincidence time resolution FWHM values obtained by leading edge discrimination at different threshold levels followed by an averaging of the first few timestamps. Note that we are only displaying a subset containing the best performing threshold levels and numbers of timestamps for clarity, though we did investigate the full range from 0.1 p.e. to 5 p.e. and from 1 to 10 timestamps.

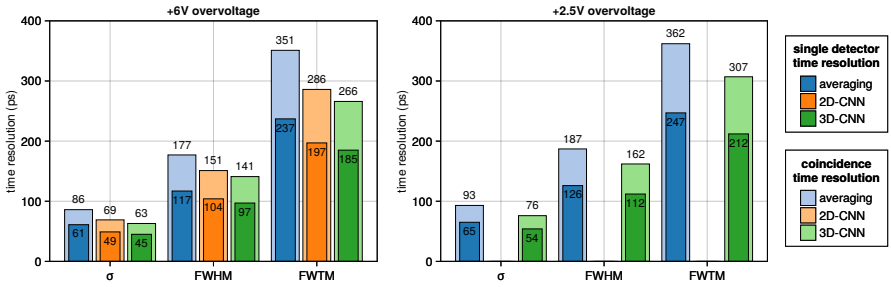


Figure 5.3: Overview of both the single detector and coincidence time resolutions obtained for different overvoltage values and prediction methods.

	0°	10°	20°	30°	40°	50°
σ	63 ps	63 ps	63 ps	64 ps	64 ps	65 ps
FWHM	141 ps	142 ps	142 ps	142 ps	144 ps	145 ps
FWTM	266 ps	266 ps	267 ps	268 ps	271 ps	273 ps

Table 5.4: Coincidence time resolution for datasets of different gamma incidence angles with respect to the front surface normal of the detector, compared to perpendicular (angle = 0°) incidences. Results were obtained for signals simulated at an overvoltage of +6V, with the same network trained on only perpendicular incidences being used in all cases.

5.3.2 Oblique incidences

Table 5.4 shows how the 3D CNN, trained only on perpendicular incidences, performs on the additional test datasets for oblique incidence angles. We only notice a minor deterioration of the time resolution, with an overall degradation of 3% going from perpendicular incidence angles to 50° oblique incidence angles. Therefore, we conclude that it suffices to perform a single calibration using only perpendicular incidences.

5.3.3 Spatial dependency

Figure 5.4 compares the spatial dependency of the time resolution for the 3D-CNN, 2D-CNN and timestamp averaging methods. We consider the time resolution both as a function of the 2D and depth-of-interaction positions. Note that here the position of a gamma event specifically refers to its first interaction position. In addition, we are now only measuring

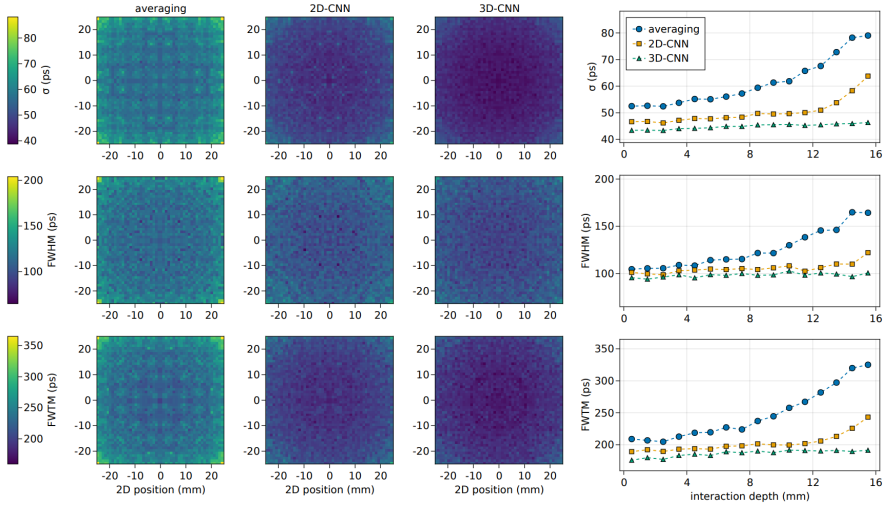


Figure 5.4: Spatial dependency of the single detector time resolution for the 3D-CNN, 2D-CNN and traditional averaging methods. The first interaction position is used in case of scattered events. The results given are those for perpendicular incidences only, using an SiPM overvoltage of +6 V. Each row shows a different metric (σ, FWHM and FWTM), with the first three columns showing the 2D spatial dependency, and the last column showing the depth of interaction dependency. In this last column, the crystal top surface is located at 0 mm, and the SiPM surface is located at 16 mm.

the single detector time resolution. The CTR would not be well-defined since two gamma photons can interact in different positions in each detector. All methods show a degradation near the edges and corners, although it is less pronounced for the neural network methods. More notably however, the averaging method shows a steep degradation of the time resolution for deeper interactions, i.e. interactions closer to the SiPM surface, which is less severe for the 2D-CNN, and not present in the 3D-CNN method. Altogether, the deep learning methods provide better detector uniformity in addition to the overall improved performance.

5.3.4 Effect of sampling rate

We have found that acceptable time resolutions could still be achieved by the 3D-CNN when switching to considerably lower sampling rates, see

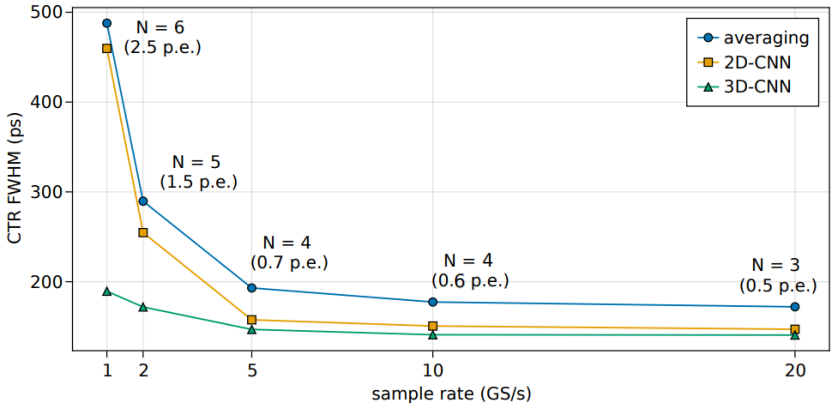


Figure 5.5: Obtained coincidence time resolution for the 3D-CNN, 2D-CNN and timestamp averaging methods, using the waveforms at an overvoltage of +6V, digitized at different sampling rates from 1GS/s to 20 GS/s. We show the number of timestamps (N) and the threshold level (in p.e.) used for the timestamp averaging method at each sampling rate. The same threshold levels are used for the 2D-CNN method.

Figure 5.5. On the other hand, the timestamp averaging methodology, as well as the 2D-CNN, suffer considerably from decreasing the sampling rate. For the 3D-CNN, we use the same network architecture but vary the time window for signal truncation to maintain the input size of 8x8x31. So for example, using a sampling rate of 1 GS/s, corresponding to 1 ns time steps, we use a time window from 5 ns prior to the signal onset until 25 ns after. For the timestamp averaging method we have again performed a parameter sweep over different leading edge threshold levels and numbers of timestamps for averaging, and are only reporting the results of the best performing combination at each sampling rate. The same leading edge thresholds are used for the 2D-CNN. While waveform digitizers will only provide improvements over fully analog readouts at sufficiently high sampling rates, the use of deep learning to extract information directly from the raw detector waveforms could considerably lower these sampling rate requirements.

5.4 Discussion

5.4.1 Method comparison

The improved time resolution achieved by our 3D-CNN can be attributed to a better use of both temporal and spatial information which otherwise remains unused. Indeed, leading edge discrimination at such low threshold levels is really only influenced by the first scintillation photon detected by each SiPM. Then, during the averaging of the first few timestamps, no attention is given to the relative positions of the SiPMs in question. In comparison, the 3D-CNN takes into account a time window which includes a fair fraction of the early scintillation photons, and is able to take into account the spatial distribution of scintillation light. The 2D-CNN loses some of this information, but retains access to the relative positions of the timestamps.

Another limitation of the timestamp averaging method was shown in Figure 5.4. The time resolution quickly degrades for interaction positions close to the SiPM surface. For these events, the majority of scintillation light is absorbed by the SiPM pixel directly underneath. This negatively impacts the timing performance due to the inclusion of less relevant timestamps, obtained from SiPMs which received only a small fraction of the scintillation light. A similar phenomenon can be observed in the 2D spatial dependency of the time resolution, most notably visible when considering the standard deviation σ. There are recurring peaks of degraded time resolution which coincide with the centers of SiPM pixels, where again a substantial fraction of the scintillation light is absorbed by a single SiPM. The 3D-CNN is however better capable of dealing with this varying importance of different SiPMs, resulting in improved detector uniformity.

We can try to improve the timestamp averaging method by optimizing the leading edge threshold and number of timestamps for each DOI layer individually. We split the test set (+ 6V overvoltage) into 16 DOI layers of 1 mm each, according to the first interaction position. The single detector time resolution FWHM for the best performing combination of leading edge threshold and number of timestamps for each DOI layer is reported in Figure 5.6, and compared to the previously obtained results from Figure 5.4. The time resolution of deep interactions can indeed be improved by reducing the number of timestamps, although an overall degradation is still observed for DOIs closer to the SiPM array. While the concentration of scintillation light reaching individual SiPMs will be

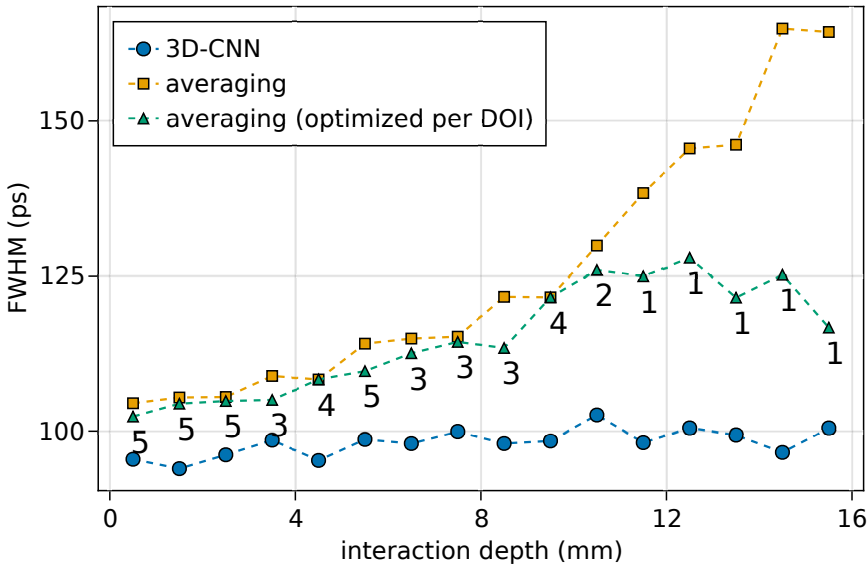


Figure 5.6: Single detector time resolution FWHM for different DOIs, obtained by optimizing the leading edge threshold and number of timestamps for each DOI layer individually. The previous results of Figure 5.4 (using 4 timestamps and a threshold of 0.6 p.e) are also included for comparison. Only the optimal number of timestamps is shown for each DOI layer, as the optimal leading edge threshold did not change much.

smaller for shallow interactions and therefore result in noisier timestamps, the inclusion of multiple timestamps allows to somewhat average out these noise contributions. This results in superior time resolution compared to deeper interactions that only use a single timestamp, despite the higher concentration of scintillation light reaching that specific SiPM pixel. This shows that knowledge of the DOI can help to improve the time resolution by giving a clue about which timestamps contain relevant information. The neural network methods show reduced degradation for deep interactions, as they already have indirect access to this DOI information, through analysis of the spatial distribution of scintillation light and timestamps.

5.4.2 Effect of overvoltage

Our results have shown that, of the two overvoltage values simulated, +6 V resulted in a better timing accuracy for both prediction methods. This can be attributed to the additional 12% photon detection efficiency, allowing us to measure earlier scintillation photons which would have otherwise remained undetected. Meanwhile, the inclusion of a baseline correction helps to minimize the primary drawback of the higher overvoltage, the increased dark count rate. Given the importance of photon detection efficiency for good spatial resolution [128], this implies that we are capable of optimizing for both timing and spatial resolution. It should be noted that this may not necessarily be true for a fully analog readout, where the influence of dark counts is not as easily removed.

5.4.3 Influence of the reference detector

Although the reference detector introduces noise to the targets used for network training, it has no influence on the achievable CTR. The same results would have been obtained without the reference detector, i.e. by using the ground truth gamma arrival times as targets. It was only included in our simulations for completeness. Indeed, it has been shown that the addition of zero mean Gaussian noise to the targets of a neural network does not alter the solution of the mean-squared error optimization problem [130]. Even if a time bias were introduced by our reference detector, resulting in a non-zero mean noise contribution, the same bias would be present in all calibrated detectors. This identical bias would then simply cancel out during the time-of-flight calculation between two coincident detectors.

5.4.4 Limitations of the simulation model

Finally, we would like to discuss some simplifications used in our simulation model. First, we do not include any saturation effects of the SiPM. We can safely ignore these since the number of microcells in these particular devices (22292 per SiPM) is much larger than the number of incoming photons per gamma event, which remains in the order of a few hundred per SiPM pixel. Neither do we simulate afterpulsing, as it has a relatively low probability (0.75% for +2.5 V and 5% for +6 V overvoltage) and little impact on the timing performance of the SiPM, which is primarily affected by the leading edge of the pulse. We also do not model any time-skew between the SiPM pixels themselves. In a real

setting, this could easily be corrected for during a prior calibration step. Alternatively, if the time-skews are not too large, it is likely that the neural network would learn to automatically account for them.

We also note that in the selection of our datasets, we only take into account those events that deposit their full energy, i.e. exactly 511 keV, in the crystal. So all non-scattered events, and in the case of scattering, only those events that end their interaction chain with a photoelectric absorption, are included. Scattered events where the gamma photon leaves the crystal with a portion of its energy are excluded. In reality, energy windows are however not perfect. This would in a practical setting result in the inclusion of some events without full energy deposit, and likewise, the exclusion of some events with full energy deposit.

All things considered, we have tried to limit the potential shortcomings of our simulation model as much as possible. While we can likely expect that an experimental evaluation of these methods would result in an overall degradation of the time resolution, we do believe our model to be sufficiently realistic to show overall trends and draw general conclusions from.

5.5 Conclusion

We performed a simulation study showing that, making use of waveform digitizers in monolithic PET detectors, superior time resolution can be achieved by switching to more advanced time prediction methods such as 3D-CNNs, capable of extracting more relevant information from the raw SiPM signals. Using 100 ps time steps for signal digitization, we reached a CTR of 141 ps FWHM for a monolithic 50x50x16 mm^3 LYSO crystal, coupled to an 8x8 readout array of SiPMs. This is a 26% improvement compared to the traditional methodology of leading edge discrimination followed by an averaging of the first few timestamps, for which a CTR of 177 ps FWHM was obtained. A 2D-CNN method which makes use of those same timestamps, reached a CTR of 151 ps FWHM, a 17% improvement compared to timestamp averaging, providing a middle ground in terms of performance and complexity.

Part II will focus on the development of an image reconstruction framework for a new, dual-panel, long AFOV PET scanner (the Walk-Through PET) under construction at our research group, based on monolithic detector technology. The studies performed in part II use a scanner design based on monolithic BGO, although ultimately, it has

been decided to use LYSO for the final design. In the future, it is the intention to implement a methodology similar to the 2D-CNN described in this chapter on the physical scanner.

Part II

Image reconstruction for the Walk-Through PET

Chapter 6

The Walk-Through PET

The Walk-Through PET (WT-PET) is a new, long AFOV PET scanner under development at our research group, see Figure 6.1 [4], [131]. The goal is to offer an affordable, high-throughput alternative to existing total-body or long AFOV PET scanners. It is based on monolithic detector technology, and departs from the cylindrical geometry of conventional PET scanners, instead making use of two flat detector panels, spaced 50 cm apart. With an AFOV of 106 cm, it is capable of simultaneous head and torso imaging at high sensitivity, offering the possibility of short, sub-one minute scan durations. The flat panels are positioned vertically, rather than horizontally, so that patients are scanned in a standing position for increased patient throughput.

The flat panel design of the system allows for the detectors to be placed closer to the patient (50 cm gap as opposed to $\sim$ 80 cm diameter), resulting in improved geometrical coverage for a lower number of detectors. Doing so, the number of detectors can be reduced by a factor of 1.9 while maintaining similar sensitivity as a cylindrical design [131]. This directly translates into a cost reduction for the system by almost half, and the use of monolithic detectors enables further cost savings due to reduced losses from crystal cutting. A final cost reduction is introduced by making use of BGO rather than LYSO, as BGO is roughly 3x cheaper per unit of volume. Overall, the aim is to keep the system cost in the same range of, or slightly more expensive than, conventional short AFOV PET/CT systems.

Meanwhile, scanning patients in a standing position allows for increased patient throughput by eliminating the procedure of positioning the patient on the bed. As patient throughput in current long AFOV

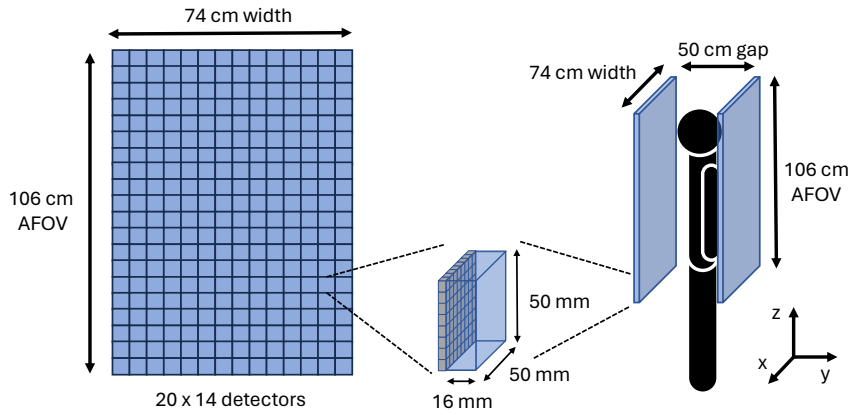


Figure 6.1: Design and dimensions of the WT-PET system. The axis definitions (x, y, z) as used throughout the following chapters are also given.

PET scanners is primarily held back by the patient preparation procedure, rather than the already short scan duration, this would result in substantial throughput gains, not only compared to short AFOV systems, but also compared to other long AFOV systems. It is estimated that 9 - 10 patients could be scanned per hour on the BGO-based WT-PET, compared to 7 - 8 per hour on other 106 cm long AFOV LSO-based PET systems, and 3 - 4 per hour on short AFOV systems [131]. The gap between the panels allows for patients to enter and exit the scanner by themselves, or with minimal assistance. By giving more independence to patients, we also expect that this can be translated into a reduction of required personnel.

We acknowledge that a certain percentage of PET patients is either bedridden or too immobile, and could therefore not be scanned on the WT-PET. From patient tests that were carried out at CHU de Liège (a hospital in Liège, Belgium), we have found that this was only a small percentage of PET patients, roughly 10% [131]. As several hospitals with PET scanners do have at least two systems, the WT-PET could therefore be used as a primary high-throughput scanner, with a second conventional short AFOV scanner as backup for bedridden patients.

6.1 Flat panel design: advantages and disadvantages

The main purpose of the flat panel design is to decrease the cost by reducing the number of detectors. However, it also has some other, both desirable and undesirable, side effects as well. For one, the design leads to missing projection angles along the transverse view, which can result in artifacts manifesting in the reconstructed images. Their impact can be reduced by the introduction of TOF, which gives information on the annihilation position otherwise lost in the missing projection data. We will discuss this in more detail in Chapter 10, as well as develop an algorithm to reduce these artifacts. Another effect of the flat panel design and the proximity of the detectors to the patient, is that more oblique LORs will be measured compared to other long AFOV scanners. If it were not for the DOI capabilities of monolithic detectors, this would result in a severe degradation of the spatial resolution through the parallax effect. Instead, an increase in oblique LORs actually turns out advantageous, as the effective detector thickness seen by the LOR increases, resulting in improved sensitivity. Furthermore, by primarily measuring coincidences along the anterior-posterior axis, the loss in detections due to attenuation is reduced compared to cylindrical long AFOV PET systems, as humans are typically thinner front-to-back compared to side-to-side. Finally, the proximity of the detectors also reduces the effects of photon non-collinearity, which further improves the already high spatial resolution enabled by monolithic detectors.

6.2 Scanner parameters

The exact design of the WT-PET has, over the course of its development, undergone some modifications. The majority of the work in this dissertation is done based on the original design, which consists of two flat panels, each containing 14 x 20 (= 280) monolithic BGO detectors (see Figure 6.1). The scintillation blocks are 50 x 50 x 16 mm^3 in size, coupled to an 8 x 8 array of $\sim$ 6 x 6 mm^2 SiPMs. Leaving 3 mm gaps between the detectors in each direction, results in a panel size of 74 x 106 cm^2. BGO was chosen for its lower price, and studies showing that good time resolutions (CTR of 327 ps FWHM) could be achieved through simultaneous position and time estimation with neural networks [132].

Eventually, we have instead decided to use LYSO for its superior

time resolution, which is especially important for the WT-PET due to the presence of limited angle artifacts. In making this switch, we have decided to use a sparse detector configuration to offset the increased cost of LYSO, keeping 70% of the detector rows by removing every third row. The loss in sensitivity due to the missing detectors and the smaller stopping power of LYSO, is partially offset by the increase in effective sensitivity due to better TOF. That being said, we will focus on the BGO design from now on. The algorithms developed throughout the remainder of this thesis are, in any case, applicable to either design.

The most important scanner parameters of the WT-PET, as used throughout the simulations in further chapters, are shown in Table 6.1. The parameters of another, conventional long AFOV PET system, the Siemens Biograph Vision Quadra, are also given for reference [21]. The detector parameters for the WT-PET (spatial resolution, energy resolution and time resolution) are obtained from the aforementioned study [132]. We use a value of 400 ps TOF, rather than 327 ps, as a more realistic estimate for a system-level TOF resolution, given that the TOF value of 327 ps was obtained on a bench-top experimental setup. We note that the TOF kernel obtained in this study was close to Gaussian, in contrast to the tailed distributions often observed in Cerenkov-based time estimation. Likely, this is because the time estimation is primarily relying on the slower scintillation photons, but nonetheless a good time resolution is achieved due to the neural network based approach for simultaneous position and time estimation. A dead time of 370 ns was chosen to match the experimentally obtained maximum count rate of one million counts per second. For the Quadra, there are no experimentally reported values available for the dead time. We chose a value of 320 ns, which has previously been used in a simulation study of the Biograph Vision [133], where it was chosen to match the experimentally reported sensitivity values. In addition, a value of 320 ns gives us a good agreement between simulation and experiment for the NECR curves [4], [21]. In further chapters, we will make some comparisons with the Quadra, as it is currently the most widely used long AFOV PET scanner, and has the same AFOV as the WT-PET. We do note that the reconstructions for the Quadra will not necessarily be representative of images obtained on the physical Quadra system, and should therefore rather be seen as reconstructions for a system based on the Quadra geometry. While we have tried to model the Quadra as accurately as possible in the simulation, there are likely to be certain parameters that remain unaccounted for,

	WT-PET	**Quadra**
AFOV	106 cm	106 cm
inner gap / diameter	50 cm	82 cm
panel dimensions	74 x 106 cm	-
number of crystals	560	243 200
scintillation material	BGO	LSO
crystal size	50 x 50 x 16 mm^3	3.2 x 3.2 x 20 mm^3
detector resolution	1.3 x 1.3 x 2 mm^3 [a]	= crystal size
energy resolution	15%	11%
energy window	434 - 645 keV	435 - 585 keV
TOF resolution	400 ps	228 ps
CTW	5 ns	4.7 ns
detector dead time	370 ns	320 ns

[a] Values are the FWHM of a Gaussian, where the DOI resolution is 2 mm.

Table 6.1: Summary of scanner parameters for the WT-PET and the Quadra.

and the reconstructions themselves are not done with Siemens' software.

6.3 Alternative configurations

In addition to the WT-PET design discussed above, we will be considering two other configurations, for a total of three designs (see Figure 6.2):

- **WT-PET**: The original configuration of the WT-PET, consisting of 14 x 20 monolithic BGO detectors per panel, with a gap of 50 cm between both panels.
- **Sparse WT-PET**: A sparse configuration of the WT-PET, consisting of 14 x 14 monolithic BGO detectors per panel (by removing every third row of detectors of WT-PET), with a gap of 50 cm between both panels.
- **Rotating WT-PET**: A rotating configuration of the WT-PET, consisting of 14 x 20 monolithic BGO detectors per panel, with a gap of 70 cm between both panels, and the panels rotating around the central axis, so that the panels rotate 180° during one acquisition.

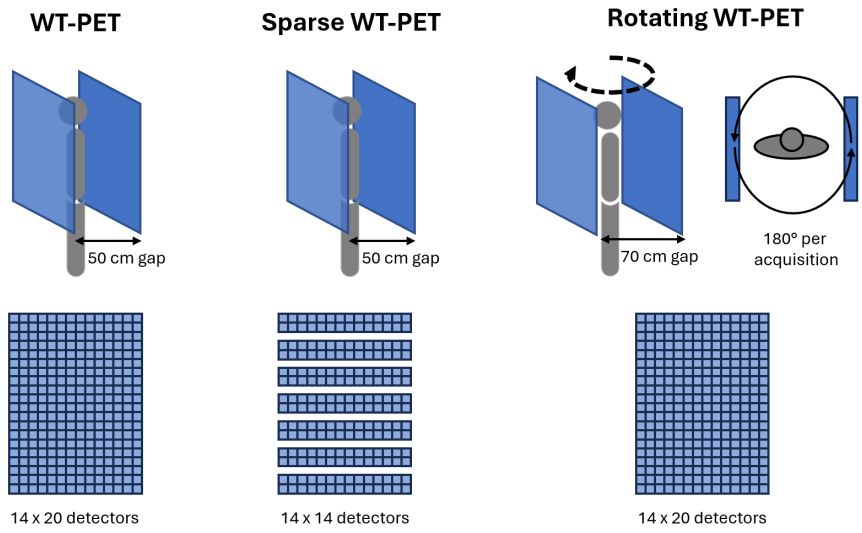


Figure 6.2: Different configurations of the WT-PET system.

The default **WT-PET** configuration forms the starting point of our comparison and is used to investigate the influence of the flat-panel design and monolithic detector technology on image quality and reconstruction. The **sparse WT-PET** is included to investigate and quantify the degradation that occurs when introducing sparsity in the design. Not only does sparsity reduce the sensitivity, but certain artifacts may also arise from the added gaps in the projection data. As the final LYSO design under construction is also sparse, any such artifacts should be eliminated. The **rotating WT-PET** is used to investigate scanner performance in the absence of limited angle artifacts, by providing full angular coverage with rotation of the panels. The rotation speed is chosen so that the panels rotate 180° during one acquisition, providing equal sensitivity along all projection angles. The gap between both panels however needs to be increased, in order to provide space for the patient's shoulders.

We are not planning to add such a rotation mechanism to the panels in the final design, in part because the increased gap size decreases sensitivity, and in part because it would add substantial mechanical complexity. Nonetheless, it would be feasible to instead position the patient on a rotating platform, which is substantially simpler in terms of mechanics. Rotating the patient would be easier than rotating the

panels, due to the heavy weight of the panels (the scintillation material itself already weighs $\sim$ 80 kg per panel), the increased complexity in connecting the electronic cables and cooling system to rotating panels, and the fact that the panels should remain aligned very precisely with each other. Rotation of the patient could be limited to a few tens of degrees, which would allow for the panels to remain close to each other, reducing limited angle artifacts, albeit not entirely removing them.

6.4 Sensitivity

Before moving to image reconstruction, we consider a key performance characteristic any PET scanner: the sensitivity. We do this based on GATE Monte Carlo simulations of the NEMA test, and compare the default BGO WT-PET configuration to the Quadra. The spatial resolution and image quality are left for the following chapter, where they will be discussed in greater detail, after introducing the image reconstruction software. We do not consider the NECR or scatter fraction here, but those results can be found in our NEMA characterization study for the WT-PET [4].

To evaluate the sensitivity of the WT-PET and the Quadra, a 70 cm long ^{18}F line source with an activity of 1 MBq is placed at the center of each scanner, as well as offset by 10 cm in the radial direction. The line source is surrounded by five concentric aluminum sleeves of the same length. Given the non-cylindrical configuration of the WT-PET system, the line source is offset along two different radial directions, once parallel to the detector panels (along the x-axis) and once toward the detector panels (along the y-axis). Since both the WT-PET and the Quadra have an AFOV of 106 cm, all simulations are additionally repeated with a 106 cm line source. We note that the physical Quadra originally used a maximum ring difference (MRD) cut, where any coincidences between detectors further than 85 rings apart would be discarded. We therefore consider both the Quadra with this MRD cut (MRD 85), and without it (MRD 322).

The obtained sensitivity profiles for both scanners are shown in Figure 6.3, and the total sensitivities are reported in Table 6.2. For the WT-PET, a total sensitivity of 154.0 cps/kBq is obtained for the centered 70 cm line source. This means that 15.4 % of all emissions are detected. Using MRD 322, the Quadra reaches a higher sensitivity than the WT-PET (179.7 cps/kBq), but it is severely reduced when the MRD cut is applied

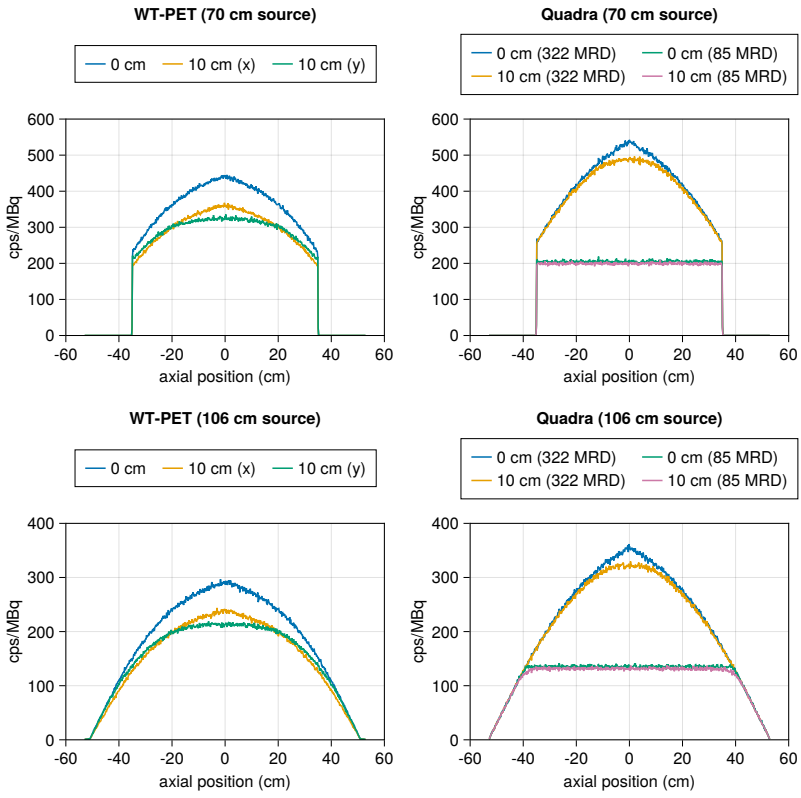


Figure 6.3: Sensitivity profile of the WT-PET system compared to the Quadra.

(87.0 cps/kBq). We also note that the values for the Quadra are in good agreement with the experimentally observed values [21].

6.5 Conclusion

The WT-PET therefore offers a unique combination of high spatial resolution (through DOI capable monolithic detectors), high sensitivity (due to long AFOV) and high throughput (enabled by standing patient design). The price is considerably reduced compared to conventional long AFOV scanners, by lowering the detector costs with the flat panel design and the use of BGO or sparse LYSO.

This design brings with it new challenges, such as the complex calibra-

line source	sensitivity (cps/kBq)		
	WT-PET	**Quadra (MRD 85)**	**Quadra (MRD 322)**
70 cm (center)	154.0	87.0	179.7
70 cm (x = 10 cm)	127.1	84.5	173.2
70 cm (y = 10 cm)	124.4	-	-
106 cm (center)	117.2	75.6	137.8
106 cm (x = 10 cm)	97.0	73.5	133.4
106 cm (y = 10 cm)	96.6	-	-

Table 6.2: Total sensitivity values for the WT-PET and the Quadra.

tion procedure required for monolithic detectors. Calibration is required not only for time resolution (as covered in part I), but also spatial resolution (which has been studied previously in our group [49], [100]). As the system is being built, detector calibration will therefore be one of the first steps towards making the system operational.

Another challenge concerns the image reconstruction. The departure of a cylindrical geometry based on pixelated detector technology makes existing reconstruction software unsuitable for our system. The remaining chapters will therefore focus on the development of an image reconstruction framework for the WT-PET, and an evaluation of the attainable image quality. The framework is introduced in Chapter 7, and is expanded upon to include methods for CT-less scatter correction (Chapter 8), CT-less attenuation correction (Chapter 9), and an approach to deal with limited angle artifacts arising from the flat panel geometry (Chapter 10).

Chapter 7

Image reconstruction

7.1 Introduction

The configuration of the Walk-Through PET is unique, both in terms of its flat panel geometry and the use of monolithic detector technology. This requires some adaptations to the image reconstruction procedures used in conventional cylindrical PET scanners, based on pixelated detectors. While open-source software for PET image reconstruction exists, such as for example CASToR (Customizable and Advanced Software for Tomographic Reconstruction) [134] or STIR (Software for Tomographic Image Reconstruction) [135], such packages would likely require a large amount of customization, additions and/or trade-offs to work well with the Walk-Through PET. Furthermore, performance can be suboptimal due to CPU-based implementations, with often limited support for GPU.

For this reason, we have opted to develop our own image reconstruction software, optimized for non-conventional geometries with continuous LOR representation. The software is written in Julia, a high-level, just-in-time (JIT) compiled programming language with excellent performance, on par with C++. The more computationally expensive parts of image reconstruction are implemented on the GPU, making use of Nvidia's CUDA toolkit. This includes the ray tracing algorithm, but also some other components of the system matrix, or necessary conversions between list-mode and sinogram data. Other computationally intensive procedures less suited for GPU are instead parallelized on the CPU. While the reconstruction code is developed with the WT-PET in mind, it makes little assumptions about the geometry and can therefore be used for other, more conventional scanners as well (e.g., for the Quadra).

This chapter introduces the reconstruction framework, which is then used to evaluate the imaging performance of the WT-PET, based on simulated acquisitions. To this end, we make extensive use of the NEMA IQ phantom for quantitative evaluation, and the anthropomorphic XCAT phantoms for qualitative assessment. These results are also compared with simulations and reconstructions for the Quadra. We also explore how some choices made for the image reconstruction framework affect imaging performance. The general reconstruction framework and simulation methodology are presented in this chapter, with further chapters focussing on the development of extensions to the reconstruction framework dealing with some more advanced data corrections. Therefore, in the methodology of later chapters, only specific modifications or additions to the reconstruction or simulation will be stated, to avoid redundancy. The results in this chapter also provide a reference point for comparison with the subsequent Chapters 8, 9 and 10.

7.2 Reconstruction framework

7.2.1 Update equation and LOR representation

The iterative MLEM update equation presented earlier (equation 2.3) is valid for sinogram data, where the vector $\mathbf{y}$ contains the number of occurrences for each LOR in the measured data. When the number of potential LORs is however very large (e.g., larger than the number of measured LORs), this results in a sparse vector $\mathbf{y}$, where most LORs are not measured at all. This can be very inefficient, both in terms of memory requirements and computational speed. In these cases, storing data in list-mode (LM) format is often preferred, where each line of data contains a single measured LOR. Equation 2.3 can be extended for list-mode data, and explicitly writing the summations over the matrix elements, it becomes:

$$\lambda_j^{(k+1)} = \frac{\lambda_j^{(k)}}{\sum_{i\in I} H_{ij}} \sum_{i\in E} \frac{H_{ij}}{\sum_{j'\in J} H_{ij'}\lambda_{j'}^{(k)} + e_i} \tag{7.1}$$

where the sum over I is a sum over all potential LORs and the sum over E is a sum over the detected LORs. The sum over J is a sum over the voxels.

A consequence of using monolithic detectors with continuous event positioning, is that we can not represent a LOR by only the two detector

IDs. In theory, it would be possible to discretize the detectors into virtual detector elements and perform image reconstruction as it is done for systems with pixelated detectors. However, given our high detector spatial resolution (1.3 mm FWHM for 2D and 2 mm FWHM for DOI resolution), the discretized elements would need to be very small (e.g. 0.5 x 0.5 x 1 mm^3) to limit the loss of system resolution. This would require an unfeasible number of virtual detector elements and potential LORs ($\sim$ 90 x 10^6 virtual detectors and $\sim$ 2 x 10^{15} unique LORs for the full system with 280 monolithic detectors per panel). Therefore, LORs are instead directly represented by a pair of 3D coordinates, saved in list-mode format. This continuous LOR representation also results in an infinite number of potential LORs I, which must therefore be sampled somehow in order to obtain the sensitivity image.

We opt to use a virtual scanner method, where LOR endpoints are uniform randomly sampled from the surface of a virtual geometry circumscribing the WT-PET, see Figure 7.1. Similar methods have been used before to store the system matrix of non-conventional or DOI capable scanner geometries [136], [137]. The virtual geometry is a cylinder of height equal to the scanner AFOV, circumscribing the physical scanner, with hemispherical caps on both ends. Note that we could not use a virtual inscribed cylinder, as the panel width is larger than the distance between both panels, and therefor not all LORs would fall within this smaller cylinder. The use of a circumscribed cylinder also necessitates the use of endcaps (or alternatively, sufficiently extending the height of the cylinder) in order to include very oblique LORs, as shown in Figure 7.1 on the right. One benefit of using hemispherical endcaps rather than extending the virtual cylinder height is that completely non-cylindrical geometries could also be modelled, e.g., when an additional detector panel would be placed above the head. When sampling, only LORs that intersect the real scanner geometry on both ends are included in I, whereas others are discarded. This sampling method, as opposed to directly sampling LOR endpoints from the detectors themselves, simplifies the computation of the system matrix, as will be made clear in Section 7.2.2. It is also easily extendable to other scanner geometries, and we use the same method to perform reconstructions of the Quadra system.

7.2.2 System response matrix

The system response matrix (SRM) can, to good approximation, be represented as a factorization of different, individual effects. These ef-

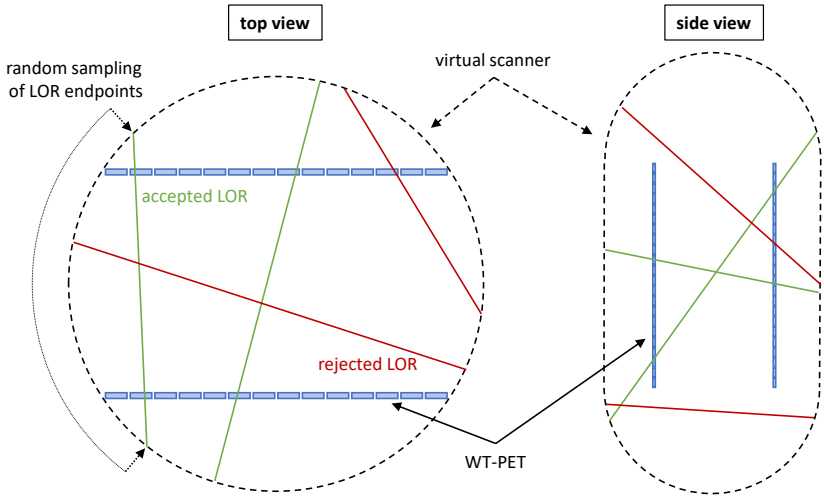


Figure 7.1: Method used for generating random LORs I, making use of a virtual scanner encasing the scanner geometry, from which endpoints are sampled.

fects and their contribution to the SRM can be calculated using either analytical, numerical, or empirical techniques, or a combination thereof. Our approach to modelling the SRM is to model as many effects as possible using analytical or efficient, non Monte Carlo, numerical techniques. These can be calculated exactly for every individual LOR without the need for LOR binning or discretization, and are not influenced by noise due to limited count statistics of empirical or Monte Carlo techniques. In our model, the system response matrix is represented as a factorization of a geometric sensitivity, a TOF weighting factor, an attenuation correction factor, a detector attenuation and a normalization factor:

$$H_{ij} = H_{geo,ij} H_{TOF,ij} H_{atn,i} H_{det,i} H_{norm,i} \tag{7.2}$$

which will all be explained in more detail in the following paragraphs. The subscripts i and ij indicate if the factor depends only on the LOR i, or on both the LOR i and the voxel j.

Geometric sensitivity

$H_{geo,ij}$ is the geometric sensitivity, representing the probability that an emission from voxel j will be emitted along LOR i. It is proportional

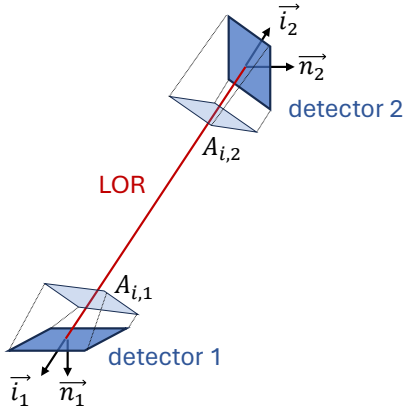


Figure 7.2: Visualization of the projected detector areas $A_{i,k}$.

to the intersection length l_{ij} of the LOR with the voxel, and the solid angle by which the LOR endpoints are seen from voxel j. Disregarding TOF information, for a LOR spanning two detectors 1 and 2, it can be computed as [138]:

$$H_{geo,ij} = \frac{l_{ij}}{V_j} \frac{A_{i,1} A_{i,2}}{2\pi r_i^2} \tag{7.3}$$

where V_j is the volume of the voxel, $A_{i,k}$ is the area of the detector element k projected on a plane perpendicular to the LOR (see Figure 7.2), and r_i is the distance between both detector elements. Here, the detector elements are virtual elements on the circumscribed cylinder. Their position is known for the sampled LORs I and easily computed analytically for the detected LORs E.

Due to the uniform sampling of the endpoints of LORs I on the cylinder's surface, the projected areas $A_{i,k}$ (both for sampled LORs I and measured LORs E) can be calculated as:

$$A_{i,k} = (\vec{i}_k \cdot \vec{n}_k) \frac{S_{tot}}{\sqrt{N}} \tag{7.4}$$

where $\vec{i}_k$ and $\vec{n}_k$ are normalized vectors representing the direction of LOR i and the outward unit normal of the virtual scanner geometry at the intersection point, respectively. S_{tot} is the total surface of the virtual geometry, and N is the total number of sampled LORs, both rejected and accepted. Therefore, $\sqrt{N}$ is the total number of virtual detector

elements and $S_{tot}/\sqrt{N}$ represents the area of a single virtual element. We note that the solid angle is independent of the emission voxel, and can therefore be pre-computed once per LOR. The simplicity of equation 7.4 also shows why we opt to use a virtual scanner geometry for LOR sampling. Calculation of the solid angle would be much more complex if LORs were sampled directly from the physical scanner, where we would need to take into account the possibility of LORs entering through the side surfaces of the detectors. Inclusion of DOI would also result in one LOR hitting multiple virtual detector elements (stacked on top of each other), further complicating matters.

Attenuation correction factor

The attenuation correction factor $H_{atn,i}$ represents the probability that neither gamma photon in the coincidence event is attenuated by the object, and both can therefore reach the detectors. With object attenuation μ, it is calculated as a discretized line integral along the LOR:

$$H_{atn,i} = e^{-\sum_{j \in J} l_{ij} \mu_j} \tag{7.5}$$

Here, the attenuation coefficient μ represents the total attenuation, which includes both attenuation due to photoelectric absorption and attenuation due to scattering. Although the calculation of $H_{atn,i}$ depends on the voxels j, the final result $H_{atn,i}$ is independent of j and can therefore again be precalculated once per LOR.

Detector attenuation

The detector attenuation $H_{det,i}$ is the probability that, upon reaching the detector faces, both gamma photons will be attenuated by the detectors. It is dependent on the intersection lengths of the LOR with the detectors on both ends, and the attenuation coefficient of the scintillation material. We calculate it as:

$$H_{det,i} = \left(1 - e^{-\mu_{det} l_{i,det_1}}\right) \left(1 - e^{-\mu_{det} l_{i,det_2}}\right) \tag{7.6}$$

Here, the intersection length l_{i,det_k} is the entire intersection length of the extended LOR through all detectors on one side of the emission. Even if the gamma interaction is measured at the top surface of the first detector intersecting the LOR, l_{i,det_k} is still computed taking the entirety of all remaining intersecting detectors into account. The coefficient μ_{det}

is the total (Compton + photoelectric) attenuation coefficient of the scintillation material. This means that $H_{det,i}$ represents the probability that both gamma photons interact in the detectors, which does not necessarily imply detection. Some Compton scattered interactions may still be absorbed by the photoelectric effect within the same monolithic crystal, and therefore fall within the energy window, whereas others will be scattered out of the crystal and deposit too little energy for detection. We are therefore overestimating the detector sensitivity with equation 7.6, which we will correct for in the normalization factor. Although this further correction is still required to obtain correct activity values, inclusion of $H_{det,i}$ in the system model eliminates many artifacts visible prior to normalization, especially in sparse configurations. Furthermore, the factor $H_{det,i}$ can be calculated exactly for each LOR and is noise-free, whereas normalization factors need to make use of LOR binning or discretization, and include noise due to the nature in which they are obtained.

Calculation of the detector intersection lengths l_{i,det_k} is done with a ray marching (also known as sphere tracing) approach [139], see Figure 7.3. At a certain starting point (e.g., the center of the LOR or the most likely annihilation position when TOF is available), the nearest distance to any detector in any direction is computed, representing the minimum distance d that may be traveled without hitting a detector. A new point is then selected at a distance d along the LOR, repeating the process until d reaches a preset minimum value or step length, at which point the detector surface is reached. In our work, we use a minimum step length of 0.1 mm. The algorithm can then be continued through the detector, recording the traveled distance to obtain the intersection lengths l_{i,det_k}.

This algorithm makes use of signed distance functions (SDFs), which provide an analytical solution for the minimum distance between a point in 3D space and primitive 3D shapes. For example, the SDF of a box (the shape of most other PET detectors) is given by:

$$\begin{aligned} \mathbf{Q} &= |\mathbf{P} - \mathbf{C}| - \frac{\mathbf{W}}{\mathbf{2}} \\ d &= ||\max(\mathbf{Q}, \mathbf{0})||_2 + \min(||\mathbf{Q}||_\infty, \mathbf{0}) \end{aligned} \tag{7.7}$$

representing the distance d from point $\mathbf{P}$ to a box, oriented along the primary axes, with center $\mathbf{C}$ and dimensions $\mathbf{W}$. The absolute value $|\mathbf{P}-\mathbf{C}|$ and maximum $\max(\mathbf{Q}, \mathbf{0})$ are taken element wise. Other arbitrary orientations of the shape can be included by first performing a coordinate transformation of the point $\mathbf{P}$. The term signed in SDF refers to the fact

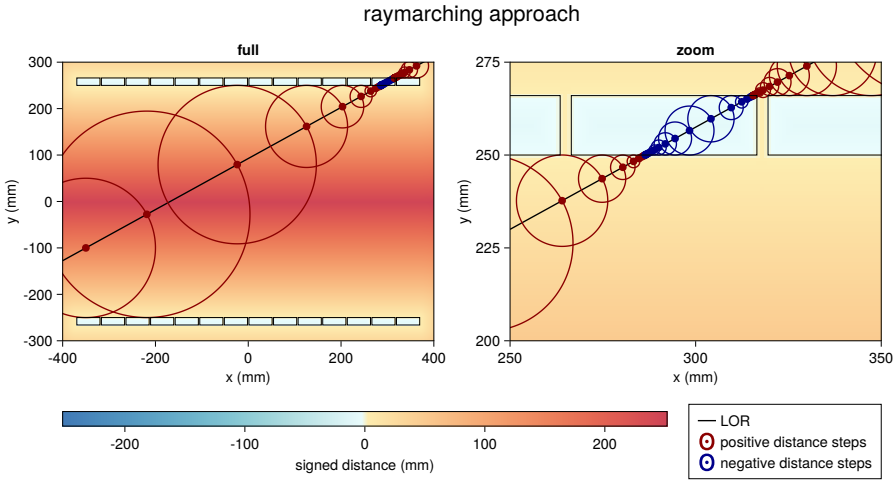


Figure 7.3: Visualization of the ray marching approach used to calculate the intersection lengths of a LOR with the detectors.

that equation 7.7 is negative for points inside the shape, which provides an easy way of accumulating the detector intersection lengths while traversing the LOR. Furthermore, a convenient property making this algorithm very performant and easily scalable to any number of detectors, is that the SDF of a collection of periodically repeated (translated and/or rotated) primitive shapes can be computed directly from the SDF of a single shape, using only a few modulo operations. Therefore, there is no need to loop over the different detectors in the geometry in order to compute the global SDF of the scanner. This same methodology is also used when sampling the LORs I in Section 7.2.1, to check which LORs intersect the physical scanner geometry.

TOF weighting factor

The TOF weighting factor $H_{TOF,ij}$ is a correction factor to be applied in case of TOF-enabled reconstruction, and is simply 1 in case of non-TOF reconstruction. It could be considered to be a part of the geometric sensitivity, but we have kept it separate here to easily distinguish between TOF and non-TOF reconstructions. Under the assumption that the time resolution of a single detector is Gaussian, which is a good approximation when the timing capabilities are not influenced by Cerenkov radiation,

the coincidence time resolution will also be Gaussian distributed and $H_{TOF,ij}$ can therefore be calculated as:

$$H_{TOF,ij} = \frac{1}{\sigma_x \sqrt{2\pi}} \exp\left(-\frac{1}{2}\frac{\Delta x_{ij}^2}{\sigma_x^2}\right) \tag{7.8}$$

where σ_x is the standard deviation of the TOF kernel measured in distance, rather than time. It is obtained by multiplying σ_t (the standard deviation of the TOF kernel in time) by a factor of $c/2$, with c the speed of light. Δx_{ij} is the distance from voxel j to the center of the TOF kernel for LOR i, i.e., the distance to the most likely annihilation point.

Just as for the LOR endpoints, the measured TOF differences are stored continuously, rather than using discretized TOF bins. We note that the TOF weighting factor does not need to be included in the calculation of the sensitivity image. In theory, for TOF reconstruction the sum over LORs I requires an additional sum over all possible TOF bins for each LOR. However, only the TOF weighting factor depends on the TOF bin, and the sum of TOF weighting factors over all TOF bins equals 1. This additional sum and the inclusion of $H_{TOF,ij}$ can therefore be skipped when generating the sensitivity image. This is of course not the case for measured LORs E, as the LOR is only measured at a specific TOF bin.

Normalization factor

The normalization factor $H_{norm,i}$ is a correction factor which accounts for additional effects not yet modelled by the system matrix, and can be derived from e.g. Monte Carlo simulation or experiment. In our case, it is primarily used to account for the overestimation of the detector sensitivity in $H_{det,i}$, and is obtained from Monte Carlo simulation in GATE. In a more realistic, experimental setting, it could also be used to account for differences in efficiency between different detectors and possible saturation effects due to dead-time. We compute $H_{norm,i}$ by forward projecting a known activity and attenuation distribution through the SRM (without normalization or TOF), and comparing it to the measured data $y_{i,MC}$ obtained from the Monte Carlo simulation:

$$H_{norm,i} = \frac{y_{i,MC}}{\sum_{j \in J} H_{geo,ij} H_{atn,i} H_{det,i} \lambda_j} \tag{7.9}$$

In this way, the normalization factor is used to account for, and only for, those effects that were not yet included in the SRM.

Here, we do run into an issue with continuous LOR representation, as none of the LORs measured during the Monte Carlo simulation will exactly match the sampled LORs I or measured LORs E. We therefore need to resort to LOR discretization. We opt to discretize LORs into sinogram bins, represented by 4 coordinates: the transverse radius r, the azimuthal angle ϕ, the axial offset z and the polar angle θ. The fraction in equation 7.9 becomes a division between counts in a specific sinogram bin, and $H_{norm,i}$ for specific LORs $i \in I$ and $i \in E$ is then obtained by interpolation of the resultant normalization sinogram.

7.2.3 Moving geometries

The image reconstruction should also be capable of dealing with a moving scanner geometry. In conventional PET systems, this would be used to image at different bed positions with for example continuous bed motion, where from the patient's point of view (and from the reconstruction FOV's point of view), the scanner is being translated axially. For the WT-PET geometry, we also wish to include rotational movement of the panels or patient around the central axis. The reconstruction code can be extended to a moving geometry with some minimal modifications.

The LORs I are still sampled from the fixed geometry, and elements of the SRM that do not depend on the voxel j (and are therefore independent of the patient position with respect to the scanner) can still be pre-computed, regardless of scanner motion. Importantly, this includes the normalization sinogram $H_{norm,i}$, as otherwise a new normalization scan or simulation would be required for every possible movement pattern. The sensitivity image itself, which does depend on scanner motion, is then computed by randomly sampling a new scanner position for each LOR $i \in I$, and translating/rotating the LOR accordingly, prior to calculation of the voxel dependent SRM elements. Similarly, for the measured LORs $i \in E$, prior to calculation of the SRM elements, these are translated/rotated according to the current scanner position, derived from the timestamp attached to each LOR.

7.3 Simulation study

7.3.1 Simulation parameters

We use the GATE Monte Carlo software to acquire realistic simulations of the PET acquisition process [26]. Unlike part I, optical photon simulation

is now turned off to speed up computations. Sources are simulated as positrons with energy corresponding to ^{18}F, so that the simulation includes the positron range and photon non-collinearity physics. The photoelectric effect, Compton and Rayleigh scattering are all enabled, both within the phantom and within the detectors. The grouping of individual hits (gamma interactions in the detectors) into singles, and the coincidence sorting of these singles, are both handled by GATE. For the WT-PET, this grouping is done over each monolithic detector, whereas for the Quadra, it is done over a detector block consisting of 10 x 20 pixels. Energy blurring of the singles is also done in GATE, where we use a Gaussian energy resolution of 15% for BGO and 11% for LSO. A paralyzable dead time on the detector level is also included, which is set to 320 ns in the Quadra and 370 ns in the WT-PET.

Detector spatial blurring is done as a post-processing step after the Monte Carlo simulation. GATE provides the exact gamma interaction positions in the case of purely photoelectric detections, or an energy weighted position in the case of detections with one or more Compton scatters. For the WT-PET, which has DOI capable monolithic detectors with continuous event positioning, this interaction point is then blurred with a Gaussian in 3D. The values for smearing are equal to the average detector resolution is each direction, namely 1.3 mm FWHM parallel to the front surface and 2 mm FWHM along the DOI [132]. In case the blurring would position the interaction point outside the detector, the point is instead placed at the very edge. For the Quadra, where the resolution is limited to the pixel size, spatial blurring is done by positioning the interaction point in the central plane of the pixel, and uniformly randomizing its 2D position within that plane. The randomization is done to counteract artifacts that may arise from the otherwise discretized interaction positions.

Time blurring is also done post simulation, by smearing the TOF difference of the coincidences using a Gaussian with FWHM equal to the CTR (400 ps for the WT-PET and 228 ps for the Quadra). It is done post simulation so that we can easily investigate different TOF resolutions without needing to re-run the simulation or needing to save the singles. In theory, this affects the coincidence sorting procedure by not performing time blurring at the singles level, prior to sorting. However, the overall statistics should remain unaffected as long as the CTR is sufficiently small compared to the CTW.

Unless otherwise specified, we always use **30 s acquisitions** in our

simulations, in this and future chapters, given that we are aiming for sub-one minute scan times in the WT-PET.

7.3.2 Phantoms

Point sources

To investigate the spatial resolution, six ^{18}F point sources with a diameter of 0.5 mm and an activity concentration of 15.28 MBq/ml (30 s acquisitions) are simulated at the positions proposed by NEMA, see Figure 7.4. Three points are located at the central transverse slice of the scanners (1, 10, and 20 cm in the radial direction), and another three at 1/8 of the AFOV. Since we are making use of iterative reconstruction, these point sources are embedded in a warm background, as we would otherwise overestimate the spatial resolution. For conventional PET scanners with cylindrical geometry, these coordinates provide an estimation of the spatial resolution for any angle around the scanner axis. Given the unique configuration of the WT-PET, however, an additional series of six ^{18}F point sources are simulated, with similar radial and axial arrangements, but rotated 90 degrees around the scanner axis. The spatial resolution along each axis is then calculated by drawing line profiles through the reconstructed point sources, and reporting the FWHM and FWTM (full width at tenth maximum). The point source positions, as well as the axis definitions for the scanners.

Uniformity phantom

Before investigating more complex phantoms, we wish to study the uniformity of the reconstructions, with and without normalization. This gives us an idea of how well the system response matrix describes the actual scanner, and makes it easier to spot any artifacts arising from incorrect modelling. For the uniformity phantom, we use a rounded box, 80 cm in height, 40 cm in width, and 30 cm thick, filled uniformly with ^{18}F in water. As the activity concentration is uniform, ideally, so should be the reconstruction. We will test the uniformity phantom on the sparse WT-PET configuration, as we expect most issues with non-uniformity there.

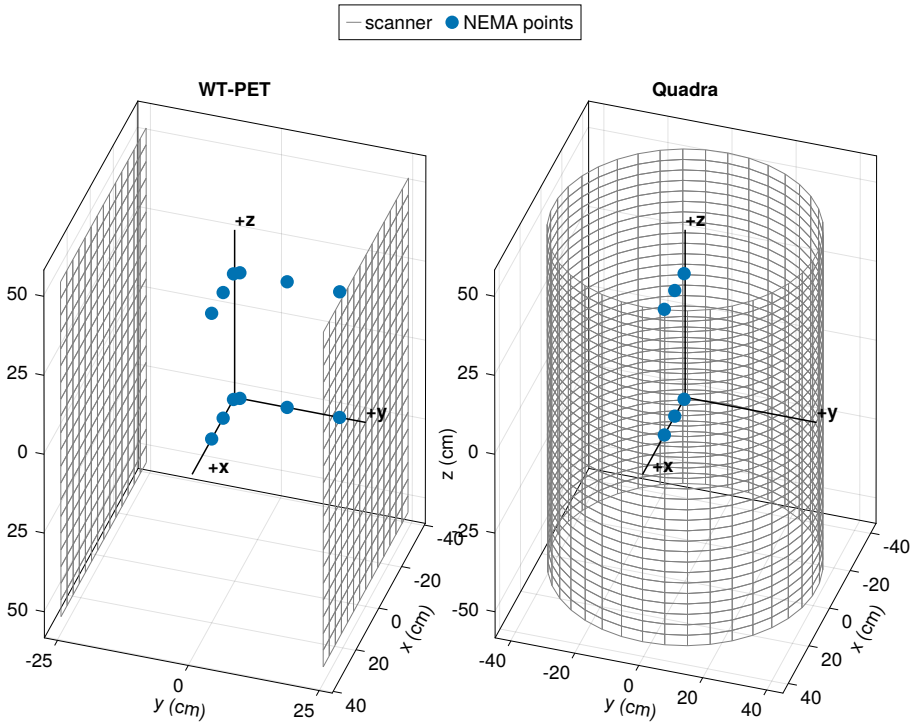


Figure 7.4: NEMA point source positions and scanner axis definitions for the WT-PET and Quadra.

IQ phantom

Next, we have the NEMA IQ phantom as a standardized test to evaluate image quality, see Figure 7.5. It contains a central lung insert surrounded by six spheres of varying diameter (10 mm - 37 mm), and is filled with a background activity concentration of 5.3 kBq/ml, using a sphere-to-background activity concentration ratio of 4:1. It is positioned so that the spheres coincide with the central transverse slice (z = 0) of the scanner.

For each sphere k, a 2D region of interest (ROI) is drawn around the sphere in the central transverse slice. In addition, for each sphere, this ROI is repeated 60 times in the background of the IQ phantom, positioned according to the NEMA guidelines [24]. Figure 7.5 shows the activity map, attenuation map and ROI positions for the IQ phantom. Only 12 background ROIs are shown, as they are repeated at -2, -1, +1 and +2 cm along the axial direction, for a total of 60 ROIs. We then

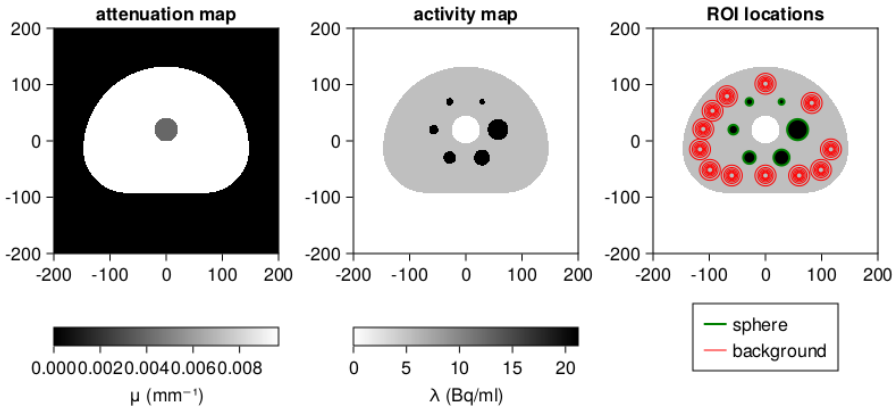


Figure 7.5: Activity and attenuation distribution for the NEMA IQ phantom, with visualization of the sphere and background ROIs.

calculate the contrast recovery coefficient (CRC) for each sphere k as:

$$\mathrm{CRC}_k = \frac{\frac{\mu_{H,k}}{\mu_{B,k}} - 1}{\frac{a_H}{a_B} - 1} \tag{7.10}$$

where a_H and a_B are the known activity concentrations of the spheres and background respectively, so that in our case $a_H/a_B = 4$. $\mu_{H,k}$ represents the average counts in the ROI of sphere k, and $\mu_{B,k}$ the average counts over all background ROIs for sphere k. CRC values closer to 1 imply a better recovery of the exact activity (contrast) in each sphere.

We do not yet include any metrics that specifically measure the noise contribution, as we will reconstruct using true coincidences only. We will include these metrics in the following chapter on scatter correction, where reconstruction is done using prompt coincidences.

XCAT phantom

Finally, we use a humanoid XCAT phantom for a more qualitative evaluation of the image quality with more complex features. Anthropomorphic XCAT phantoms are advanced computational models that simulate realistic human anatomy for various medical imaging studies [58]. These provide a valuable tool for assessing imaging system performance, image reconstruction algorithms, and optimizing imaging protocols. These phantoms incorporate anatomical details such as organ shapes, sizes,

and tissue densities, allowing for accurate simulations of different patient populations. The specific phantom we use in this chapter is a female of age 52, with a height of 179 cm and a weight of 72 kg (BMI of 22.47). We use a scan dose of 3 MBq/kg. The phantom is positioned such that the top of the head is 5 cm below the top of the scanner AFOV. For reference, the ground-truth activity map used for the simulation is given in Figure 7.6. The XCAT phantoms used for evaluation in the following chapters are also displayed here.

7.3.3 Reconstruction parameters

Images are reconstructed using the aforementioned reconstruction procedure. The normalization scan for the computation of $H_{norm,ij}$ is obtained from a GATE simulation of a uniform source in air. In the computation of the normalization factors, we use relatively large sinogram bins corresponding to a total of $\sim$ 900 million discretized LORs. Images are reconstructed using 2 x 2 x 2 mm^3 voxel sizes, except for the point source reconstructions, where we use 0.5 x 0.5 x 0.5 mm^3 voxel sizes. We do not use subsets in the reconstructions, and visualize images at the tenth iteration. The tenth iteration was chosen based on the convergence of the spatial resolution, see Figure 7.7 in the results. For attenuation correction, the ground truth attenuation map at 511 keV is used. Random and scattered coincidences are removed from the data prior to reconstruction, making use of the tags available from the GATE simulation. Reconstruction is therefore done using true coincidences only, and the factor e_i is set to zero for all LORs.

7.4 Results

7.4.1 Spatial Resolution

We first investigate convergence of spatial resolution by measuring the spatial resolution for both the WT-PET and Quadra for two NEMA point source positions, in function of the number of iterations, see Figure 7.7. We note that in all cases, the spatial resolution has converged by the tenth iteration. Along with a visual assessment of image quality of other phantoms, we have deemed the tenth iteration to be a reasonable middle ground between the sharpness of, and noise level in, the image.

A comparison of the spatial resolutions for all NEMA point source positions, obtained at the tenth iteration and measured in FWHM and

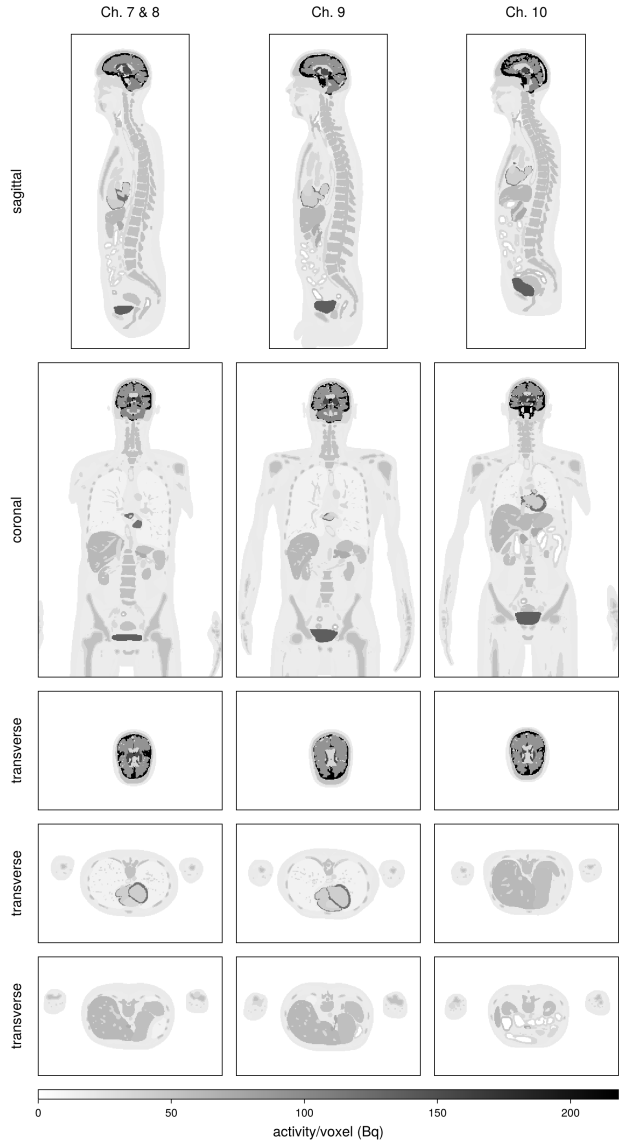


Figure 7.6: Ground-truth activity maps of the XCAT phantoms used in Chapter 7 and 8 (left), Chapter 9 (center) and Chapter 10 (right).

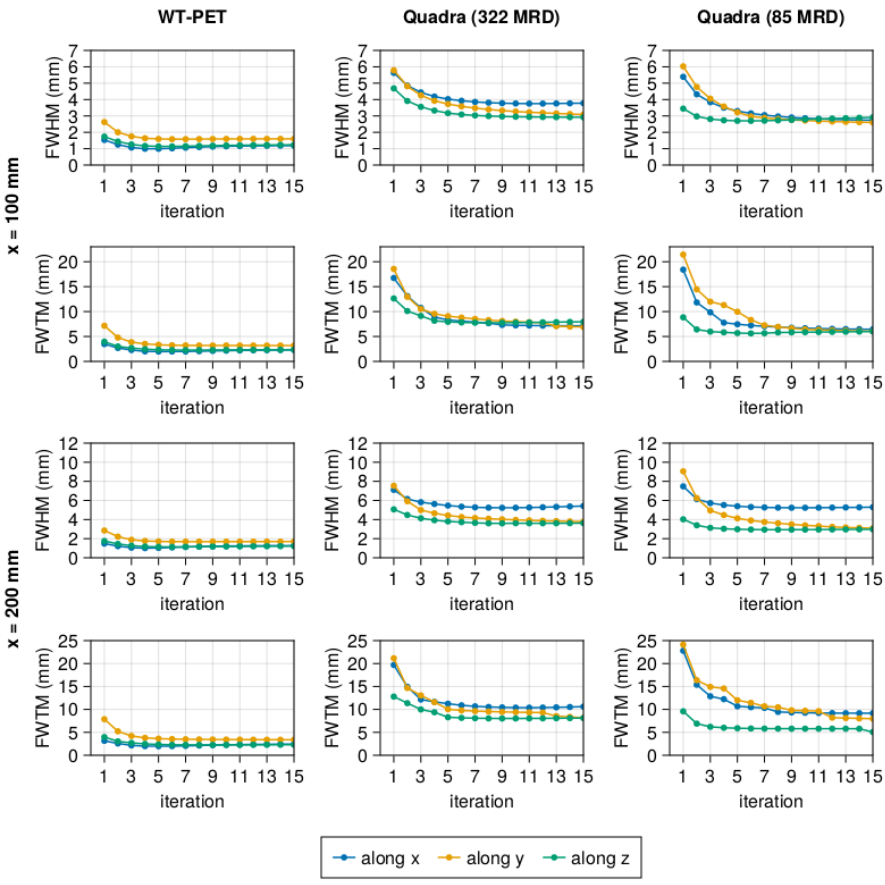


Figure 7.7: Spatial resolution measurements of the NEMA point sources at (10, 0, 0) cm and (20, 0, 0) cm, in function of iteration number.

FWTM, along the x (parallel to panels), y (toward panels) and z (along AFOV) axes, are shown in Table 7.1 for the WT-PET and the Quadra. We note that the spatial resolution for the WT-PET remains below 2 mm FWHM for the majority of the FOV, except along the y direction for points towards the edge of the AFOV. The spatial resolution of the Quadra, on the other hand, is consistently above 2 mm FWHM, and shows substantial degradation along the radial direction for points moving radially outward (in this case, along x).

We note that the spatial resolution results for the simulated Quadra system are considerably better than those reported experimentally [21],

scanner	position source (cm)	FWHM (mm)			FWTM (mm)		
		x	y	z	x	y	z
WT-PET	(1, 0, 0)	1.20	1.62	1.30	2.87	4.14	3.06
	(10, 0, 0)	1.16	1.90	1.12	2.81	4.83	2.89
	(20, 0, 0)	1.17	1.94	1.19	2.60	4.36	2.81
	(1, 0, 39.75)	1.14	2.13	1.24	2.86	5.22	3.20
	(10, 0, 39.75)	1.21	2.48	1.29	3.19	7.11	3.06
	(20, 0, 39.75)	1.26	2.52	1.05	2.08	5.58	2.08
	(0, 1, 0)	1.18	1.75	1.34	2.43	3.79	2.54
	(0, 10, 0)	1.22	1.92	1.33	2.54	4.20	2.54
	(0, 20, 0)	1.32	2.25	1.41	2.71	7.08	2.71
	(0, 1, 39.75)	1.29	2.24	1.16	2.58	5.71	2.59
	(0, 10, 39.75)	1.52	2.71	1.46	2.91	6.54	3.35
	(0, 20, 39.75)	1.65	3.36	1.88	3.76	8.18	3.96
Quadra MRD 322	(1, 0, 0)	2.55	2.62	2.85	5.68	6.45	7.03
	(10, 0, 0)	3.76	3.27	2.95	7.24	7.98	7.78
	(20, 0, 0)	5.24	3.97	3.62	10.39	9.39	8.03
	(1, 0, 39.75)	2.09	2.35	2.24	4.76	5.70	4.68
	(10, 0, 39.75)	2.78	3.14	2.89	5.85	7.64	5.03
	(20, 0, 39.75)	4.63	3.31	3.20	8.61	6.81	5.76
Quadra MRD 85	(1, 0, 0)	2.62	2.70	2.72	5.02	5.66	4.82
	(10, 0, 0)	2.86	2.72	2.78	6.66	6.43	5.86
	(20, 0, 0)	5.24	3.40	2.94	9.27	9.71	5.79
	(1, 0, 39.75)	2.22	2.44	2.40	4.80	5.71	4.84
	(10, 0, 39.75)	3.31	3.02	2.71	6.35	6.67	5.95
	(20, 0, 39.75)	4.29	2.66	2.23	8.55	10.78	5.82

Table 7.1: Spatial resolution measurements of simulated point sources at the NEMA positions for the WT-PET and Quadra. The x-axis is parallel to the WT-PET panels, the y-axis is toward the WT-PET panels, and the z-axis is along the axial direction of both scanners.

often $\sim$ 0.5 - 1 mm better in terms of FWHM. There are a few potential reasons for this: First, a different reconstruction algorithm was used. The experimental study used a 3D TOF direct inversion Fourier transform backprojection algorithm, an analytic reconstruction algorithm, whereas we used an iterative reconstruction algorithm. As mentioned, we use a warm background to counteract the usual overestimation of spatial resolution occurring in iterative algorithms, but some degree of

overestimation may nonetheless remain. Furthermore, we used a small voxel size (0.5 mm^3), whereas the experimental study used a voxel size of (1.65 mm^3). This can affect the measured spatial resolution, especially when the voxel size comes too close to the spatial resolution. In addition, certain settings used in the GATE digitizer chain, such as the energy winner-takes-all approach for deciding the gamma interaction position in events with multiple crystal interactions, may not entirely coincide with the reality. This, combined with the fact that the pixel with the highest energy deposit is known exactly in the simulation, may result in an overestimation of the resolution, even if we are applying detector spatial blurring within the pixels themselves. Finally, there may be a number of scanner non-idealities, such as slight detector misalignments, which contribute to spatial resolution degradation, but are not modelled in the simulation.

To show the impact that LOR discretization would have had on the spatial resolution, we re-reconstruct the NEMA point source at (1, 0, 0) cm, only this time we further subdivide the monolithic crystals into virtual pixels. After performing the intrinsic detector spatial resolution blurring for the WT-PET as usual, we assign the new, blurred LOR endpoints to the corresponding virtual pixel within the monolithic detector, and then randomize the position within that pixel. The randomization is again done in order to counteract artifacts that otherwise arise when simply assigning all LORs to the center of the pixel. Even though the reconstruction uses 3D coordinates as usual, we have concealed the continuous LOR positioning information. This is similar to what would happen if we were to discretize measured LOR endpoints in order to use reconstruction software developed specifically for pixelated detectors. The spatial resolution results are shown in Table 7.2, in function of iteration, for different choices of the virtual pixel size. We see that, even when going to a pixel size below the detector spatial resolution, a loss of system spatial resolution remains.

7.4.2 Uniformity

Figure 7.8 shows the reconstructions of the uniformity phantom for the sparse WT-PET configuration, once with and without inclusion of the detector attenuation factor H_{det}, and once with and without the normalization factor H_{norm}. We note that, when neither factor is included, a visible oscillating pattern becomes visible in the reconstruction. Inclusion of H_{det} removes this pattern, even prior to normalization. In

discretization	FWHM (x)	FWHM (y)	FWHM (z)
continuous	1.20 mm	1.62 mm	1.30 mm
0.625 x 0.625 x 1 mm^3	1.26 mm	1.93 mm	1.48 mm
1.25 x 1.25 x 2 mm^3	1.65 mm	2.08 mm	1.60 mm
2.5 x 2.5 x 4 mm^3	2.66 mm	3.92 mm	2.62 mm

Table 7.2: Spatial resolution measurements of the NEMA point source at (1, 0, 0) cm, 10th iteration, using different levels of LOR discretization.

fact, the pattern is still slightly visible when applying normalization without modelling the detector attenuation. This is likely an artifact of using insufficient statistics and/or sinogram bins when computing the normalization factors. At the edges, we can also see blurring occur, as a result of the limited projection angles.

7.4.3 IQ phantom

In Figure 7.9, we show the reconstructions of the IQ phantom for the three WT-PET scanner configurations and the Quadra (MRD 322), all for 30 s acquisitions. The tenth iteration is visualized, and the CRC is calculated for all spheres in function of iteration. Quantitatively, we obtain better results for most spheres in the WT-PET configurations than in the Quadra, especially for the rotating configuration and the smaller spheres. Here, the high spatial resolution of the WT-PET plays a large role in recovering the activity in the smaller spheres. Quantitatively, however, we can observe the limited angle artifacts in the default and sparse WT-PET configurations. We also notice that the reconstruction for the rotating WT-PET is noisier than the one for the Quadra, which is to be expected due to the lower sensitivity.

In Figure 7.10, we again show the tenth iteration of the reconstruction of the IQ phantom for the default WT-PET configuration, but we now vary the TOF, both in the post-processing and reconstruction. Comparing with the non-TOF reconstruction, we can see how much TOF helps in reducing the limited angle artifacts. We also observe that the CRC values for the non-TOF reconstruction are rather nonsensical: the reconstruction is no longer uniform throughout the background, resulting in CRC values that are heavily influenced by the placement of the background ROIs. To a lesser extent, this is still the case for the 400 ps TOF reconstructions, and should therefore be kept in mind when comparing CRC values of reconstructions with and without limited angle artifacts.

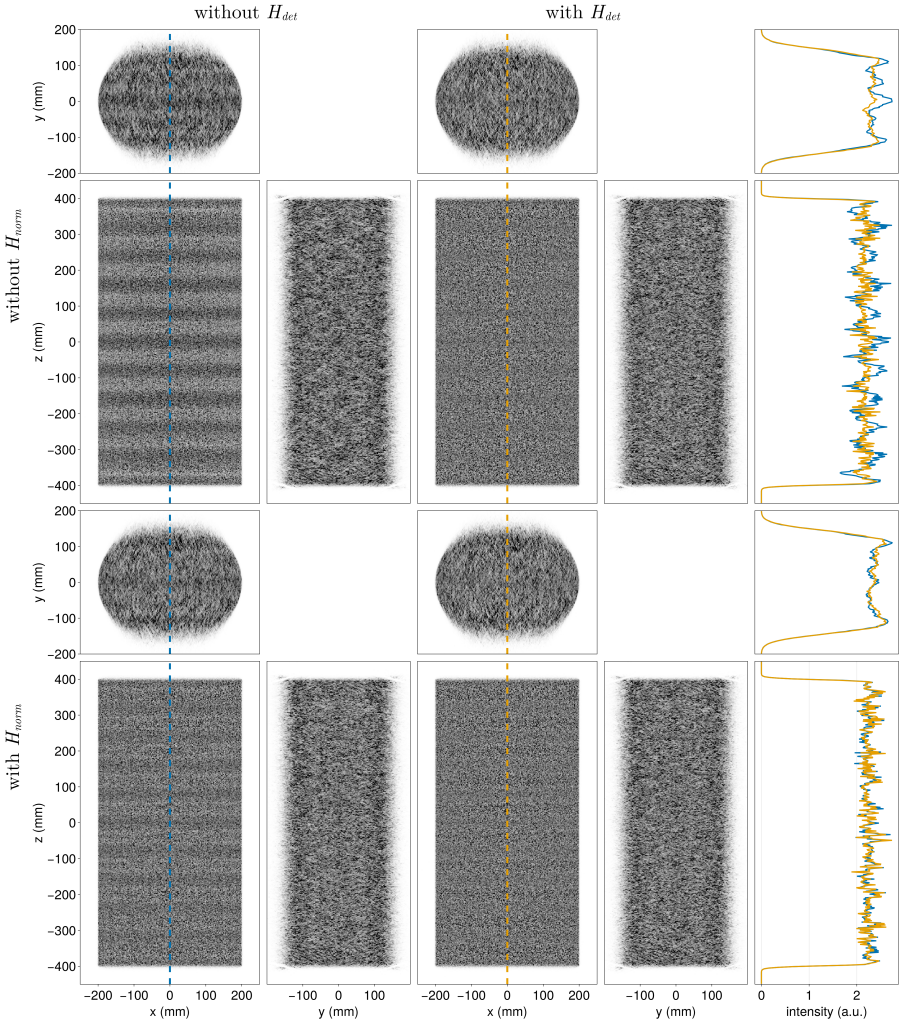


Figure 7.8: Reconstructions of the uniformity phantom in the sparse WT-PET, with and without modelling for the detector attenuation H_{det} and with and without inclusion of the normalization factor H_{norm}.

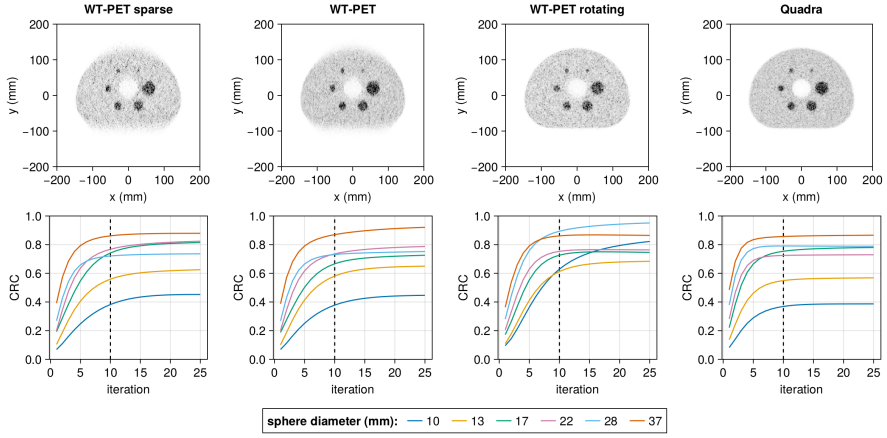


Figure 7.9: IQ reconstructions and sphere CRC values for the three WT-PET configurations and the Quadra. The tenth iteration is visualized.

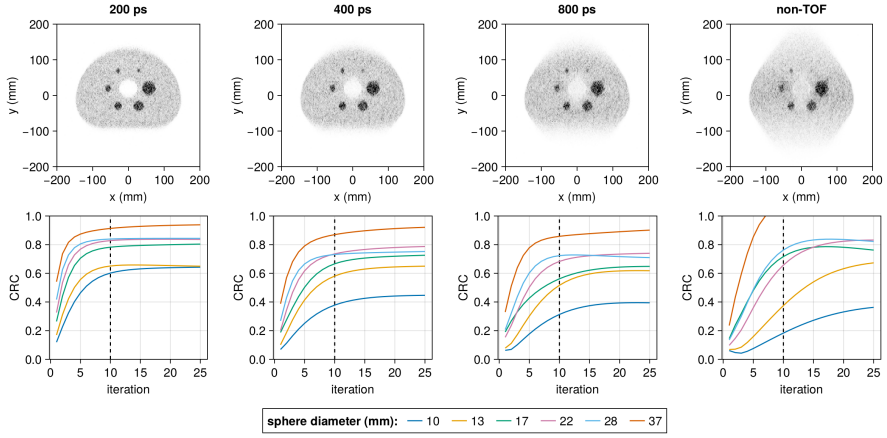


Figure 7.10: IQ reconstructions and sphere CRC values for the default WT-PET configuration with different TOF values (stated on top in FWHM). The tenth iteration is visualized.

7.4.4 XCAT phantom

Finally, for a visual assessment, the reconstructions of the XCAT phantom are shown in Figure 7.11.

7.5 Discussion

7.5.1 Reconstruction quality

The point source reconstructions show one of the biggest advantages of using monolithic detectors: we not only obtain very good spatial resolution, but the resolution is also very uniform over the FOV due to the DOI capabilities of the detectors. We do observe some degradation along the y-direction, especially for those points located at 1/8 of the AFOV. This can primarily be attributed to the limited projection angles, causing a blurring towards the panels, as was also observed in the other phantoms' reconstructions. Fewer projection angles are available for points further away from the center of the scanner (see also Chapter 10), causing the observed pattern of degradation. Compare this to the Quadra results, where we observe a considerably larger degradation of spatial resolution measured along the x-direction, for points moving along the x-axis (i.e., radially outward). This is caused by the parallax effect, due to the non-DOI capable, pixelated detectors. The spatial resolution results for the Quadra in both MRD modes are very similar, although slightly better for MRD 85. This is due to more oblique LORs being measured in MRD 322, which exacerbates issues due to lack of DOI. A more comprehensive analysis of the spatial resolution, with more point source positions for both scanners and therefore better mapping of the spatial resolution in function of position, can be found in the NEMA study for the WT-PET [4]. We also note that the results for the Quadra in Table 7.1 are slightly better than the experimentally reported spatial resolution results [21]. Even though we have used a warm background, this is likely due to the use of an iterative reconstruction algorithm rather than FBP. The results for the WT-PET could therefore also be overestimated compared to what would have been obtained using FBP.

Table 7.2 also showed why it is important to perform the reconstructions using a continuous LOR representation, in order to fully leverage the high spatial resolution provided by the monolithic detectors. We would need to discretize LORs using very small virtual detectors in order to minimize spatial resolution losses, resulting in too many potential

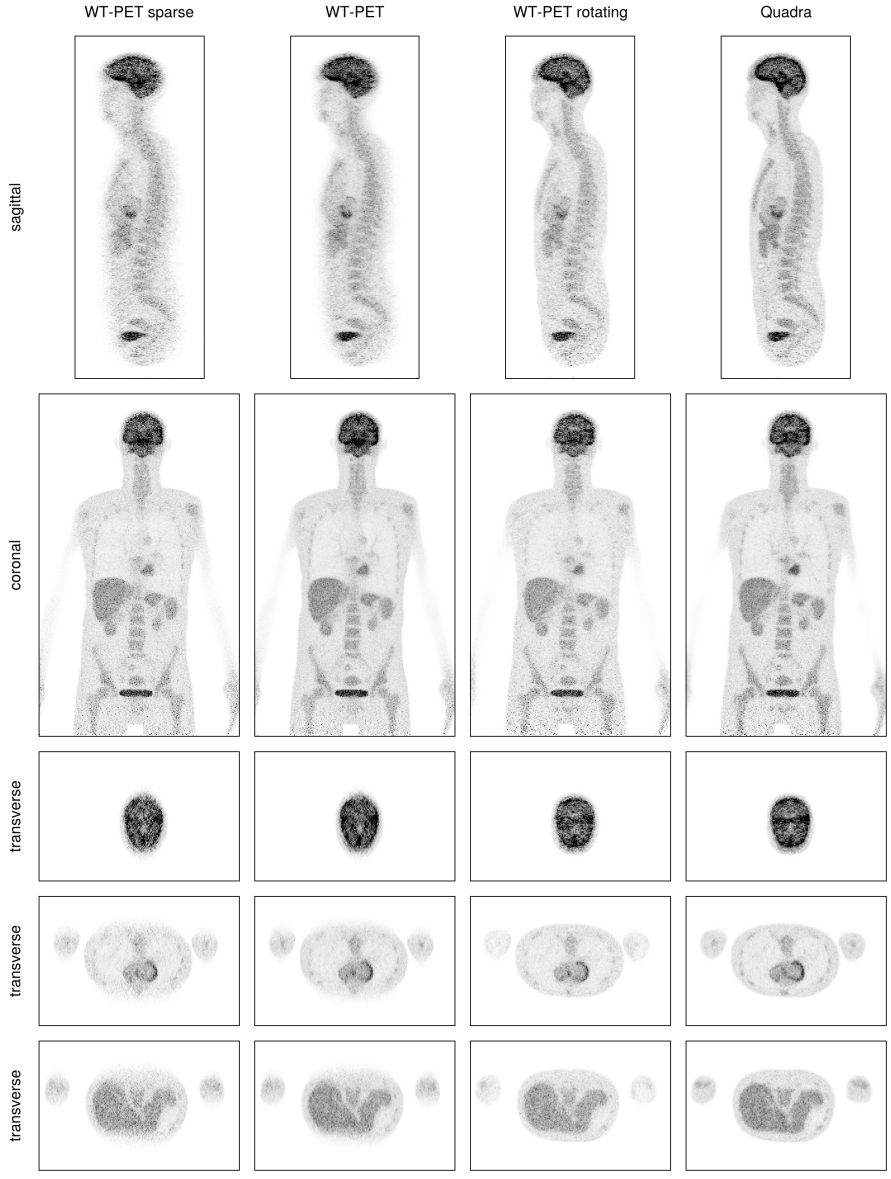


Figure 7.11: XCAT reconstructions for the three WT-PET configurations and the Quadra. The tenth iteration is visualized.

LORs to feasibly compute the system matrix for all of them.

Similarly, the uniformity phantom shows how modelling of the detector attenuation within the system matrix provides a much more uniform reconstruction prior to normalization. Even though this could in theory be entirely incorporated within the normalization factor, this puts more strain on the normalization in terms of required statistics and sinogram bin sizes required to obtain good reconstructions. We see this in Figure 7.8, where we actually get better uniformity modelling the detector attenuation without normalization, than including normalization without detector attenuation. Although not visualized, this procedure also gives very uniform reconstructions for the default WT-PET, the rotating WT-PET and the Quadra prior to normalization.

The IQ and XCAT reconstructions show that decent images can be obtained within just a 30 s acquisition, although the missing projection angles do result in some artifacts. As the reconstructions were done using true coincidences only, we can also expect higher noise levels in reality. Furthermore, the XCAT phantom, while complex, is not entirely representative of real patient data. For example, the background tissue and many organs are assigned a single uniform activity concentration, which can result in overly uniform reconstructions when compared to experimental data. In experiment, there is also the issue of patient motion causing a blurring in the reconstructions, which was not simulated here. Especially for the high spatial resolution of the WT-PET, this could have a noticeable impact.

7.5.2 Reconstruction speed

Finally, we would like to mention a few notes on the speed of reconstructions. All reconstructions were performed on a workstation with an AMD Ryzen 9 5950X CPU (16 physical / 32 virtual cores) and an NVIDIA GeForce RTX 4090 GPU (24 GB of VRAM), with 64 GB of RAM memory. Table 7.3 shows how long different steps of the reconstruction took for a 60s acquisition of the XCAT phantom in the default WT-PET configuration (using all coincidences, $\sim$ 420 million measured LORs). The sensitivity image was computed using 1 billion LORs, and the reconstruction FOV was 62 x 40 x 106 cm^3 (2 mm^3 voxels). In total, reconstruction took $\sim$ 10 minutes for 10 iterations, without using subsets.

component	duration
sensitivity image	2m30s
system matrix components for measured LORs	1m45s
H_{ang}	negligible
H_{det}	50s
H_{atn}	45s
single iteration	30s

Table 7.3: Computation times for different steps of the reconstruction of an XCAT phantom with 420 million measured LORs.

Chapter 8

Energy based scatter correction

8.1 Introduction

In the previous chapter, we have disregarded the effect of random and scattered coincidences by removing them from the data prior to reconstruction. This is possible when dealing with simulated data because we have access to the ground-truth, which is not the case in experimental data. Neither is it possible to somehow derive with absolute certainty whether a specific coincidence was true, scattered or random, and therefore, they cannot be removed from the data prior to reconstruction. The average contributions of random and scattered coincidences to the measured data can, however, be estimated and incorporated into the iterative image reconstruction procedure, as shown in equation 7.1.

As mentioned in the introductory chapters, random correction is fairly straightforward, and we have the option of either estimating random rates from the single count rates, or using the delayed coincidence time window method. On the WT-PET, we are planning to do coincidence sorting on the fly to limit data transfer rates, in which case we will not have access to the singles data. Therefore, we will base ourselves on the delayed coincidence time window method for random estimation.

As for scatter correction, in the introductory chapters we have introduced the SSS algorithm as a common method for scatter estimation. While powerful, it does have a few drawbacks, including the lack of multiple scatter modeling, the need for tail fitting, and the need for an estimate of the activity distribution without the effect of randoms and

scatters. The latter point implies we need to alternatingly update the activity distribution with MLEM and the scatter estimate with SSS, resulting in long computations. Instead of the SSS algorithm, we will implement and evaluate an energy based scatter correction method. In such a framework, the scatter contributions are estimated from the dual energy distributions of scattered and unscattered coincidences. This requires energy information to be stored for every coincidence, and is therefore only possible when using list-mode data. It does offer a few advantages compared to the SSS algorithm: it inherently takes into account the effects of multiple scatter, it estimates scatter rates exactly (without the need for tail fitting), and does not require an estimate of the activity or attenuation distributions. It can therefore be done without requiring a first reconstruction to estimate λ, and can be performed in the absence of a CT-scan. Primarily for the latter reason, we focus our efforts on energy based scatter correction, rather than the SSS algorithm, for the WT-PET. As the first prototype of the WT-PET will not have a CT in place (at least at the start), it would complicate scatter correction methods such as SSS. Even once the CT is installed, we still wish to have the option of performing PET scans without the CT for dose reduction purposes.

In this chapter, although scatter estimation will be done without the attenuation map, we will still use the ground truth attenuation map for attenuation correction during the reconstruction. Chapter 9 will cover some methods to eliminate the CT for attenuation correction as well. The combination of both chapters then provides a true CT-less PET reconstruction framework, where neither scatter correction nor attenuation correction requires an attenuation map derived from the CT. We however investigate both methodologies separately, in order to better understand their influence on image quality independent of one another.

For energy based scatter correction, we largely follow the methodology as proposed in [140], [141]. Our implementation is for the most part the same as in the original work, other than some additional sinogram bin normalizations to handle the continuous LOR representations of monolithic detectors. The main differences are in the acquisition data used for evaluation. Our work uses the WT-PET for evaluation, which ofcourse has a different geometry (dual-panel, rather than cylindrical), and makes use of BGO (rather than LYSO) detectors, resulting in an inferior energy resolution. We've also included XCAT simulation results, as a test on a more complex phantom covering the entire FOV.

8.2 Materials and Methods

8.2.1 Energy based scatter correction

We will first give a short overview of the energy based scatter correction procedure here. An overview is given in Figure 8.1.

In a collection of coincidence events (e.g., the events within a certain sinogram bin), each of the two gamma photons can either be scattered or unscattered, giving us four distinct groups: a fraction of events σ_{uu} where neither photon was scattered, a fraction σ_{us} where only the second photon scattered, a fraction σ_{su} where only the first photon scattered, and a fraction σ_{ss} where both photons scattered. Keeping the contribution of random coincidences separate, the dual (2D) energy distribution for this collection of coincidence events is then given by the form:

$$\begin{aligned} P_C(E_1, E_2) = \sigma_{uu} U(E_1) U(E_2) + \sigma_{us} U(E_1) S(E_2) \\ + \sigma_{su} S(E_1) U(E_2) + \sigma_{ss} S(E_1) S(E_2) \\ + P_R(E_1, E_2) \end{aligned} \tag{8.1}$$

where $U(E)$ and $S(E)$ are the energy probability densities for unscattered and scattered photons, respectively. These are only defined over the energy window ϵ and are therefore normalized so that $\int_\epsilon U(E)dE = \int_\epsilon S(E)dE = 1$ over ϵ. Random coincidences are considered separately in $P_R(E_1, E_2)$, so that $U(E)$ and $S(E)$ only refer to singles from a non-random coincidence. If $U(E)$ and $S(E)$ are known, it is then possible to use a statistical estimator to derive the fractions σ_{uu}, σ_{us}, σ_{su} and σ_{ss}, where the fraction of true coincidences is given by σ_{uu}, and the fraction of scattered coincidences is given by $\sigma_{us} + \sigma_{su} + \sigma_{ss}$.

The probability densities $U(E)$ and $S(E)$ can be derived from the single (1D) energy histogram:

$$P_S(E) = \sigma_u U(E) + \sigma_s S(E) + \sigma_r R(E) \tag{8.2}$$

where $R(E)$ is the energy probability density for a single (be it scattered or unscattered) belonging to a random coincidence event. The fractions of unscattered, scattered and random singles are given by σ_u, σ_s and σ_r, respectively. The energy probability density $U(E)$ is known: it is the detector energy response to 511 keV gamma photons, and can therefore be modelled as a Gaussian centered around 511 keV, and with FWHM equal to the energy resolution of the detector. The probability density $R(E)$, and the accompanying fraction σ_r, can be obtained directly from

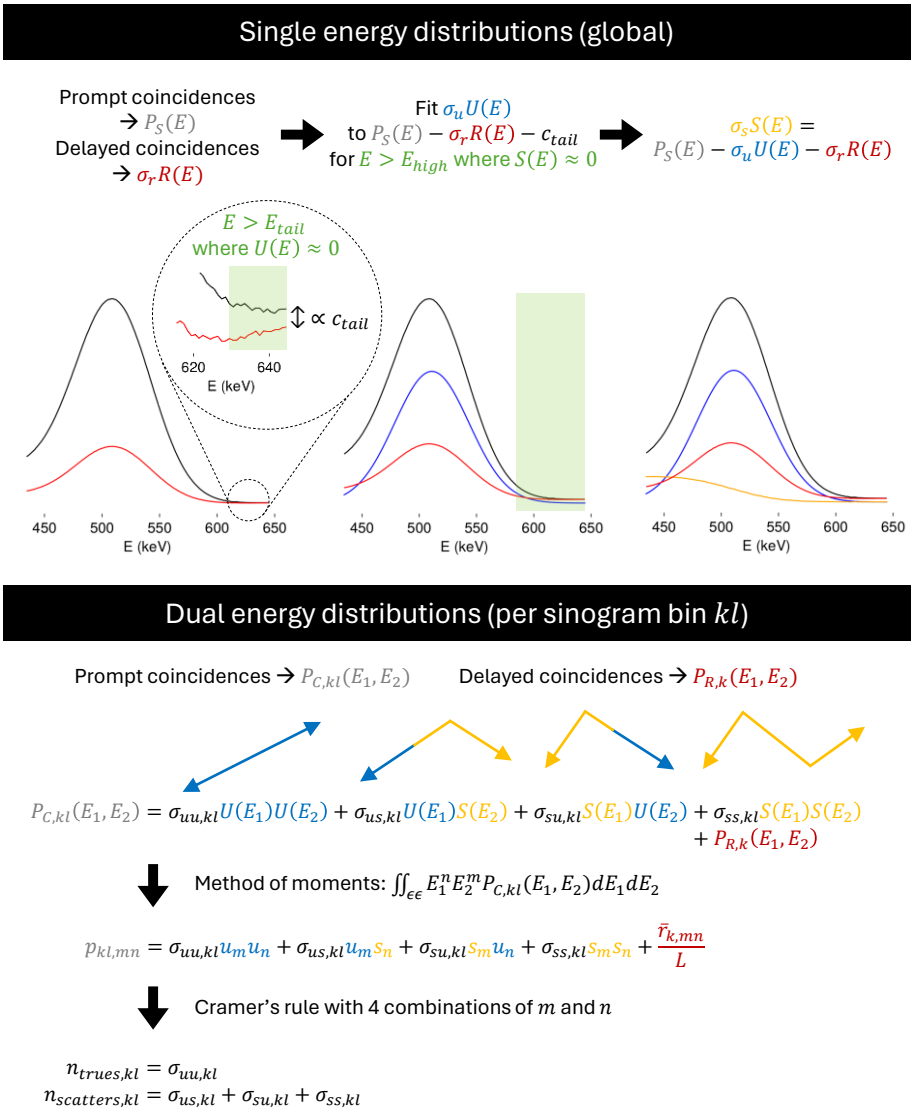


Figure 8.1: Schematic overview of the energy based scatter correction procedure.

the delayed coincidences. Only $S(E)$ is unknown, but can be derived from equation 8.2 if either σ_u or σ_s is known.

We assume that the high energy range (at some point past the photopeak) is dominated by unscattered photons, and that therefore $S(E > E_{\text{high}}) = 0$. This allows us to fit a Gaussian to this high-energy range, after subtraction of $\sigma_r R(E)$, giving us the value for σ_u. The shape of $U(E)$ is however not strictly Gaussian, even in GATE simulation, due to effects related to inter-crystal scatter, partial energy deposition, or energy sharing. This results in a longer tail than would be observed in a Gaussian distribution. The effect is sufficiently small that we can still model $U(E)$ as a Gaussian, however the tail contribution should be taken into account in order to obtain the correct value for σ_u, and therefore the correct shape of $S(E)$. We correct for this effect by first subtracting these events in the tail (for which $E > E_{\text{tail}}$, with $E_{\text{tail}} > E_{\text{high}}$), prior to performing the fit of $U(E)$:

$$\int_{E>E_{\text{high}}} \sigma_u U(E)dE \approx \int_{E>E_{\text{high}}} P_S(E)dE - \int_{E>E_{\text{high}}} \sigma_r R(E)dE - c_{\text{tail}} \quad (8.3)$$

with:

$$c_{\text{tail}} = (n_{\text{prompt,tail}} - n_{\text{random,tail}}) \frac{E_{\text{max}} - E_{\text{high}}}{E_{\text{max}} - E_{\text{tail}}} \quad (8.4)$$

where $n_{\text{prompt,tail}}$ and $n_{\text{random,tail}}$ are the total number of singles in the tail energy region, belonging to a prompt or random coincidence, respectively. E_{max} is the upper threshold of the energy window.

We note that the probability density $S(E)$ is object dependent, and furthermore, its low energy region also varies with the position in projection space, i.e., the sinogram bin [140]. The higher energy region, but still starting below 511 keV, is however largely position independent, resulting in more or less a unique shape for $S(E)$ per object/acquisition. Therefore, as long as we choose the lower threshold of the energy window to be sufficiently large, it suffices to use the same distribution $S(E)$ for any sinogram bin in equation 8.1.

With this knowledge, we can now derive the fractions σ_{uu}, σ_{us}, σ_{su} and σ_{ss} for a group of coincidences belonging to a sinogram bin k and TOF bin l. We use the method of moments, where we match the moments

of the acquisition data with those of the model in equation 8.1:

$$\begin{aligned} p_{kl,mn} - \frac{\bar{r}_{k,mn}}{L} &= \int_{\epsilon}\int_{\epsilon} E_1^n E_2^m P_{C,kl}(E_1, E_2) dE_1 dE_2 \\ &= \sigma_{uu,kl} u_m u_n + \sigma_{us,kl} u_m s_n \\ &\quad + \sigma_{su,kl} s_m u_n + \sigma_{ss,kl} s_m s_n \end{aligned} \tag{8.5}$$

where $u_m = \int_{\epsilon} E^m U(E) dE$ and $s_m = \int_{\epsilon} E^m S(E) dE$. The moment of order (m, n) of the prompt coincidences in sinogram bin k and TOF bin l is given by $p_{kl,mn} = \sum_{i \in kl} E_{i1}^m E_{i2}^n$. Similarly, the moment of the random coincidences in sinogram bin k is estimated from the delayed coincidences as $\bar{r}_{k,mn} = \sum_{i \in k} E_{i1}^m E_{i2}^n$, which is divided by the total number of TOF bins L. For the randoms, we have aggregated over the TOF bins, as the rate of delayed coincidences is independent of the TOF bin. Choosing two distinct values for m and n, gives us a set of four equations with four unkowns ($\sigma_{uu,kl}$, $\sigma_{us,kl}$, $\sigma_{su,kl}$ and $\sigma_{ss,kl}$), which can be solved using Cramer's rule. We note that the predicted number of random coincidences within a sinogram bin is simply $\bar{r}_k := \bar{r}_{k,00}$.

8.2.2 Implementation details

We have kept the same energy window as before (434 - 645 keV), and choose the high energy threshold above which $S(E)$ is assumed to be zero, to be two standard deviations of the energy resolution above 511 keV ($E_{\text{high}} = 576$ keV). The tail threshold E_{tail} is set to 630 keV. For the method of moments, we use orders 0 and 1 for m and n. As was necessary for the computation of H_{norm}, we need to discretize our continuous list-mode data, for which we again use projection space coordinates r, ϕ, z, θ and t. The obtained scatter sinogram ($\sigma_{us}+\sigma_{su}+\sigma_{ss}$) is further smoothed with a Gaussian blurring kernel, given the slowly varying contribution of scatters in projection space. Computation of these fractions σ gives us the total number of scatters (and randoms from the delayed coincidences) within a certain sinogram bin. In order to obtain the random and scatter contribution per LOR in the emission data, it is however necessary to normalize this based on the LOR sampling scheme proposed back in Figure 7.1. As it stands, the "size" of a sinogram bin does not match the size of a LOR in projection space. Therefore, the random and scatter contributions $r_i + s_i = e_i$, as used in the list-mode MLEM update

equation 7.1, are computed as:

$$\begin{aligned} r_{i,kl} &= \frac{1}{n_{I,k}} \frac{T_{\text{cont}}}{T_{\text{discr}}} \frac{\bar{r}_k}{L} \\ s_{i,kl} &= \frac{1}{n_{I,k}} \frac{T_{\text{cont}}}{T_{\text{discr}}} \left(\sigma_{us,kl} + \sigma_{su,kl} + \sigma_{ss,kl}\right) \end{aligned} \tag{8.6}$$

with $n_{I,k}$ the number of LORs in sinogram bin k in the sampled LOR space, T_{discr} the length of the discretized TOF bin, and T_{cont} the length of the continuous TOF bin. The first fraction corrects for the fact that there are different numbers of sampled LORs for each sinogram bin, as the sampling scheme we used does not match the discretization scheme. In practice, we use relatively large sinogram bins due to the smoothly varying contribution of scatters, so that $n_{I,kl} >> 1$. The second fraction corrects for the difference between the length of the discretized TOF bin and the continuous TOF bin. T_{cont} depends on the normalization chosen for the computation of the TOF weighting factor $H_{TOF,ij}$, which in our case was chosen such that $cT_{\text{cont}}/2 = 1$ mm, with c the speed of light. Rather than simply assigning the contributions $\bar{r}_{k,mn}/L$ and $\sigma_{us,kl} + \sigma_{su,kl} + \sigma_{ss,kl}$ from sinogram bin kl to LOR i, we instead obtain a unique contribution for each LOR i by cubic B-spline interpolation of the corresponding sinograms.

Simulation and reconstruction

We evaluate the performance of energy based scatter correction on GATE simulations of the IQ and XCAT phantoms. The IQ phantom again uses a 4:1 activity concentration ratio of spheres to background. We also use the same XCAT phantom as before: a female of age 52, with a height of 179 cm and a weight of 72 kg (BMI of 22.47), using a scan dose of 3 MBq/kg (see Figure 7.6 for the ground-truth activity map). All simulation parameters remain the same as in Chapter 7, still using 30s acquisitions, and in addition to recording prompt coincidences, we also record delayed coincidences using a 50 ns delay. Reconstruction parameters also remain the same, but it is now done using all prompt coincidences, by including the estimated random and scatter contributions in the MLEM update equation.

8.2.3 Evaluation metrics

For the IQ phantom, in addition to the CRC, we now also include metrics that evaluate noise in the image: the background variability (BV) and

the contrast-to-noise ratio (CNR). We use the same (2D transverse) ROIs as proposed by NEMA, and as used for the computation of the CRC in equation 7.10.

The background variability provides a measure for the noise in the background, and is computed for each sphere k as [24]:

$$\mathrm{BV}_k = \frac{\sqrt{\sum_{l=1}^{L} \left(\mu_{B,k,l} - \mu_{B,k}\right)^2 / (L-1)}}{\mu_{B,k}} \tag{8.7}$$

with $L = 60$ the total number of background ROIs for sphere k, $\mu_{B,k}$ the average background counts over all ROIs and $\mu_{B,k,l}$ the average background counts for ROI l.

We also compute the contrast-to-noise ratio (CNR), as a metric that better measures how "detectable" a sphere is by taking into account both the contrast recovery and the noise statistics:

$$\mathrm{CNR}_k = \frac{\mu_{H,k} - \mu_{B,k}}{\sigma_{B,k}} \tag{8.8}$$

Here, $\mu_{H,k}$ is the average counts in the ROI of sphere k, and $\mu_{B,k}$ and $\sigma_{B,k}$ are the average and standard deviation of all counts in the corresponding background ROIs.

8.3 Results

8.3.1 IQ phantom

The predicted single energy distributions for the IQ phantom, scaled by the fractions σ_u, σ_s and σ_r, are shown in Figure 8.2 and compared to the ground-truth obtained from Monte Carlo simulation. We observe that the energy probability densities are predicted with minimal error.

In Figure 8.3, we show the predicted random ($\bar{r}_{k,mn}/L$) and scatter ($\sigma_{us} + \sigma_{su} + \sigma_{ss}$) contributions along a random selection of sinogram bins for the IQ phantom, in function of the radial distance r. The shown scatter contributions are already post-smoothed with a Gaussian blurring kernel. Here, we do observe some small mismatches between the predicted and true scatter contributions along certain sinogram bins. The maximum deviation in any sinogram bin (difference between yellow lines in Figure 8.3), including those not visualized, was 126 counts. For reference, the maximum number of prompt coincidences in any sinogram bin (black line) was 1903.

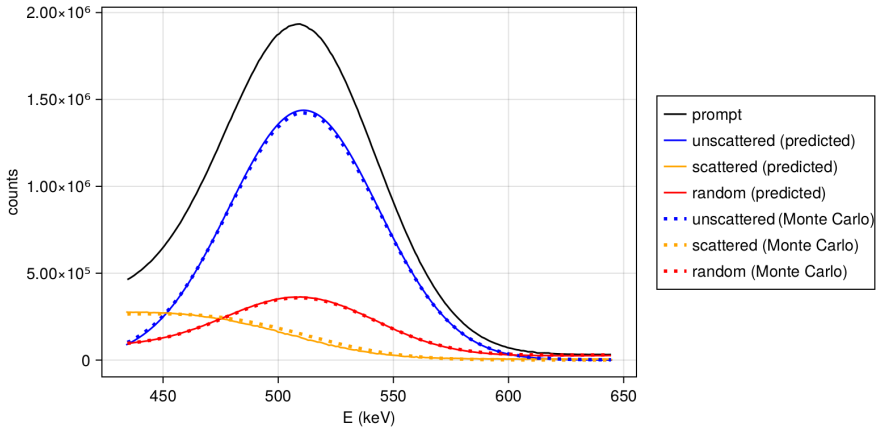


Figure 8.2: The scaled single energy distributions $\sigma_u U(E)$, $\sigma_s S(E)$ and $\sigma_r R(E)$ for the IQ phantom, compared with the ground-truth histograms obtained from Monte Carlo.

Figure 8.4 shows the reconstruction of the IQ phantom at the tenth iteration, using the energy based scatter correction method, compared to a reconstruction using only true coincidences, and a reconstruction without random/scatter correction. All images use the same color scale. We again show the CRC for the different sphere sizes in function of iteration, but now also include the BV and CNR. Note that the CNR increases at first (due to increasing contrast as observed in the CRC), but as the noise keeps growing in further iterations, the CNR reaches a maximum, after which it drops off again. Visually, the scatter corrected image looks very similar to the trues only reconstruction, although the smaller spheres do show some contrast loss. This is also confirmed by the metrics. The uncorrected image overestimates the activity, especially near the center of the phantom. Figure 8.5 shows a horizontal line profile through the reconstructions shown on top of Figure 8.4, through the center of the two middle spheres. Here we can see that while the corrected and trues only reconstructions match closely, there is still some activity in the central lung insert that is not entirely corrected for.

8.3.2 XCAT phantom

Figure 8.6 shows the predicted single energy distributions for the XCAT phantom, scaled by the fractions σ_u, σ_s and σ_r. They are again compared

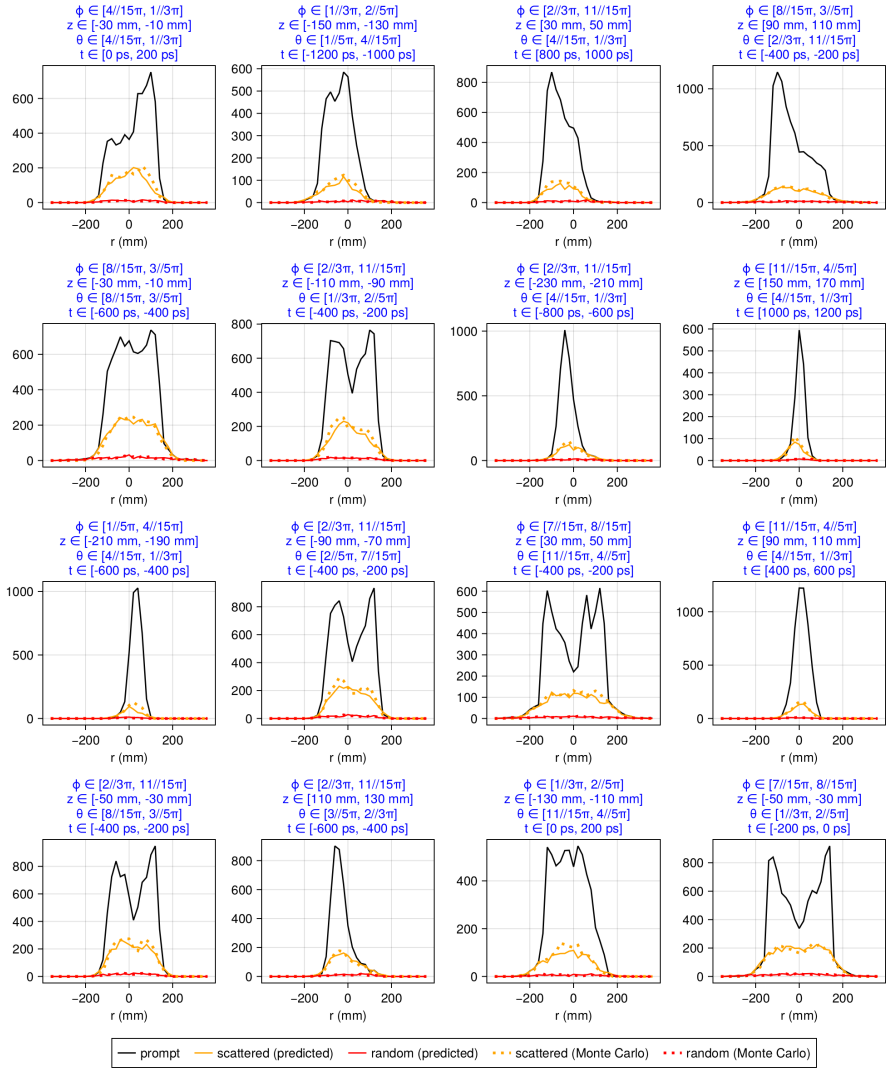


Figure 8.3: Predicted number of random and scattered coincidences along specific sinogram bins, compared to the ground-truth obtained from the GATE simulation.

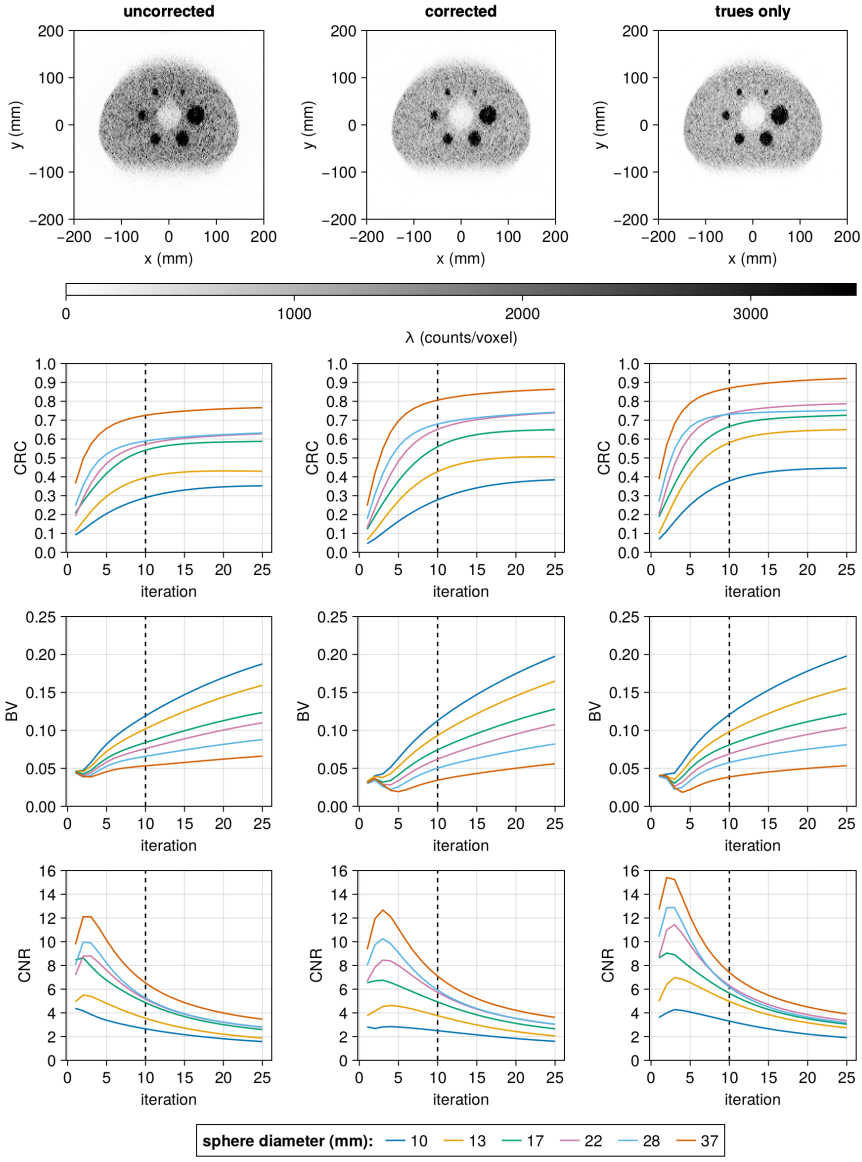


Figure 8.4: Comparison of IQ reconstructions using prompt coincidences with no random/scatter correction (left), using prompt coincidences with energy based random/scatter correction (center), and using only true coincidences (right). The tenth iteration is visualized on top, and the contrast recovery coefficient (CRC), background variability (BV) and contrast-to-noise ratio (CNR) are computed for the different sphere sizes.

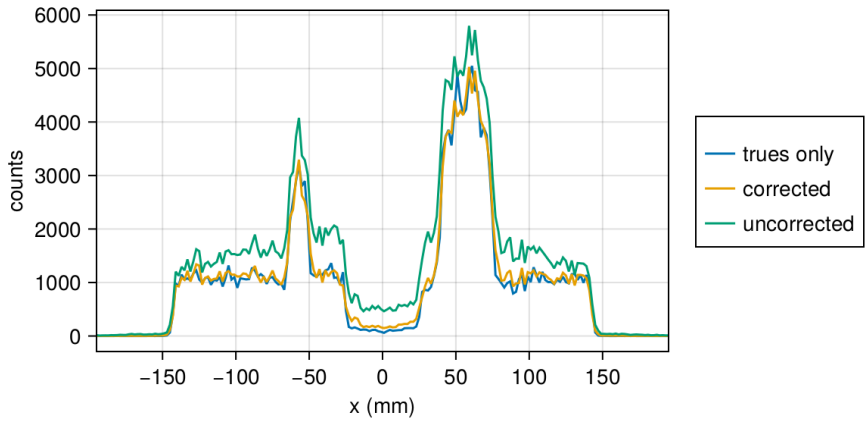


Figure 8.5: Line profile along the x-axis through the two central spheres in the reconstructed IQ phantoms shown on top in Figure 8.4.

to the ground-truth obtained from Monte Carlo simulation. The predicted energy density functions again match well match with the ground-truth data.

Finally, in Figure 8.7, the reconstruction of the XCAT at the 10th iteration is shown, with and without random/scatter correction, and compared to the trues only reconstruction. In addition, we show the error map (corrected image - trues only image) to investigate any systematic error introduced by the scatter correction. The error map has been post-processed with a 3D Gaussian filter, with $\sigma = 4$ mm, in order to better highlight regions of over and/or underestimated activity. Some degree of detail (contrast in sharp features) is lost when comparing the corrected image to the trues only reconstruction. We also observe that some leftover activity remains in the background. Activity in the lung and spine is somewhat underestimated, whereas liver, brain and hip activity are slightly overestimated. Noting the color scales however, these effects are relatively minor.

8.4 Discussion

The scatter-corrected reconstructions show good overall agreement with the trues only reconstructions, although there is some loss of detail as expected. As mentioned, energy based scatter correction offers some

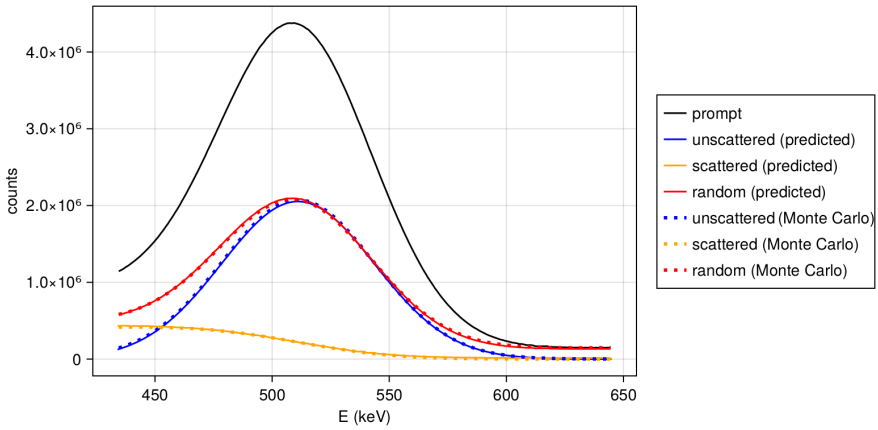


Figure 8.6: The scaled single energy distributions $\sigma_u U(E)$, $\sigma_s S(E)$ and $\sigma_r R(E)$ for the XCAT phantom, compared with the ground-truth histograms obtained from Monte Carlo.

key advantages over SSS: an activity/attenuation map is not required, multiple scatter is inherently modeled, and there is no need for tail fitting, making it an attractive alternative to SSS when using list-mode data.

Since no knowledge of the activity distribution is required, there is also no need to alternatingly update the activity distribution and scatter estimations as is done in SSS. This considerably speeds up the process, as only a single scatter estimation and reconstruction are required for energy based scatter correction. Scatter correction is also possible without a CT, since we do not require an attenuation map. In theory, we could still use SSS without a CT by using an attenuation map derived from a transmission source or from the emission data itself. However, as will be seen in Chapter 9, these attenuation maps are much noisier and less detailed than the CT. Furthermore, these methods only provide an attenuation map in the PET reconstruction FOV, and may therefore not cover the whole patient. For accurate modelling of scatters, however, the SSS algorithm requires the entire attenuation distribution, not only the portion within the FOV. This would therefore pose a limitation for the SSS algorithm in the absence of a CT. Finally, energy based scatter correction inherently models multiple scatter and requires no tail fitting. While the tail fitting procedure can be a source of noise in SSS, we do note that it can simultaneously be beneficial, as it serves to partially correct for the lack of multiple scatter modelling, which would otherwise

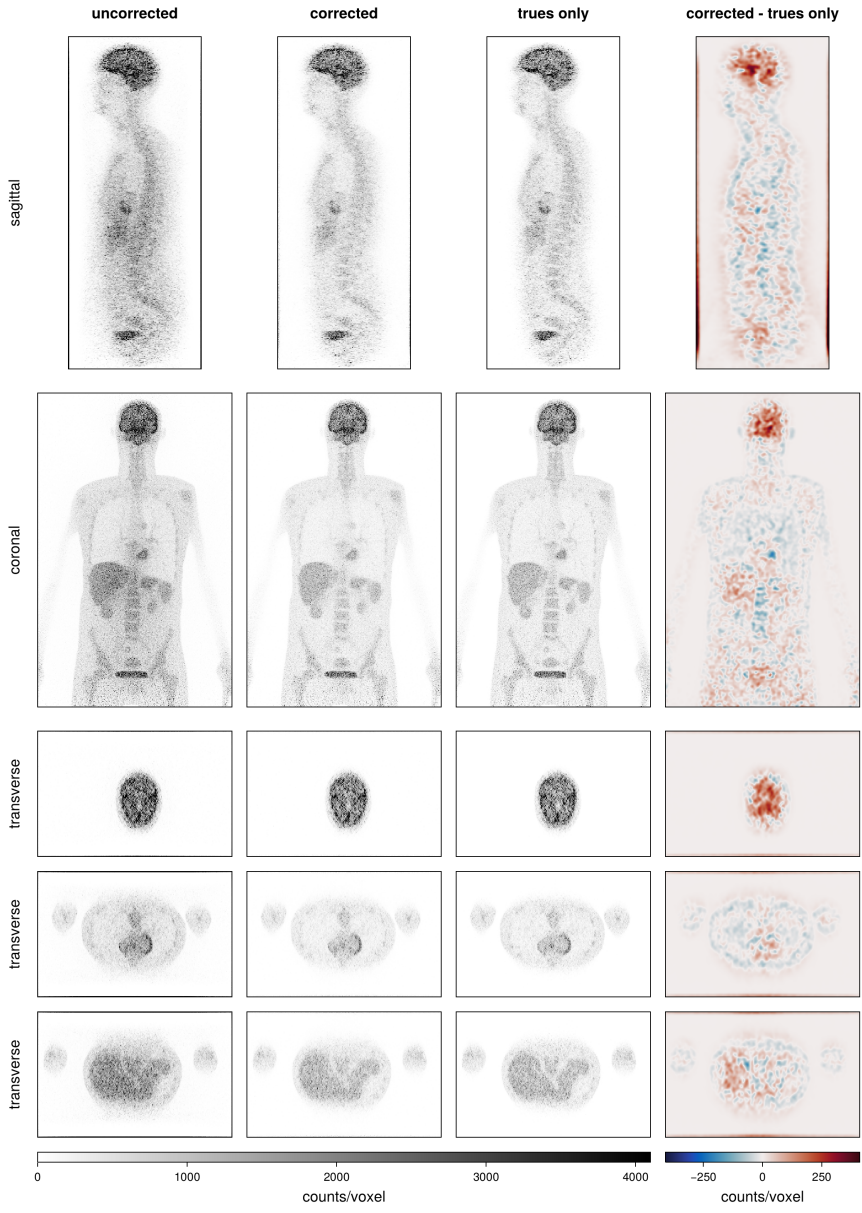


Figure 8.7: XCAT reconstruction (10th iteration) comparison for prompt coincidences without random/scatter correction (left), for prompt coincidences with energy based random/scatter correction (center left), and for true coincidences only (center right). The error map between corrected and trues only reconstructions is shown on the right.

result in an underestimation of the total number of scatters.

The accuracy of energy based scatter correction is dependent on how well the single energy probability density $S(E)$ can be estimated, which in turn depends on the choice of the high energy threshold E_{high}. In our specific case, we have found a value of 2σ above 511 keV to work well. While the energy response of a physical detector is normally well approximated by a Gaussian distribution, the goodness of the fit may be overestimated in simulation, since the energy blurring in GATE is done by sampling from a Gaussian distribution in the first place. Effects such as crystal scatter can of course alter this, resulting in a non-Gaussian distribution for simulated data as well. The non-Gaussian nature is already partially taken into account by subtracting the leftover events in the tail (c_{tail}), but if the actual detector response is too different, it may be better to use a non-Gaussian shape for $U(E)$. This shape could for example be obtained from experiment, using a scatter-free source.

Related to this, one potential pitfall with monolithic detectors is that the energy distribution of true coincidences can be dependent on the interaction position within the detector, due to differences in light spread between the center and edges of the detector, or differences between various DOIs. In this case, the energy distribution $U(E)$ would no longer be the same for all sinogram bins. In our current simulation, this was not modelled. However, once we have experimental detector results, it could be interesting to investigate how large these variations in energy distribution are, and if we can keep using a single energy distribution $U(E)$, or should instead use a number of position (i.e., sinogram bin) dependent distributions.

Another potential pitfall for energy based scatter correction, is that the accuracy of the estimation of the fractions σ in equation 8.5 depends on the number of measured prompt and delayed coincidences within that specific sinogram bin. A low number of measurements results in a reduced confidence that they represent an accurate approximation of the true dual energy distribution. Given that we are using 30 s acquisitions, this may be a considerable source of noise, resulting in the mismatches observed in Figure 8.3. Furthermore, this is not offset by the increased sensitivity of the WT-PET compared to short AFOV scanners. Although we obtain more coincidence events within 30 s, the number of possible sinogram bins is also considerably larger. This would not be an issue in SSS, where the LORs and scatter points are sampled, and we can therefore increase the statistics per sinogram bin as we see fit. Short

scan durations do, however, reduce the accuracy of the estimated activity distribution, which in turn also impacts the SSS algorithm.

In general, we believe the advantages of energy based scatter estimation outweigh the disadvantages, at least in those cases where a CT is unavailable, or when it is preferable to eliminate the CT for dose reduction purposes.

Chapter 9

CT-less attenuation correction

9.1 Introduction

Nowadays, attenuation correction in PET is almost always done using the accompanying CT-scan, which is already there to provide additional anatomical information. Anatomical information from the CT is however not always required, as is often the case in brain imaging, or follow-up studies where a prior CT is already available. In these cases, a low dose CT can instead be obtained, used only for the purpose of attenuation and/or scatter correction. For dose reduction purpose, it would however be desirable to completely eliminate the CT when possible. In Chapter 8, we have already seen how scatter correction can be done in the absence of a CT scan, using energy based scatter correction. Similarly, there exist methods to do attenuation correction without the CT, thereby enabling CT-less PET image reconstruction.

CT-less attenuation correction can for example be done using deep learning algorithms to estimate the attenuation map from uncorrected PET images [142], [143], or from other structural information such as an MRI, when available [144]. Attenuation (and scatter) correction can then be done as usual using the "synthetic" CT image. Another approach is to directly estimate the corrected PET images from the uncorrected images [145]. In this case, attenuation and scatter correction are both done simultaneously in image space, post-reconstruction of the uncorrected image.

As for non-deep learning approaches, when available, MRI can still be

used to estimate the attenuation map using for example template-based (co-registering an average patient attenuation map to the MRI image) or segmentation-based (e.g., clustering of intensity values) approaches [146]. If only PET is available, one option is to use an external 511 keV source to obtain a transmission scan, as was done in the early days of PET. With the introduction of TOF-PET, it is now possible to perform this transmission scan simultaneously with the actual PET acquisition (emission scan) itself. Transmission data, originating from the transmission source placed close to the detectors (and therefore far from the FOV center) will have large TOF differences. Emission data on the other hand, originating from the patient (and therefore closer to the FOV center) will have smaller TOF differences. It is therefore possible to distinguish simultaneously acquired transmission and emission data from one another based on their TOF difference [147]. Of course, this does again result in additional radiation dose for the patient, although smaller than with CT. It may also offer some other advantages compared to CT, outlined below. Finally, the attenuation can be estimated from the TOF emission data itself, essentially considering the emission data as a transmission source, using algorithms that simultaneously update the activity and attenuation maps in an interleaving fashion. In this chapter, we will focus on these two, non-deep learning options.

Besides dose reduction, these CT-less attenuation correction methods offer the advantage that the attenuation map is derived from data obtained during the PET scan itself. This accelerates the acquisition procedure, and ensures that no mismatch will occur due to possible patient movement between the PET and CT acquisitions. Furthermore, the attenuation is derived directly using 511 keV gamma photons, so that no errors are introduced when scaling from the CT energy to the PET energy. The CT can often only provide an approximation of the attenuation at 511 keV, due to necessary simplifications, such as for example, incorrectly assuming CT energy to be monoenergetic.

Such CT-less attenuation correction methods could be of great use for the WT-PET specifically, where the goal is high throughput and the system may see frequent use for follow-up studies that do not require a CT.

9.2 Materials and Methods

9.2.1 Maximum likelihood transmission

When using an external transmission source, the idea is to derive the attenuation map by comparing the projection space data acquired with the attenuating object present, to an acquisition without any attenuation. A common choice for a transmission source would be a ^{68}Ge rod, which is a positron emitter and therefore produces two back-to-back 511 keV gamma photons upon electron-positron annihilation. If the rod is placed somewhere inside the scanner, but close to the detectors, one of the photons can make it to a detector undisturbed, whereas the other may pass through the patient and can therefore be attenuated. Performing a scan once with, and once without, the patient, gives rise to two sets of projection data: a transmission scan T with attenuation, and a blank scan B without attenuation.

The attenuation map can then be derived iteratively using maximum likelihood methods, where we again use the likelihood for Poisson distributed values, due to the single photon counting nature of our measurements [148]. Adapting the maximum likelihood transmission (MLTR) update equation from [149] to list-mode data, we obtain:

$$\begin{aligned} \bar{y}_i^{(k)} &= e^{-\sum_{j\in J} l_{ij}\mu_j^{(k)}} \\ \mu_j^{(k+1)} &= \mu_j^{(k)} + \alpha \frac{\sum_{i\in B} l_{ij}\bar{y}_i^{(k)} - \sum_{i\in T} l_{ij}}{\sum_{i\in B} l_{ij}\bar{y}_i^{(k)} \sum_{j'\in J} l_{ij'}} \end{aligned} \tag{9.1}$$

This equation is only valid in the absence of random and/or scattered coincidences. Here, μ_j^k is the attenuation coefficient in voxel j at iteration k. LORs $i \in T$ belong to the transmission scan and LORs $i \in B$ belong to the blank scan. The elements l_{ij} are again the intersection lengths of LOR i with voxel j, and $\alpha > 0$ is a relaxation parameter that controls the speed of convergence, which we have set to 1 in all experiments. The elements $\bar{y}_i^{(k)}$ are the predicted counts along LOR i, given the current estimate of the attenuation map. In our specific case, due to the continuous LOR estimation, none of the LORs in the transmission scan T will exactly match those in the blank scan B. This however poses no issue, as we only compare the sum of LORs, rather than individual LORs, of the transmission scan and the blank scan. We also note that since the summation over T does not depend on the attenuation coefficients,

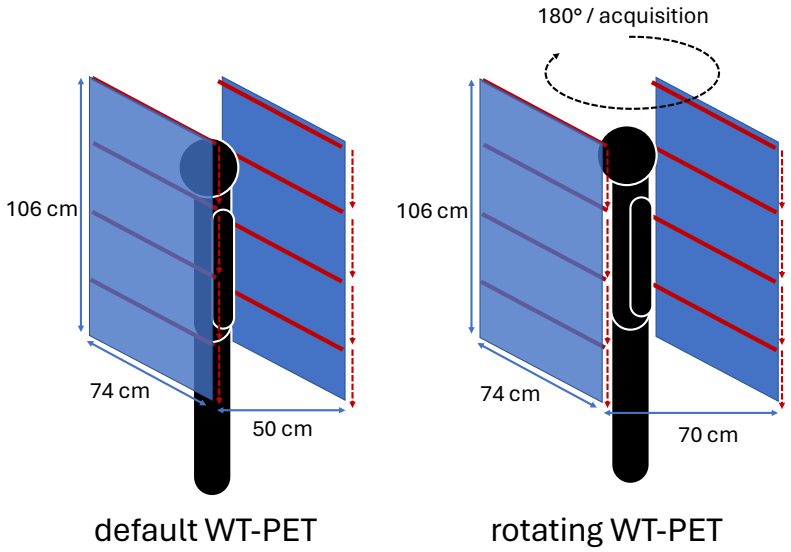


Figure 9.1: Placement of the transmission sources in the default and rotating configurations of the WT-PET.

it can be precomputed at the start, and only the LORs from the blank scan are required in further iterations to update the attenuation map.

In our simulations for the WT-PET, we opt for 4 horizontal, isotropically emitting rod sources (1 mm diameter) per panel (8 total), each with a total activity of 3 MBq, see Figure 9.1. The activity was chosen low enough so that no appreciable detector saturation would occur due to the transmission sources. The sources stretch the full width of the panel and are placed 1 cm in front of it. They travel downwards during a 30 s scan, so that each line source sweeps over a quarter of the panel during the scan duration. In the case of the rotating WT-PET configuration, the sources rotate together with the panels while translating downwards. In GATE, the rod source is modelled as a back-to-back gamma photon emitter, rather than a positron emitter. Other than the lack of positron range modelling, this should have no effect on the results.

As mentioned, the transmission scan can be done simultaneously with the PET acquisition, by separating the emission and transmission data based on TOF information. Our reconstructions were, however, obtained using a separate simulation for the transmission and emission

scan. Nonetheless, we will investigate the feasibility of TOF separation, in the presence of randoms and scatters, on a simultaneous acquisition of the IQ phantom, but without reconstruction. To this end, we assign coincidences to the transmission scan based on two conditions, which should simultaneously be true. First, the absolute TOF difference should be within a certain range where we do not expect to find (many) true emission coincidences, for which we use the range from 1000 to 4000 ps. Second, the LOR should pass through one of the transmission sources, which we consider to be the case if the closest point of the LOR to any of the transmission sources is less than 3 mm. This second condition allows us to not only further reject emission data within the aforementioned TOF difference range, but it also allows us to reject the majority of random and scattered coincidences, which would otherwise contaminate the transmission data.

9.2.2 Maximum likelihood activity and attenuation

Rather than relying on an external transmission source, it is also possible to use the emission data itself as the transmission source, enabling simultaneous estimation of the activity and attenuation [150]. The difficulty with such approaches in the past has however been that they lack a unique solution and severe "cross-talk" can be observed between the activity and attenuation reconstructions. It has however been shown that the inclusion of TOF information offers a unique solution, up to a constant scaling factor, and thereby also removes the issue of cross-talk [151].

The maximum likelihood activity and attenuation (MLAA) algorithm uses an interleaved updating of the activity and attenuation, where in every iteration, first the activity is updated while keeping the attenuation fixed, and then the attenuation is updated while keeping the activity fixed [152]. The update equations, adapted to list-mode data, can be

expressed as:

$$\begin{aligned} a_i^{(k)} &= e^{-\sum_{j\in J} l_{ij}\mu_j^{(k)}} \\ \lambda_j^{(k+1)} &= \frac{\lambda_j^{(k)}}{\sum_{i\in I} a_i^{(k)}\tilde{H}_{ij}^{(k)}} \sum_{i\in E} \frac{a_i^{(k)} H_{TOF,ij}\tilde{H}_{ij}}{\sum_{j'\in J} a_i^{(k)} H_{TOF,ij'}\tilde{H}_{ij'}\lambda_{j'}^{(k)}} \\ \bar{y}_i^{(k+1)} &= a_i^{(k)} \sum_{j\in J} \tilde{H}_{ij}\lambda_j^{(k+1)} \\ \mu_j^{(k+1)} &= \mu_j^{(k)} + \frac{\sum_{i\in I} l_{ij}\bar{y}_i^{(k+1)} - \sum_{i\in E} l_{ij}}{\sum_{i\in I} l_{ij}\bar{y}_i^{(k+1)} \sum_{j'\in J} l_{ij'}} \end{aligned} \tag{9.2}$$

which are again only valid in the absence of random and/or scattered coincidences. Here, $\tilde{H}_{ij}$ is the system matrix without inclusion of the attenuation correction factor or the TOF weighting factor. We recognize the MLEM update equation 7.1 for λ_j^{k+1}, where the current estimate of the attenuation map $\mu_j^{(k)}$ is used to calculate the attenuation correction factors $a_i^{(k)}$. The MLTR update equation 9.1 (with $\alpha = 1$) is also visible, where the emission data E is used as the transmission scan, and the sampled space of potential LORs I is used as the blank scan.

Since these update equations only provide the correct solution for λ and μ up to a constant scaling factor, some form of constraint based on prior knowledge should be put on λ and/or μ to obtain the correct values. In our case, we normalize the activity at every iteration, imposing that the total activity should equal the one obtained from the MLEM reconstruction. Furthermore, we set the attenuation to zero for voxels outside the phantom support, which is derived from the activity estimate, as we otherwise obtain non-zero attenuation outside the phantom.

9.2.3 Simulation and reconstruction parameters

We implement and evaluate the MLTR and MLAA algorithms within our reconstruction framework, again using GATE Monte Carlo simulations of 30 s acquisitions for the IQ and XCAT phantoms. Both the default and the rotating configuration of the WT-PET are utilized, with the rotating configuration primarily included to investigate the algorithms' performance in the absence of limited angle artifacts. The IQ phantom again used a 4:1 sphere to background activity concentration ratio with a background activity concentration of 5.3 kBq/ml, but we have used a different XCAT phantom than before. The XCAT phantom is now a

male of age 31, with a weight of 77.9 kg and a height of 185.2 cm (BMI = 22.71), see Figure 7.6 for the ground-truth activity map. The scan dose is again 3 MBq/kg. Reconstructions are done using true coincidences only, where random and scattered coincidences have been removed from the data based on the Monte Carlo labels. Both the attenuation map μ and activity map λ are reconstructed using 2 x 2 x 2 mm^3 voxels. For the MLTR reconstruction, we use a total of 25 iterations (no subsets), given the slower convergence due to the lack of TOF information. For the same reason in MLAA, we perform five updates of the attenuation map, for every update of the activity map. We do not use subsets, and the activity reconstructions are all visualized at the tenth iteration.

9.3 Results

The various reconstructions for the IQ phantom in the default WT-PET configuration are shown in Figure 9.2. Here, we compare the uncorrected (NAC) image, the MLEM reconstruction using the MLTR derived attenuation map, the MLAA reconstruction, and the MLEM reconstruction using the ground-truth attenuation map. The corresponding attenuation maps are also visualized on top, and the CRC for each sphere is calculated in function of the iteration. Severe limited angle artifacts are observed for the attenuation map predicted by MLTR, which is not the case in MLAA. Nonetheless, both MLTR-based and MLAA activity reconstructions closely match the MLEM reconstruction obtained with the ground-truth attenuation map, although some differences can be observed in terms of the CRC values. The same IQ phantom reconstructions, but now for the rotating WT-PET, are shown in Figure 9.3. As expected, the limited angle artifacts are no longer visible, resulting in a much better reconstruction of the attenuation map for MLTR.

Reconstructions for the XCAT phantom on the default WT-PET configuration are shown in Figures 9.4 (activity maps) and 9.5 (attenuation maps). We also show the difference image between the MLTR-based and MLEM reconstruction (λ_{MLTR} - λ_{MLEM}), and the MLAA and MLEM reconstruction (λ_{MLAA} - λ_{MLEM}). The difference images have been post-smoothed with a Gaussian blurring filter (σ = 4 mm) to better show the overall tendencies. We again observe severe limited angle artifacts for the attenuation map obtained with MLTR, although the activity map is predicted quite well. For MLAA, however, we do observe more systematic differences between the MLAA and MLEM reconstructions, both

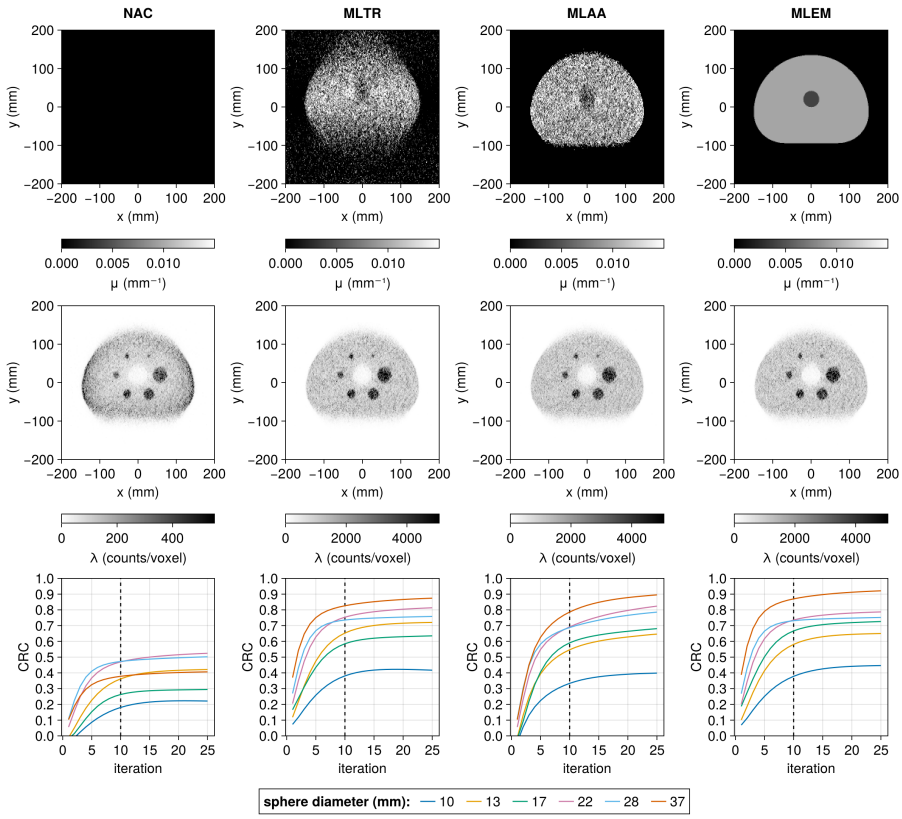


Figure 9.2: IQ phantom reconstructions on the **default WT-PET** configuration, comparing CT-less attenuation correction methods (MLTR and MLAA) to uncorrected (NAC) or ground-truth attenuation corrected images (MLEM).

in terms of activity and attenuation. While iterating further did reduce this "drift", it did not completely eliminate the issue, and produced too noisy images without regularization. Finally, Figures 9.6 and 9.6 show the activity and attenuation reconstructions of the XCAT phantom in the rotating WT-PET configuration. The systematic error in MLAA is reduced compared to the default WT-PET configuration, especially in the arms, although not entirely gone.

Finally, in Figure 9.8 we show the potential of simultaneous emission and transmission by TOF separation, on a simulation of the IQ phantom in the default WT-PET configuration. Looking at the true

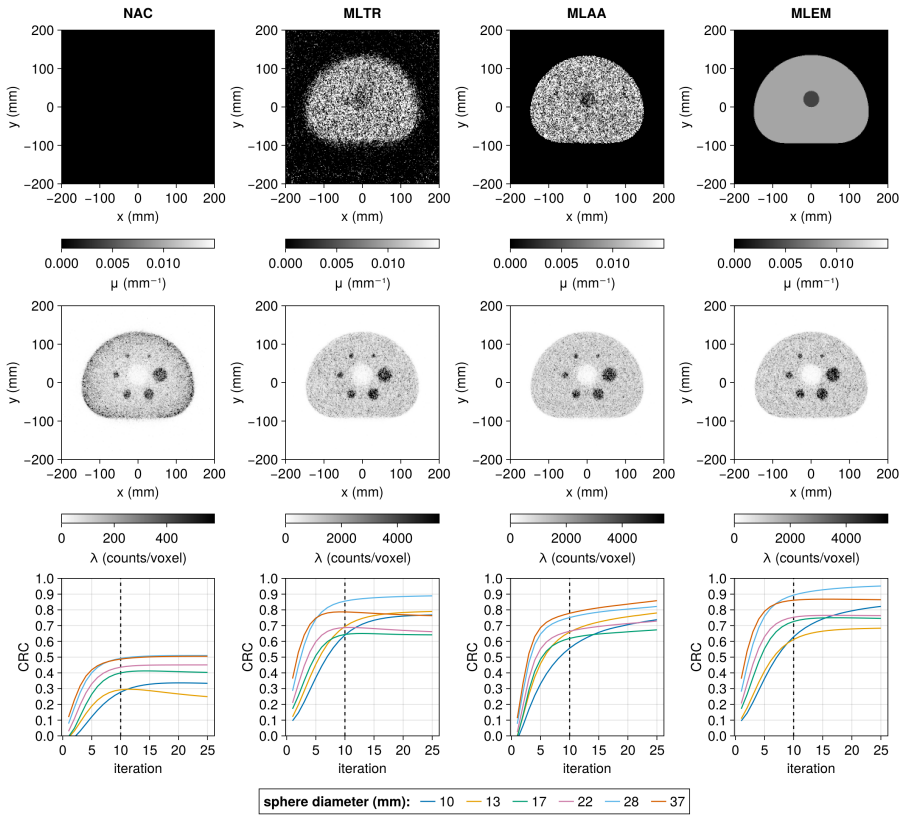


Figure 9.3: IQ phantom reconstructions on the **rotating WT-PET** configuration, comparing CT-less attenuation correction methods (MLTR and MLAA) to uncorrected (NAC) or ground-truth attenuation corrected images (MLEM).

coincidences of transmission and emission data, we see that there is almost no overlap in the TOF histogram (using 400 ps TOF). Separating the transmission LORs from the prompt coincidences based on their TOF difference and the closest distance to any of the transmission sources, we observe good agreement between the predicted and true coincidences in the transmission scan. More quantitatively, only 6.8% of true transmission coincidences are missed in the predictions, 9.5% of predicted true transmission coincidences were in fact random coincidences, and 1.7% of predicted true transmission coincidences were in fact scattered coincidences.

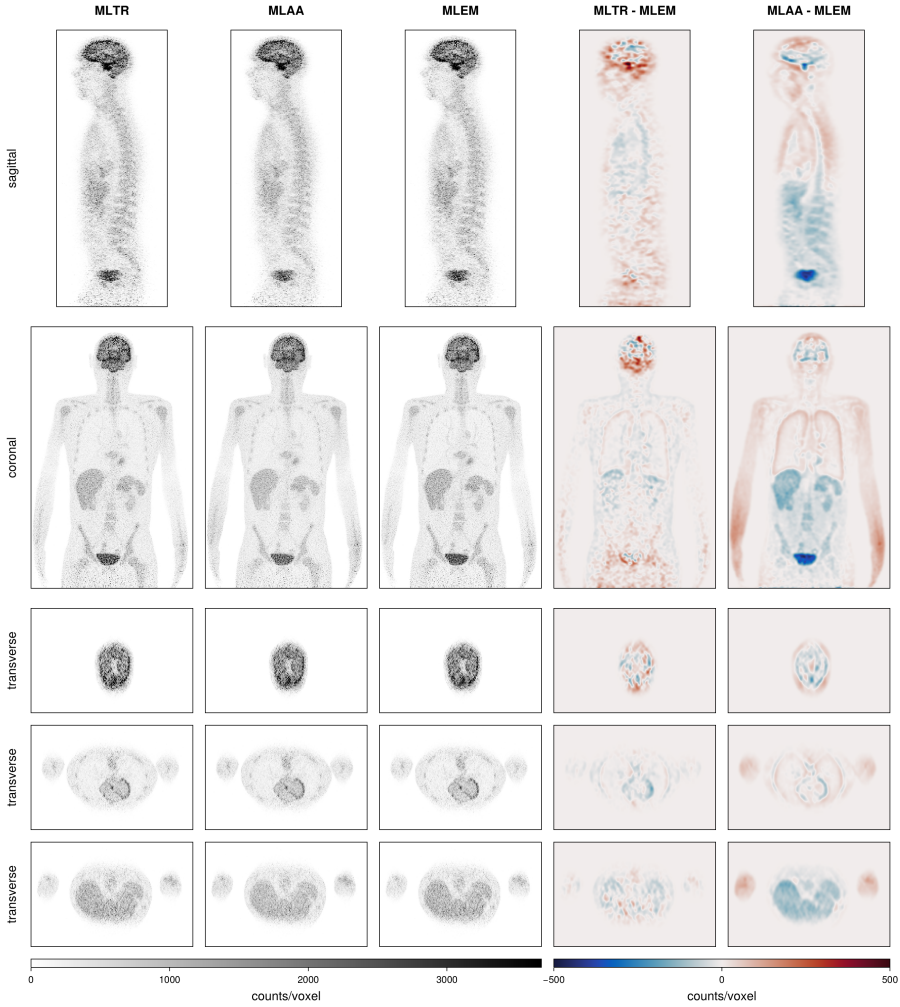


Figure 9.4: Activity reconstructions for the **XCAT phantom** in the **default WT-PET** configuration, comparing different CT-less attenuation correction methods with ground-truth attenuation correction.

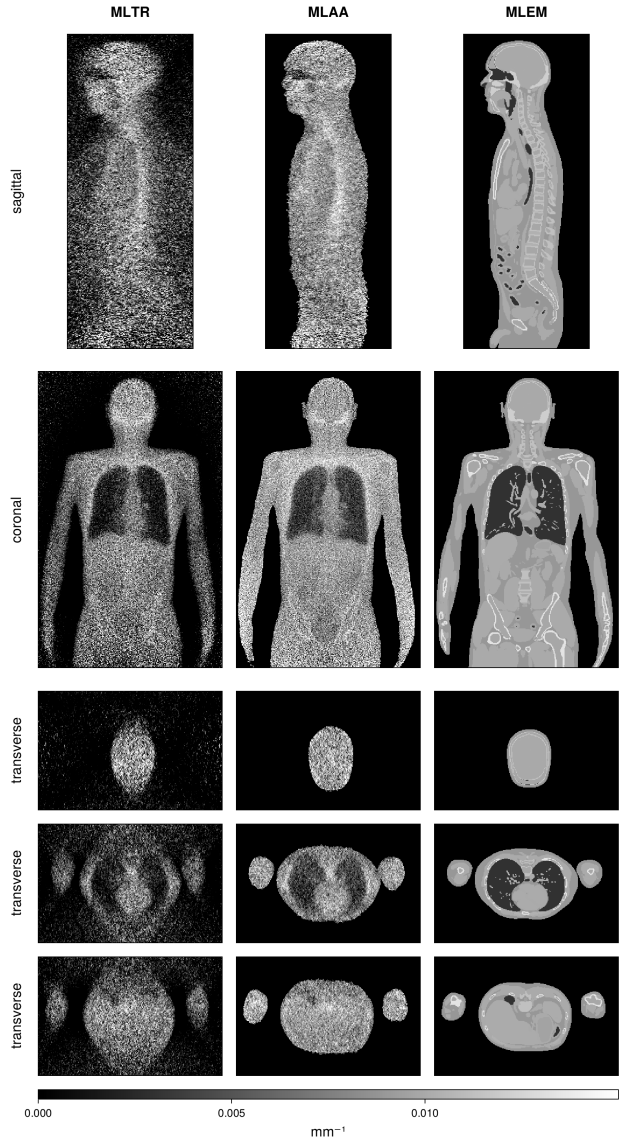


Figure 9.5: Attenuation reconstructions for the **XCAT phantom** in the **default WT-PET** configuration, comparing with the ground-truth attenuation map.

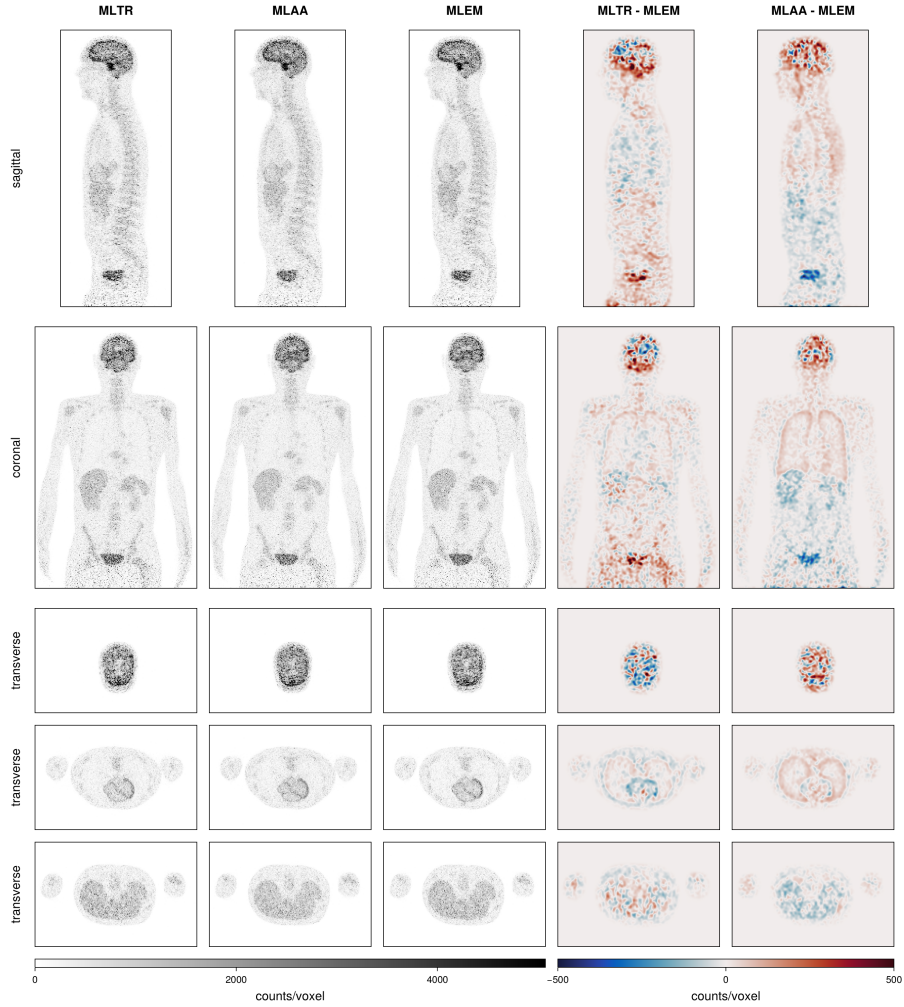


Figure 9.6: Activity reconstructions for the **XCAT phantom** in the **rotating WT-PET** configuration, comparing different CT-less attenuation correction methods with ground-truth attenuation correction.

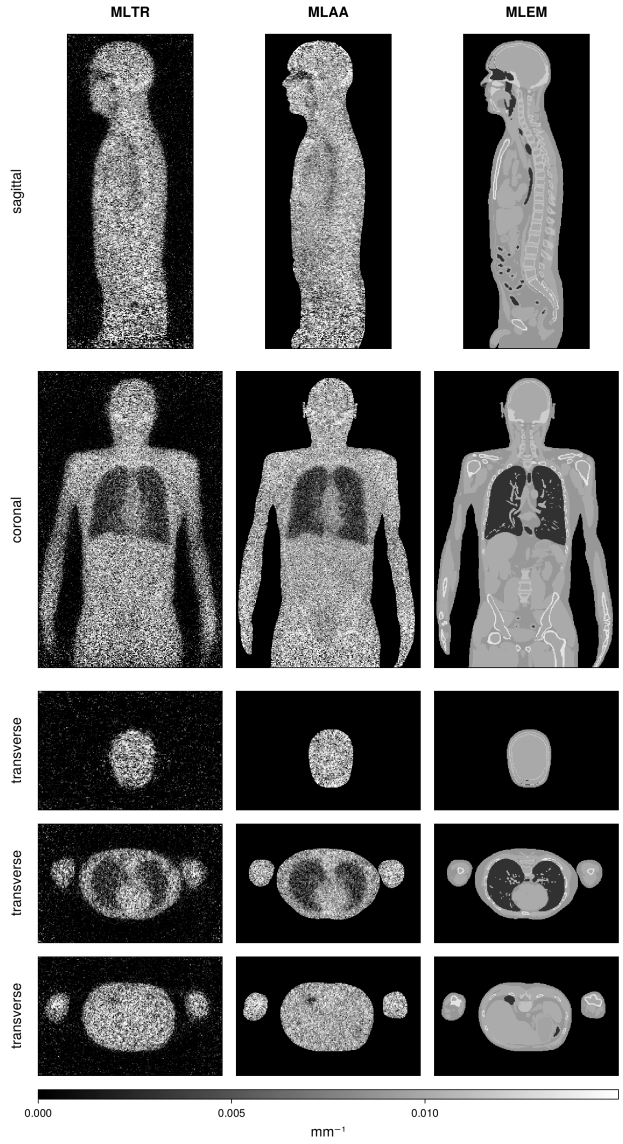


Figure 9.7: Attenuation reconstructions for the **XCAT phantom** in the **rotating WT-PET** configuration, comparing with the ground-truth attenuation map.

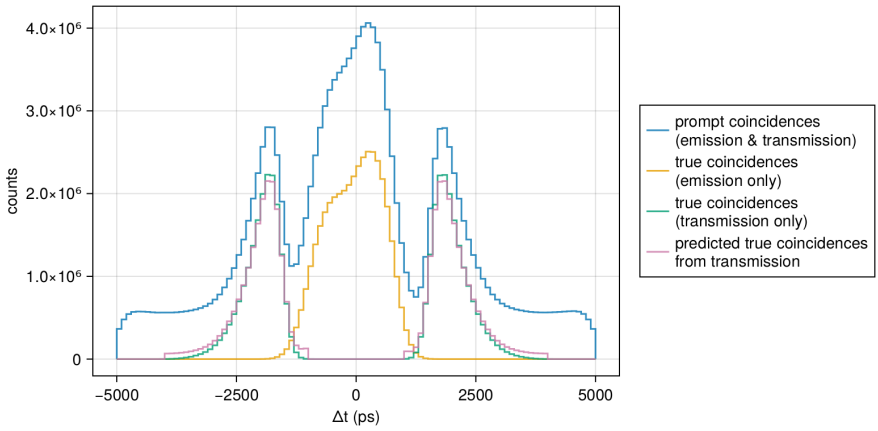


Figure 9.8: TOF histogram of a simultaneous transmission and emission scan on the default WT-PET configuration.

9.4 Discussion

Visually, comparing the reconstructions using the attenuation map generated by either MLTR or MLAA with the reconstruction using the ground-truth attenuation map, we only see minor differences with some loss of detail. Inspection of the difference maps for the XCAT phantom, however, show that there is some systematic error introduced, especially in MLAA. In part, this is because MLAA requires more iterations for convergence, due to the interdependence between activity and attenuation reconstructions. The current lack of regularization, both in the activity and attenuation reconstructions, however prohibits using much higher iteration numbers. A regularization term in the attenuation reconstruction could also be used to penalize unrealistic attenuation coefficients, as we know they should belong to either air, lung tissue, soft tissue or bone. This may solve some of the issues we notice in MLAA, where we for example see that the attenuation coefficients in the arms consistently blow up in value.

Besides some remaining issues with the activity reconstructions for MLAA, we can also see that the obtained attenuation maps have very little detail, and are therefore no substitute to diagnostic CT. Furthermore, in the default WT-PET configuration, the attenuation map obtained with MLTR shows severe limited angle artifacts, due to the lack of TOF information for transmission data. Nonetheless, it results in a good activ-

ity reconstruction, as the attenuation correction factors are still properly estimated, due to the emission LORs lying along the same directions as the transmission LORs. However, this does mean that the attenuation map is unsuitable for e.g., scatter correction with SSS, as scattered LORs no longer lie along the same directions as transmission LORs, and the location of potential scatter points would be incorrectly sampled. These limited angle artifacts are reduced in the case of MLAA, as we can use the TOF information of the emission data to put constraints on the attenuation map.

9.4.1 Advantages and disadvantages of MLTR and MLAA

While the use of a transmission source is undesirable due to the increased radiation dose, the added dose remains rather small. In the specific case we simulated, there is a total of 8 x 3 = 24 MBq of additional activity present, but since this is emitted isotropically, the majority of photons will not pass through the patient. In addition, the patient is only exposed to this dose for the scan duration of 30 s, or realistically, a little longer to include time for the patient to enter and exit the scanner. The additional radiation dose is therefore small compared to a normal patient dose of 3 MBq/kg, to which the patient is exposed for the entire lifetime of the isotope (half-life of 110 minutes for ^{18}F). Although some fraction of the injected dose will be excreted prior to decaying, this radiation is emitted from within the body, so that all emitted photons pass through the patient. The transmission sources do add some additional complexity to the scanner in terms of mechanics, although of course much simpler than the inclusion of a CT. MLTR does show good convergence and has low computational complexity, offering fast reconstructions of the attenuation map. We also note that in theory, we can further reduce the noise in the reconstruction by performing a much longer blank scan, as in that case there are no concerns of radiation dose, and we do not care about scan duration as it only needs to be performed once prior to any patient scans. In such a case, we should apply an appropriate weighting factor to the LORs of the blank scan in equation 9.1.

MLAA on the other hand is more complex, both in terms of convergence and computation time. The need for additional prior information to obtain appropriately scaled activity and attenuation maps poses a major bottleneck for MLAA. Here, we have imposed that the total activity should match that of the MLEM reconstruction (i.e., the total injected activity within the FOV). In practice, this information is of course not

available, except for whole-body scans where the total injected activity equals the activity within the FOV. For the WT-PET scanner, where we image at least the hips, torso and head, it could be feasible to estimate the activity within the FOV from the injected activity, given that the legs contain primarily background activity. Nonetheless, this would introduce some amount of error. Another option is to place an additional object with known activity inside the FOV, and normalize the total activity based on that. This could provide good accuracy, but would of course incur an additional radiation dose and is inconvenient to prepare.

Another disadvantage of MLAA is that a reconstruction takes considerably longer than a normal MLEM reconstruction. We need to update the attenuation map multiple times at every iteration, due to its lack of TOF data leading to slower convergence for the attenuation map. In addition, the MLEM update equation itself also takes longer. The sensitivity image needs to be recomputed for every iteration due to the updating attenuation map, whereas before, it could be precomputed. This adds considerable time, as the number of sampled LORs I, especially for shorter scans, can be substantially larger than the number of emission LORs E. Furthermore, the attenuation correction factors need to be recalculated at every iteration, which requires us to compute the intersection lengths with all voxels intersecting the LOR, whereas with precomputed attenuation correction factors, this could be sped up considerably, as we only need to compute the intersection lengths where the TOF kernel support is non-zero. Despite its drawbacks, MLAA remains an attractive option for CT-less attenuation correction, due to not requiring any additional setup or acquisitions, and not increasing radiation dose.

9.4.2 Random and scatter correction

Although for our reconstructions, the transmission and emission scans were simulated separately for MLTR, Figure 9.8 does show the feasibility of separating the transmission data from the emission data. Moreover, by including a second condition based on the closest distance between the LOR and any transmission source, we can reject the majority of random and scattered coincidences, allowing us to use equation 9.1 valid for true coincidences, even when dealing with data containing randoms and scatters. These rejected random and scattered coincidences from the transmission data will, however, contaminate the emission data used for the activity reconstruction. Removing these from the emission data

based on their TOF difference is not possible, as scattered coincidences have more spread out TOF differences, and random coincidences show no correlation with TOF difference. Instead, we can simply perform the activity reconstruction assuming that the simultaneous emission and transmission scan are, combined, a normal emission scan. Random and scatter correction can then be performed as usual, based on all acquired data.

As for MLAA, the update equations 9.2 can be modified to include the random and scatter contributions r_i and s_i [152]. The noise contributions are added to the update equation for λ as is done in MLEM, but also the update equation for μ needs to be modified. Again, energy based scatter estimation would be a good choice here, as it does not require an attenuation map.

9.4.3 Alternative options

As previously mentioned, another option is to use deep learning approaches to estimate either the attenuation map, or the attenuation corrected PET image, from an uncorrected image. Similar to MLAA, this would not incur any additional radiation dose, and inference would be very fast. We would however need training data, which at this point is only available from simulation. An algorithm trained only on simulated data is unlikely to perform as well on experimental data. Even though the GATE simulation is very detailed, not all sources of noise or scanner non-idealities are modelled. Furthermore, the XCAT phantoms, while based on realistic experimental data, are rather uniform between one another, which may result in overfitting on the training data. It would therefore be ideal to include experimental training data. Using data from other existing scanners may not be very useful, given the rather different geometry of the WT-PET. It would therefore be ideal to train such a network on patient data obtained from the WT-PET itself, with integrated CT. Until this is possible, it may be preferable to use these non-deep learning based methods for attenuation correction.

9.5 Conclusion

With these CT-less attenuation correction methods, and with the energy based scatter correction method from Chapter 8, we now have the necessary tools to enable CT-less image reconstruction on the WT-PET. One remaining issue regarding the image reconstruction, are the limited

angle artifacts introduced by the flat panel geometry, as seen throughout these last few chapters in reconstructions for the non-rotating WT-PET configurations. This will be addressed in Chapter 10, using a deep learning based approach embedded in the iterative image reconstruction procedure.

Chapter 10

Limited angle artifact correction

10.1 Introduction

In this chapter, we take a closer look at one of the remaining limitations of the WT-PET: the presence of limited angle artifacts due to missing projection data between the gaps of the panels. For tomography, 2D projection data over an angular range of 180° is sufficient to provide a unique solution to the inverse reconstruction problem, if we disregard physical limitations such as gaps between detectors, limited detector resolution, count statistics and other resolution degrading or noise inducing processes. In such a case, the transition to 3D projections does not provide any new information, and that data can be considered redundant for image reconstruction. However, in a real scanner with physical limitations, the additional data provided by 3D projections helps to improve the reconstructions, for example by increasing count statistics. When part of the 2D angular range is missing, as is the case in the flat panel geometry of the WT-PET, the available data is insufficient to provide a unique solution, even in the ideal case (i.e., we are dealing with data incompleteness). In that instance, additional 3D projections provide valuable information that constrain the solution space. However, unless the angular range along the third dimension is itself at least 180°, a unique solution still does not exist, resulting in the artifacts observed during the reconstruction. The observed reconstructions are the most likely (in terms of the log-likelihood) activity distributions for the given projection data, although other activity distributions exist that also agree with the

measured data (including the real, artifact-free distribution). Similar to 3D projection data, the inclusion of TOF information puts a constraint on the annihilation position along the LOR, providing information on the activity distribution perpendicular to the projection data, further constraining the solution space. With perfect TOF, a unique solution could be found from only a single projection angle, however with the TOF resolutions of current PET scanners, the spatial resolution along the LOR remains limited to at least a few cm ($\Delta x = c\Delta t/2$).

For the WT-PET, this missing projection data results in a smearing of the reconstruction, largely perpendicular to the panels, as observed in previous reconstructions. We could also see that these artifacts are more pronounced in the brain region, which is located near the edge of the AFOV. The available angular projection range, in the transverse and sagittal views, is shown schematically in Figure 10.1. At the center of the FOV, along the transverse plane, only 112° / 180° (= 62 %) of the projection range is available, and in the sagittal plane, 129° / 180° (= 72 %) is available. We note that since the panels are taller than they are wide, we actually have a wider projection range in the sagittal plane, so that projections partially along the axial direction contain a considerable amount of additional information, constraining the extent of artifacts more than would be expected purely from the 2D transverse views. However, we can also see that the available projection range decreases when moving away from the center of the FOV, aggravating the artifacts. So far, this has been most notable in the brain region. As was also shown in Figure 7.10, the inclusion of TOF information reduces the extent of these artifacts. Nonetheless, even at 200 ps TOF ($\Delta x \approx 3$ cm), the artifacts are not entirely gone, resulting in reduced interpretability of the PET image. Nuclear medicine physicians are not trained and/or used to seeing images with such artifacts, which would pose an obstacle for the adoption of the WT-PET. We should therefore develop a method to correct for these artifacts.

A number of methods have previously been developed for dealing with missing projection data, such as sinogram interpolation [153] or constrained Fourier space methods [154]. These methods seem to work quite well when the angular gaps are small, but potentially numerous (i.e., sparse designs). However, it is unclear how they would perform when the gaps are as large as they are in our case. Furthermore, sinogram based methods would require LOR discretization of our continuous list-mode data, which we wish to avoid where possible, to maintain high

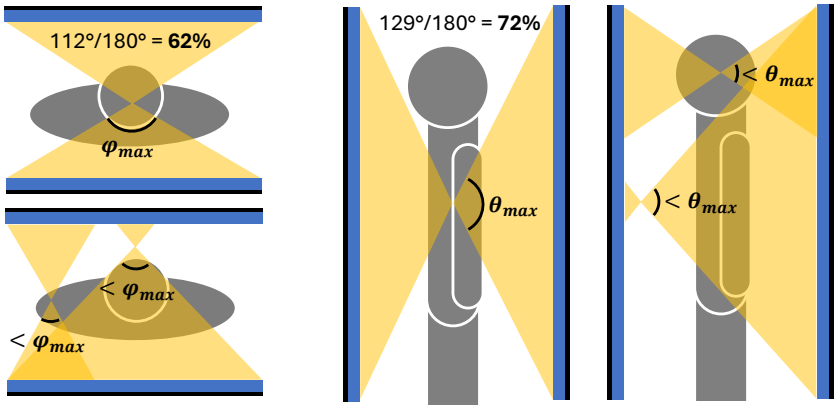


Figure 10.1: Schematic visualization of the available projection ranges in the WT-PET, showing how the available range decreases when moving away from the center of the FOV.

spatial resolution. More recently, deep learning based methods have been proposed for artifact correction in limited angle tomography, both in sinogram space [86], [87] and image space [77], [88]. For PET, most research has been on sparse designs with smaller gaps, but for CT, a number of studies have investigated larger missing angular ranges [155], [156], or even generating 3D images from 2D X-Rays [157], [158].

One important consideration with deep learning based methods is that we need to ensure that the algorithm will not make any false predictions. The fact that a predicted image looks realistic, is by itself no guarantee that the image in question offers an accurate representation of the real tracer distribution. One way to better guarantee that these algorithms will provide solutions that agree with reality, is to incorporate them into the iterative reconstruction procedure. This ensures that the predictions agree with the measured projection data, by using the measurements at every iteration to update the image. In this chapter, we will develop a deep learning algorithm for limited angle artifact removal in image space for the WT-PET, and integrate it in the reconstruction procedure as a regularization term to ensure data-consistency.

10.2 Materials and Methods

10.2.1 Implementation in reconstruction

As was discussed in the introductory chapters, a regularization term can be included in the objective function (equation 2.5) to penalize solutions that do not match certain a priori assumptions about the image properties. This transforms MLEM into a maximum a posteriori expectation maximization (MAP-EM) algorithm. We would like this regularization term R to penalize images that contain limited angle artifacts, or otherwise do not correspond to the expected, artifact-free, reconstruction. To achieve this, we can use a quadratic penalty such as:

$$R(\lambda) = \sum_{j \in J} (p_j - \lambda_j)^2 \tag{10.1}$$

with voxels $j \in J$, and $\mathbf{p}$ a prior image from which the reconstruction λ should not deviate too far [159]. In our case, we wish for $\mathbf{p}$ to be the artifact-free version of λ. Using this quadratic prior, a solution to the penalized optimization problem can be found iteratively, by a combination of the prior image $\mathbf{p}$, and the standard EM update image $\lambda_{j,EM}^{(k+1)}$ obtained from equation 7.1 [160]:

$$\lambda_j^{(k+1)} = \frac{2\lambda_{j,EM}^{(k+1)}}{(1 - \frac{\beta p_j}{S_j}) + \sqrt{(1 - \frac{\beta p_j}{S_j})^2 + 4\frac{\beta \lambda_{j,EM}^{(k+1)}}{S_j}}} \tag{10.2}$$

where $S_j = \sum_{i \in I} H_{ij}$ is the sensitivity image and β is a parameter that controls the strength of the regularization. It can be seen that when $\beta = 0$, the right-hand side reduces to $\lambda_{j,EM}^{(k+1)}$, and we are therefore left with the MLEM update equation.

In order to obtain the prior image $\mathbf{p}$, we use an artifact correcting neural network, which removes the limited angle artifacts from the current image estimate, $F : \lambda^{(\mathbf{k})} \rightarrow \lambda^{(\mathbf{k})}_{\text{artifact-free}}$:

$$\mathbf{p} = \mathbf{F}\left(\lambda^{(\mathbf{k})}\right) \tag{10.3}$$

Therefore, the prior image itself is updated at every iteration.

As we iterate, the limited angle artifacts in $\lambda^{(\mathbf{k})}$ should be reduced or eliminated due to the regularization. As the prior image $\mathbf{p}$ is obtained from $\lambda^{(\mathbf{k})}$, the network F should produce artifact-free images both for

inputs with, and without, artifacts. The extent of the artifacts may also vary through the iterations, as we expect the artifacts to gradually diminish, depending on the strength β of the regularization term. This means that the network should be trained on a variety of data with varying severity of limited angle artifacts, to ensure that the image $\mathbf{p}$ is correctly predicted at every iteration.

10.2.2 Training data

For training of the neural network F, we use image reconstructions obtained from GATE simulation of 11 different XCAT phantoms, with varying genders, heights and BMIs. A 12th phantom is kept separate for validation (to monitor training convergence), and a 13th phantom is used for testing. For the training targets, we make use of the rotating WT-PET configuration, which provides full angular coverage, but otherwise has similar properties to the default WT-PET configuration. To include the aforementioned variety in artifacts in the training data, we use as inputs not only the reconstructions on the default WT-PET configuration, but also reconstructions on the rotating configuration. On the rotating WT-PET, we can vary the angular range by limiting the rotation of the panels to less than 180°. To this end, we reuse the same simulation, and simply filter the list-mode data based on the event timestamps. Furthermore, we train the network on different iteration numbers, since computation of the prior will also be done at different iterations, and therefore, at different levels of convergence. Finally, we also train using different TOF values, as better TOF results in reduced artifacts and faster convergence, providing an additional way to vary the range of artifacts and convergence.

The GATE simulations and image reconstructions were all performed as in previous chapters, but again using only true coincidences. In order to ensure that our training labels are not noisier than the inputs, we used 45 s acquisitions for the targets, rather than the 30 s used for the default WT-PET configuration. This value was chosen because the increased gap between the panels required for the rotation results in a loss of sensitivity of about 33%. A summary of the parameters varied for generating the training data is given in Table 10.1.

	input	target
phantom	11 XCAT phantoms	= input
TOF	100, 200, 400 & 800 ps	= input
iteration	1 - 10	= input
geometry	default[a], rotating 22.5°, 45°, 90°, 180°	rotating 180°
acquisition	30 s, 5.125 s, 11.25s, 22.5 s, 45 s	45 s

[a]Non-rotating, non-sparse configuration of the WT-PET.

Table 10.1: Summary of the training data used. The target is always an artifact-free image at a similar level of convergence.

10.2.3 Network architecture and training

We use a U-Net for the neural network, with the details of the architecture shown in Figure 10.2. It is a 2D network (i.e., 2D kernels for the convolutions) that operates on three consecutive transverse slices, to remove the limited angle artifacts from the central slice. We use three input slices, rather than one, to provide some additional statistics for the input in the presence of noise. Given the structural similarity between input and target, we use a residual connection between the central input slice and the output, so that the network only needs to predict the residual artifacts ($\lambda - \lambda_{\text{artifact-free}}$). The network architecture and training is implemented in Julia, for easy integration within the reconstruction code. The Adam optimizer is used with the mean squared error (MSE) as loss function, and training is done for only a single epoch. We have found one epoch to be sufficient for convergence, given the large number of reconstructions available (11 XCAT phantoms x 4 TOF values x 10 iterations x 5 geometries = 2200), each containing 530 slices for a total of $\sim$ 1.2 million images. Therefore, although we are only using a total of 11 XCAT phantoms for training, we do end up with many data points for training in total (although admittedly, many of these will be quite similar, e.g., iteration 9 and 10 of the same slice will not show many differences).

10.2.4 Evaluation

We evaluate the performance of the algorithm on the 13th XCAT phantom, a 59-year-old female, 162.7 cm tall and weighing 63.7 kg (BMI = 24.06), see Figure 7.6 for the ground-truth activity map. This is primarily done visually, but we also include 5 artificial lesions in the liver

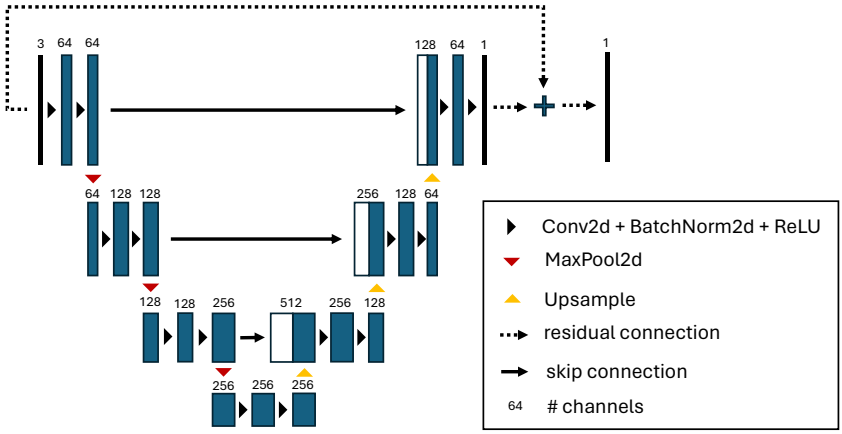


Figure 10.2: 2D U-Net architecture used for limited angle correction.

to study the effect on lesion contrast. The spherical lesions are 10 mm (x 2), 7 mm and 5 mm (x 2) in diameter, with a lesion to background activity concentration ratio of 8:1. We evaluate both applying the neural network after normal MLEM reconstruction as a post-processing step, and integrating the neural network within the iterative reconstruction procedure as outlined before. Finally, we also evaluate the performance on the IQ phantom, which was not a part of the training data.

10.3 Results

Figure 10.3 shows the visual results on the 13th XCAT phantom, where we compare the MLEM reconstruction (input) with the DL post-processed results (output) and the MAP-EM reconstruction. The 45 s acquisition on the rotating WT-PET configuration (i.e., the target if it were part of the training data), as well as a 90 s acquisition with 50 ps TOF on the rotating WT-PET, are given as reference. For MAP-EM, we used $\beta = \beta'/\lambda$ with $\beta' = 10^{-3}$. We have found that voxel-wise scaling of β, based on the current image estimate intensity, worked better to ensure that limited angle artifacts were reduced at similar rates across the FOV. Otherwise, the brain region would converge faster than the thorax region. All reconstructions shown are at the 10th iteration. We note that the lesions were not included in the rotating WT-PET simulations, and are therefore not visible. Visually, there is good agreement between the DL

results and the reference images, showing that limited angle artifacts are largely corrected for without "hallucinations" from the network. MAP-EM shows better retainment of sharpness, at the cost of increased noise. While the lesions remain visible in both cases, loss of contrast is observed, especially for the smallest lesions in the post-processed DL results. This is also confirmed quantitatively in Figure 10.4, where we compare the contrast recovery coefficient for the different lesions between MLEM, MLEM with DL post-processing, and MAP-EM (for a number of values for β').

Figure 10.5 shows the results on the IQ phantom, where we again compute the CRCs for the different sphere sizes in function of iteration, and visualize the reconstruction at the 10th iteration. Again, we observe better contrast recovery for MAP-EM compared to MLEM + DL post-processing, especially for the smaller sphere sizes. Visually, we can however see that, while limited angle artifacts are reduced, and effectively eliminated at the phantom edges, some stretching remains visible around the inner cold region.

10.4 Discussion

We have found that our methodology is capable of removing the majority of artifacts in the XCAT phantom, producing images that structurally match the artifact-free reconstructions from the rotating WT-PET configuration. We also notice that the neural network has a denoising effect, even though we did not use higher count targets. This is because the network learns to predict the average voxel values, as it is difficult for the network to fit to noise due to its inherent randomness. The network does have a tendency to introduce blurring, and thereby also reduces the contrast of lesions. This is partially counteracted when integrating the network within the iterative reconstruction procedure, as contrast is reintroduced at each iteration based on the measured projection data. The loss in contrast is also likely to be dependent on the position in the FOV, as the extent of limited angle artifacts and the strength of corrections performed by the neural network vary within the FOV. From the IQ results, we can see that, while the artifacts are reduced, they are not entirely gone: the network does not generalize well on non-XCAT data. While implementation within the iterative reconstruction does provide some confidence that the reconstruction matches the measured projection data (especially when β is relatively small), it is no guarantee that the

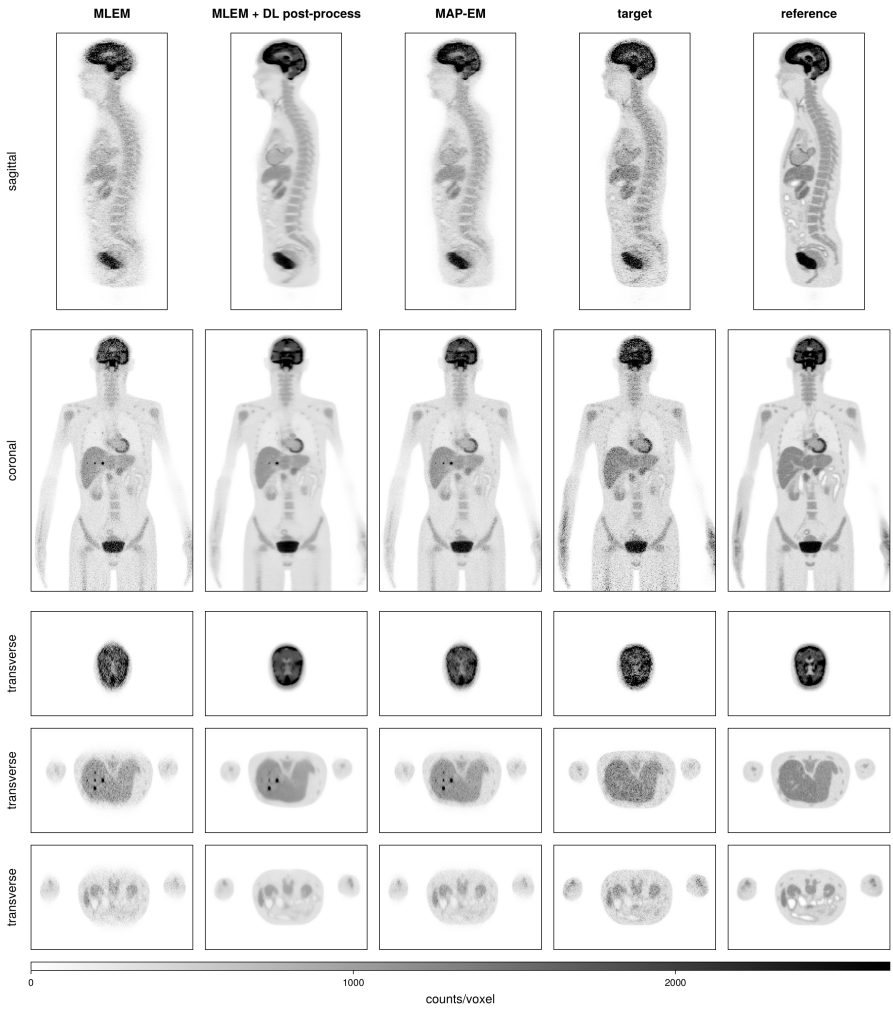


Figure 10.3: Limited angle artifact correction for the XCAT phantom, comparing implementation as a post-processing step or within the iterative reconstruction algorithm. All reconstructions are visualized at the 10th iteration. The simulation for the target and reference image did not contain any lesions, but otherwise used the same activity distribution for the phantom.

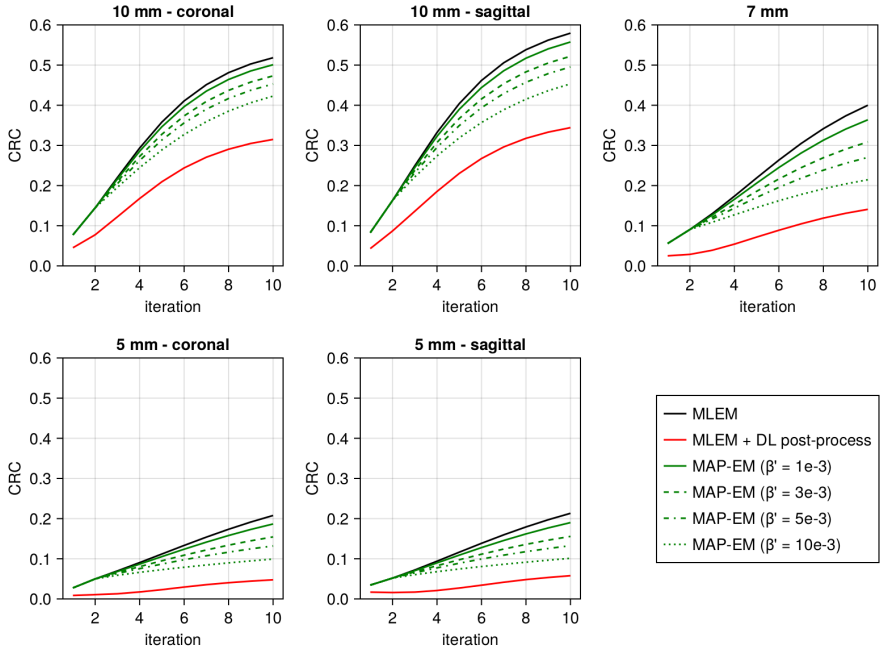


Figure 10.4: Contrast recovery coefficients (CRCs) for the 5 lesions in the liver, comparing the different limited angle artifact correction implementations to the uncorrected case (MLEM reconstruction of the default, non-rotating WT-PET configuration).

result will be artifact-free. After all, both the image with and without artifacts are a potential solution that would result in the measured data.

It also remains to be investigated whether a 3D U-Net would provide better results. On the one hand, we know that the projection data in PET is not 2D, and artifacts within a specific transverse slice are therefore expected to contain contributions (i.e., activity leakage) from adjacent slices. Although we provide three consecutive transverse slices to the network for prediction, this only equals 6 mm of axial data when using 2 mm voxel sizes. This is much smaller than the spatial accuracy provided by the TOF information ($\sim$ 6 cm at 400 ps TOF), and is therefore insufficient to contain all data that could potentially impact the transverse slice in question. On the other hand, training on 3D data is very memory intensive, which would require us to resort to patch-based approaches, depending on the number of network parameters. The TOF

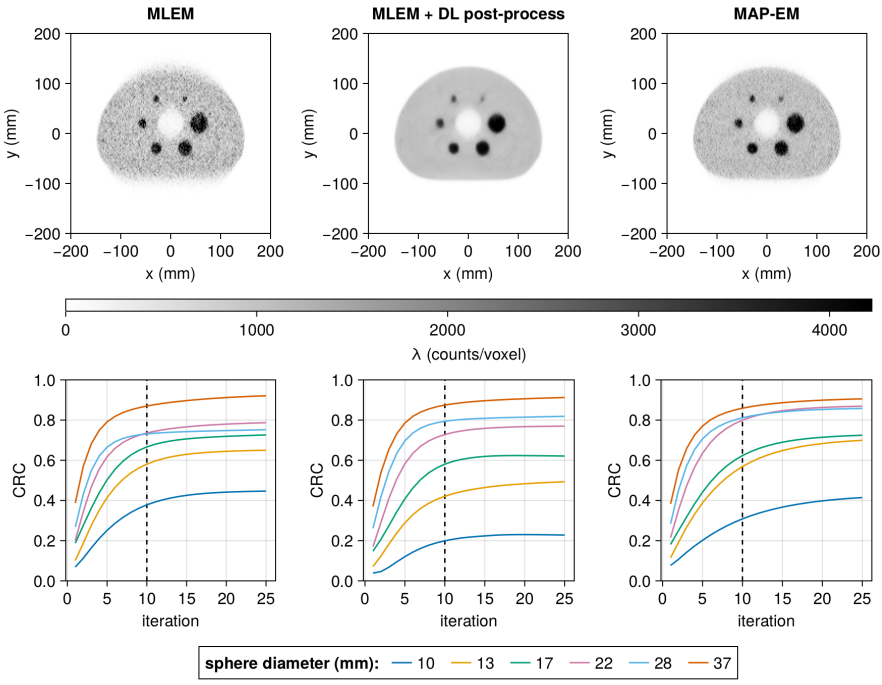


Figure 10.5: IQ phantom reconstructions with limited angle artifact correction, showing the 10th iteration, as well as contrast recovery coefficient (CRC) for the different spheres in function of iteration.

information does constrain the regions of space that can influence each other, so that we do not necessarily require the network to see the entire image at once. Nonetheless, making predictions at the edges of a patch would be difficult, seeing as there we would require information from outside the patch.

We note that an alternative to our regularization approach would have been to use the (typically more complex) algorithm unrolling methods discussed in Section 3.2.3. Instead, we have opted to start simpler, and if necessary, other approaches can be studied in the future. At this point in time, the biggest limitation is likely to be the lack of phantom variety in the training data, not the methodology itself, resulting in an over-fitting on the XCAT phantoms. While we have used different XCAT phantoms of different BMIs, they were still very uniform in terms of, for example, tracer distribution. The relative activity concentration ratios between different organs was very similar between phantoms, and the distribution

itself was entirely uniform for each tissue. Furthermore, the phantoms were all positioned the same within the scanner, with the head at the end of the AFOV. Given that the extent of the artifacts depends on the position within the FOV, this could result in the network struggling for different phantom positions.

The solution to this is, however, not so straightforward. It will be difficult to obtain good experimental training data, given that there are currently no plans to build a rotating configuration of the WT-PET. When the WT-PET is built, we could potentially obtain training data pairs by having patients scanned subsequently in the WT-PET and a normal, cylindrical PET system. However, the patient would be positioned differently in each scanner, which would render the paired data useless. Unless some form of image co-registration procedure is used, which is itself prone to error, and would be more difficult due to the presence of limited angle artifacts. Another option would be to use only data from a cylindrical PET scanner, and generate limited angle reconstructions from a normal acquisition by removing projection data. This has the issue that the limited angle artifacts will not entirely match those in the WT-PET, given that one geometry would consist of two arcs, and the other of two flat panels. Furthermore, no clinical PET system exists with a resolution as high as the WT-PET, so training would need to be done on lower resolution data. Another option could be to use simulated data, not using the XCAT phantoms, but using real PET reconstructions as the ground-truth activity distribution in GATE. While the training data would still be lower resolution, it would allow generating more realistic data specific to the WT-PET geometry.

10.5 Conclusion

With the end of this chapter, we have finished covering the efforts performed within the scope of this dissertation towards developing an image reconstruction framework for the WT-PET. While there certainly remains room for further improvements and additions, the image reconstruction software is already being used by other colleagues for their own studies regarding the WT-PET. In Chapter 11, we will provide some concluding remarks for the thesis as a whole, and cover some of these potential future developments we have in mind for image reconstruction.

Chapter 11

Conclusions and future perspectives

This dissertation has focused on two main parts: enhancing the timing resolution in monolithic detectors, and image reconstruction for the WT-PET: a flat panel, long AFOV PET system based on such monolithic detector technology.

11.1 Timing capabilities of monolithic detectors

For the scintillation material of PET detectors, there are two main choices: LYSO or BGO. LYSO is known to have superior scintillation characteristics, with a faster scintillation pulse and more photons generated per gamma event, which translates to better energy, spatial and time resolutions for the detector. The superior time resolution is important for TOF-PET, and is the reason why the vast majority of current PET scanners use L(Y)SO. For limited angle tomography, the TOF resolution becomes even more important. Meanwhile, BGO has the advantage of being considerably cheaper, and has recently gained renewed attention due to the potential for good time resolutions by making use of Cerenkov radiation. However, studies investigating Cerenkov production in BGO have predominantly focused on pixelated detectors.

In Chapter 4, we have studied the potential effect of Cerenkov radiation on time resolution in monolithic BGO, where we also looked into the effect of the crystal surface finish, and made a comparison with LYSO and pixelated detectors. We have found that the surface finish has a large impact on photon collection efficiency, that is, the fraction of

emitted photons detected by the SiPMs. As expected, a reflective side surface finish increases collection efficiency by reducing the number of photons that escape the detector. However, the surface finish on the side of the photodetector is also important. Using a rough, rather than polished, surface finish increases the transmittance of photons from the crystal to the optical coupling grease at high incidence angles, at the cost of reduced transmittance at low incidence angles. Overall, this leads to a gain in collection efficiency, due to the isotropic emission of scintillation photons, resulting in more photons of high incidence angle with respect to the photodetector surface normal. The same is observed for Cerenkov photons, which are also emitted close to isotropically due to a combination of the short mean free path of the recoil electron, the fact that the electron itself will not be emitted in the exact same direction as the incoming gamma photon, and the possibility of the gamma photon itself having been scattered. The gains in photon collection efficiency can be substantial: it is more than doubled when switching from a rough black side surface and polished photodetector surface, to a polished reflective side surface and rough photodetector surface. The corresponding gain in time resolution is, however, more modest, and primarily manifests itself in the tails of the TOF kernels, with improvements of around 10% - 15% in terms of FWTM in monolithic BGO for the aforementioned switch in surface finishes. In comparison, time resolution of monolithic LYSO is less affected by the surface finish and corresponding increases in photon collection efficiency. This is because the increased collection efficiency is specifically for reflected and/or high angle-of-incidence photons, which possess longer transit times in the crystal. Since there is already a high number of direct scintillation photons detected early on in monolithic LYSO, most of these additional photons will be detected too late to matter. In monolithic BGO, relatively few direct photons are detected early on due to its slow scintillation pulse, so that these additional indirect photons, if emitted early on, can improve time resolution.

Optimizing the surface finishes is especially important for Cerenkov based time estimation, due to the very small number (~ 18) of Cerenkov photons generated per 511 keV photoelectric event in BGO. Nonetheless, even with the ideal surface finishes, the individual SiPMs in monolithic detectors almost never detect more than a single Cerenkov photon per event due to the light spread. In order to trigger on these early Cerenkov photons, we therefore need to choose a threshold level below the single photoelectron level. This contaminates the data with dark counts, neg-

atively impacting the potential time resolution. Cerenkov based time estimation is therefore easier to achieve in pixelated BGO, where the light is concentrated onto one SiPM, so that Cerenkov photons may be observed above the single photoelectron level.

The estimated time resolutions in Chapter 4 were obtained using the first SiPM timestamp, acquired by leading edge discrimination on the SiPM signals. We only used the first timestamp there, because for BGO, we did not see any improvement when averaging over multiple timestamps. Again, this can be attributed to the lower light yield and slower scintillation pulse of BGO, which results in later timestamps containing less relevant information. As seen in Chapter 5, we do notice substantial gains in time resolution for LYSO when averaging over the first few timestamps. This could be improved by feeding the timestamp matrix to a neural network predictor (a 2D CNN), showing a 17% improvement compared to timestamp averaging: from 177 ps FWHM to 151 ps FWHM in CTR. Better yet was to supply the leading edge portion of the digitized SiPM waveforms to a 3D CNN, in order to gain access to additional temporal information, reaching 141 ps FWHM in CTR, an improvement of 26% over timestamp averaging. The primary drawback of this approach is, however, the requirement of waveform digitizers, making the electronics more expensive compared to analog readouts based on voltage thresholding. This makes the 2D CNN method more feasible in practice, while still providing better time resolution and uniformity compared to conventional methods. We do reiterate that in order to reach these values, we performed a baseline correction on the digitized signals for all methods, so that the effect of dark counts was minimized. In a fully analog readout, the effect of dark counts is likely to be larger, degrading the time resolution. Nonetheless, the method comparison should remain valid in experiment, with a 2D CNN improving performance compared to timestamp averaging.

11.2 Detectors for the Walk-Through PET

As was mentioned in Chapter 6, the original design of the WT-PET was based on monolithic BGO to reduce cost, using detectors capable of reaching 327 ps TOF by means of neural network time and positioning estimation, based on experimental results from a different group [132]. This is also why our simulations in part II were based on monolithic BGO, with a TOF of 400 ps, representing a conservative estimate of a system

level TOF resolution using those detectors. In the end, however, the decision was made to go with LYSO for better timing performance, where a system level TOF resolution of 200 - 300 ps should be achievable using suitable electronics. The superior time resolution is especially important in the WT-PET system, to limit the extent of limited angle artifacts. In doing the switch to LYSO, we introduce sparsity (missing detector rows) into the design in order to offset the increased cost of LYSO, without needing to reduce the AFOV. The accompanying sensitivity loss is partially offset by an increase in "effective" sensitivity due to improved TOF. For the electronics, we will use the TOFPET2 ASIC (application specific integrated circuit) from PETsys [117]. This ASIC provides 64 electronic channels, both for timing (using voltage thresholding) and energy (using charge integration). The ASIC also provides a built-in analog method for dark count rejection. We will be using a methodology similar to the 2D-CNN proposed in Chapter 5 for gamma interaction time estimation.

11.3 Image reconstruction for the Walk-Through PET

Due to the unique design of the WT-PET (monolithic detectors and flat-panel geometry), we could run into limitations using existing image reconstruction software, which is normally designed for cylindrical PET scanners based on pixelated detector technology. To truly benefit from the superior detector spatial resolution provided by monolithic detectors, LORs should be represented using continuous spatial coordinates. Furthermore, inclusion of DOI information and the flat-panel geometry can complicate the calculation of the system matrix and the normalization procedure. The main differences between our reconstruction software, and (most) other existing software packages, are therefore (i) the continuous LOR representation, (ii) the use of a virtual scanner geometry to simplify computation of the system matrix of complex geometries, with or without DOI, and (iii) the fact that we include both the solid angle (within $\mathbf{H_{geo}}$) and the detector attenuation ($\mathbf{H_{det}}$) within the system matrix, providing good uniformity prior to any normalization, including for sparse geometries. In addition, reconstructions are fast due to the GPU-based implementation.

The results from Chapter 7 show that good image quality can be obtained in short, 30s acquisitions, due to the high sensitivity of the

scanner. As expected though, there are limited angle artifacts along the transverse and sagittal slices. We showed how these are reduced with improved TOF, and are almost entirely gone at 200 ps TOF, at least for the IQ phantom positioned in the center of the FOV. This reaffirms our decision to make the change to LYSO for better time resolution.

The reconstructions from Chapter 7 are idealized in the sense that we have eliminated the random and scattered coincidences from the list-mode data prior to reconstruction, by making use of the ground-truth labels available from the GATE simulation. This is not possible in practice, where we need to estimate these contributions from other available information. In Chapter 8, we have implemented this using an energy-based scatter correction method. The energy-based method offers an alternative to the commonly used SSS algorithm, with the main advantage to us being the fact that it does not require an estimate of the attenuation map, normally derived from a CT scan. The ability to perform scatter correction without a CT is important, given that the CT-component for the WT-PET is only in the earlier stages of development and will not be available at the moment the PET component is built. Even when the CT is finished, it will be desirable to have the option of CT-less acquisitions for dose reduction purposes. The results for energy-based scatter correction showed good performance, with minimal degradation compared to reconstructions using true coincidences only.

Likewise, we have looked at CT-less attenuation correction methods in Chapter 9, where we focused on two methods: using an external transmission source (with the MLTR algorithm), or estimation of the attenuation map from the TOF emission data itself (with the MLAA algorithm). In general, we obtained better convergence using the external transmission source, for which we used 8 ^{68}Ge rod sources positioned horizontally in front of the panels, translating vertically during an acquisition. The MLAA algorithm, on the other hand, showed more issues with convergence and some systematic overestimation of the activity in the upper torso and arms, and underestimation in the abdominal region and hips. This seems to be caused by an over/under-estimation of the attenuation coefficients in the corresponding regions, although we are not entirely sure why this happens. Incorporating some form of regularization into the attenuation update of the MLAA algorithm could potentially improve this, for example, by giving a penalty to unrealistic values of attenuation coefficients (e.g., values that do not belong to lung tissue, soft tissue or bone). Since MLAA only determines the activity

up to a constant, an additional constraint is required, such as fixing the total activity within the FOV. Despite these issues, MLAA remains an interesting choice for further investigation, as it requires no additional radiation dose or equipment.

With the switch to LYSO for the final scanner design, a third option for CT-less attenuation correction becomes available: using the intrinsic radiation of ^{176}Lu in LYSO for a transmission scan [161]. The principle is the same as when using an external transmission source, only now, the PET detectors themselves are the transmission source. These transmission events can be distinguished from emission data based on their predictable TOF difference and specific energy spectrum, with the transmitted photon showing peaks at 202 and 307 keV. Similarly to MLAA, this would have the advantage that no additional radiation dose or equipment is required, since the intrinsic radiation is there, whether we make use of it or not. Furthermore, the switch to LYSO would also improve the MLAA algorithm due to the better TOF resolution, improving convergence and allowing us to better estimate the patient boundary for the attenuation map in case of limited angle tomography. Finally, MLAA could be combined with the transmission scan from LYSO, for example, in case the statistics of the intrinsic ^{176}Lu radiation are too low. In such a case, the attenuation coefficients derived from the intrinsic LYSO transmission could also be used as a global constraint to obtain correct activity and attenuation values from MLAA. We may need to use such an approach, since the intrinsic activity in LYSO is about 300 Bq/cm^3 (12 000 Bq/cm^3 per monolith), which amounts to $\sim$ 2.35 MBq per panel in case of 14 rows and 14 columns. This is roughly 20% of what we simulated with the 4 x 3 MBq ^{68}Ge rod sources per panel. We also note that the intrinsic gamma radiation can be absorbed by the emitting crystal itself, reducing the actually available activity. A Monte Carlo study [162] has shown that for LYSO crystals similar in size as ours (57.4 x 57.4 x 10 mm^3), the fraction of gamma photons actually escaping the crystal was 39.7% for 307 keV and 18.5% for 202 keV.

Finally, in Chapter 10 we looked into correcting the limited angle artifacts. We used a deep learning, image-based approach, where we incorporated the correction in a regularization term within the iterative reconstruction to better guarantee data fidelity. We also saw that this lead to better lesion contrast recovery than when applied as a post-processing step, as the neural network (a 2D U-Net) had a tendency to overly smooth the images. A big limitation of this study is that training

and testing was done on XCAT phantoms, which we suspect may be too uniform in nature to accurately represent a real patient scan. Including more realistic testing (and if required, training) data, such as simulations based on real PET patient scans, would therefore be a logical next step to better determine how well the method produces artifact-free images without potential mispredictions. Here, we again note that making the switch to LYSO for better TOF resolution will decrease the limited angle artifacts, which will in turn make the limited angle correcting algorithm more accurate and robust, as the relative differences between input and output are reduced. In certain cases, such as for dedicated brain scans, it will still be advisable to center the relevant organ (brain) in the FOV, not only to increase the sensitivity, but also to reduce the limited angle artifacts. This is especially true when we wish to optimize the spatial resolution to image finer structures, as we have seen in Chapter 7 that the spatial resolution degrades towards the edge of the AFOV.

11.4 Future perspectives

One aspect regarding the imaging process, which was not included in the simulation, but is unavoidable in practice, is patient motion. There can be both internal (cardiac and breathing) and external motion, which degrade (blur) the reconstructions and pose some practical limitation on the effective spatial resolution of the system. It is important to take this into account in the WT-PET to make effective use of the system's high spatial resolution, also considering the increased likelihood of motion for standing patients. Our group has already started developing a platform to provide patient support for the head, back and hands in order to limit motion as much as possible. Nonetheless, some degree of motion is expected to remain, including the internal breathing motion. Therefore, in the future, we would like to include motion correction within the reconstruction framework. Our current plan is to do this by reconstructing short (e.g., 1 s) frames, from which motion vector fields can be derived, and then use this motion data to offset individual LORs accordingly during the reconstruction, in order to reconstruct all events onto a common, fixed coordinate grid. Such patient motion can be included in the XCAT phantoms for GATE simulation, where we would again have access to ground-truth data for evaluation.

Given that the work in this dissertation has been based on simulations, and that the WT-PET scanner is currently under construction,

another important next step is to implement and validate the developed algorithms on the physical system. This starts from the detector calibration, where we need to do not only the time calibration as in Chapter 5, but also positioning calibration. Our group has previous (experimental) experience calibrating monolithic detectors for positioning, so that we can use previously developed approaches [49], [100]. Other groups have also shown that simultaneous estimation of the gamma interaction time and position can improve performance, as the optical photon detection delays in the crystal depend on the interaction position [101]. It would therefore be advantageous to combine these calibrations. This calibration procedure will ultimately need to be done for all detectors in the system, which will be a time-consuming procedure. Here, we can also look into methods to speed up this procedure, for example, by using a collimated line source, rather than a pencil beam source. Once at least a few rows of the system have been assembled and calibrated, we can start experimentally assessing the spatial resolution, image quality, accuracy of scatter correction, etc. We will also need to check here if our approximation of the system matrix (without normalization) still produces accurate images, since non-idealities such as detector non-uniformity, or the specific nature of dead-time within the PETsys electronics, could impact these results.

It will be exciting to see the continued development of the WT-PET scanner, and hopefully the developed methodologies and results obtained within this dissertation will be of use in the final system.

Bibliography

[1] M. Decuyper, J. Maebe, R. Van Holen, and S. Vandenberghe, "Artificial intelligence with deep learning in nuclear medicine and radiology", *EJNMMI Physics*, vol. 8, no. 1, p. 81, Dec. 2021, ISSN: 2197-7364. DOI: 10.1186/s40658-021-00426-y.

[2] J. Maebe and S. Vandenberghe, "Effect of detector geometry and surface finish on Cerenkov based time estimation in monolithic BGO detectors", *Physics in Medicine & Biology*, vol. 68, no. 2, p. 025 009, Jan. 2023, ISSN: 0031-9155. DOI: 10.1088/1361-6560/acabfd.

[3] J. Maebe and S. Vandenberghe, "Simulation study on 3D convolutional neural networks for time-of-flight prediction in monolithic PET detectors using digitized waveforms", *Physics in Medicine & Biology*, vol. 67, no. 12, p. 125 016, Jun. 2022, ISSN: 0031-9155. DOI: 10.1088/1361-6560/ac73d3.

[4] M. Dadgar, J. Maebe, M. Abi Akl, B. Vervenne, and S. Vandenberghe, "A simulation study of the system characteristics for a long axial FOV PET design based on monolithic BGO flat panels compared with a pixelated LSO cylindrical design", *EJNMMI Physics*, vol. 10, no. 1, p. 75, Dec. 2023, ISSN: 2197-7364. DOI: 10.1186/s40658-023-00593-0.

[5] S. Erhamamci, M. Reyhan, G. N. Nursal, N. Torun, and A. F. Yapar, "Incidental diagnosis of tumor thrombosis on FDG PET/CT imaging", *Revista Española de Medicina Nuclear e Imagen Molecular*, vol. 34, no. 5, pp. 287–294, Sep. 2015, ISSN: 2253-654X. DOI: 10.1016/j.remn.2015.04.002.

[6] T. Maurer, M. Eiber, M. Schwaiger, and J. E. Gschwend, "Current use of PSMA–PET in prostate cancer management", *Nature*

Reviews Urology, vol. 13, no. 4, pp. 226–235, Apr. 2016, ISSN: 1759-4820. DOI: 10.1038/nrurol.2016.26.

[7] C. Marcus, E. Mena, and R. M. Subramaniam, "Brain PET in the Diagnosis of Alzheimer's Disease", *Clinical nuclear medicine*, vol. 39, no. 10, e413–e426, Oct. 2014, ISSN: 0363-9762. DOI: 10.1097/RLU.0000000000000547.

[8] I. Sarikaya, "PET studies in epilepsy", *American Journal of Nuclear Medicine and Molecular Imaging*, vol. 5, no. 5, pp. 416–430, Oct. 2015, ISSN: 2160-8407.

[9] P. T. Meyer, L. Frings, G. Rücker, and S. Hellwig, "18F-FDG PET in Parkinsonism: Differential Diagnosis and Evaluation of Cognitive Impairment", *Journal of Nuclear Medicine*, vol. 58, no. 12, pp. 1888–1898, Dec. 2017, ISSN: 0161-5505, 2159-662X. DOI: 10.2967/jnumed.116.186403.

[10] A. Takalkar, A. Mavi, A. Alavi, and L. Araujo, "PET in cardiology", *Radiologic Clinics of North America*, PET Imaging II, vol. 43, no. 1, pp. 107–119, Jan. 2005, ISSN: 0033-8389. DOI: 10.1016/j.rcl.2004.09.007.

[11] L. M. Carter, A. Leon Kesner, E. C. Pratt, *et al.*, "The Impact of Positron Range on PET Resolution, Evaluated with Phantoms and PHITS Monte Carlo Simulations for Conventional and Non-conventional Radionuclides", *Molecular imaging and biology*, vol. 22, no. 1, pp. 73–84, Feb. 2020, ISSN: 1536-1632. DOI: 10.1007/s11307-019-01337-2.

[12] M. D. Harpen, "Positronium: Review of symmetry, conserved quantities and decay for the radiological physicist", *Medical Physics*, vol. 31, no. 1, pp. 57–61, 2004, ISSN: 2473-4209. DOI: 10.1118/1.1630494.

[13] P. Moskal, D. Kisielewska, C. Curceanu, *et al.*, "Feasibility study of the positronium imaging with the J-PET tomograph", *Physics in Medicine & Biology*, vol. 64, no. 5, p. 055 017, Mar. 2019, ISSN: 0031-9155. DOI: 10.1088/1361-6560/aafe20.

[14] P. Colombino, B. Fiscella, and L. Trossi, "Study of positronium in water and ice from 22 to -144 °C by annihilation quanta measurements", *Il Nuovo Cimento (1955-1965)*, vol. 38, no. 2, pp. 707–723, Jul. 1965, ISSN: 1827-6121. DOI: 10.1007/BF02748591.

[15] W. W. Moses, "Fundamental Limits of Spatial Resolution in PET", *Nuclear instruments & methods in physics research. Section A, Accelerators, spectrometers, detectors and associated equipment*, vol. 648 Supplement 1, S236–S240, Aug. 2011, ISSN: 0168-9002. DOI: 10.1016/j.nima.2010.11.092.

[16] M. Nikl, "Scintillation detectors for x-rays", *Measurement Science and Technology*, vol. 17, no. 4, R37, Feb. 2006, ISSN: 0957-0233. DOI: 10.1088/0957-0233/17/4/R01.

[17] S. Gundacker, E. Auffray, K. Pauwels, and P. Lecoq, "Measurement of intrinsic rise times for various L(Y)SO and LuAG scintillators with a general study of prompt photons to achieve 10 ps in TOF-PET", *Physics in Medicine & Biology*, vol. 61, no. 7, p. 2802, Mar. 2016, ISSN: 0031-9155. DOI: 10.1088/0031-9155/61/7/2802.

[18] "LYSO(Ce) Scintillator", Epic-Crystal, Tech. Rep.

[19] "BGO Scintillator", Epic-Crystal, Tech. Rep.

[20] F. Acerbi and S. Gundacker, "Understanding and simulating SiPMs", *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, Silicon Photomultipliers: Technology, Characterisation and Applications, vol. 926, pp. 16–35, May 2019, ISSN: 0168-9002. DOI: 10.1016/j.nima.2018.11.118.

[21] G. A. Prenosil, H. Sari, M. Fürstner, *et al.*, "Performance Characteristics of the Biograph Vision Quadra PET/CT System with a Long Axial Field of View Using the NEMA NU 2-2018 Standard", *Journal of Nuclear Medicine*, vol. 63, no. 3, pp. 476–484, Mar. 2022, ISSN: 0161-5505, 2159-662X. DOI: 10.2967/jnumed.121.261972.

[22] E. J. Hoffman, S.-C. Huang, M. E. Phelps, and D. E. Kuhl, "Quantitation in Positron Emission Computed Tomography: 4. Effect of Accidental Coincidences", *Journal of Computer Assisted Tomography*, vol. 5, no. 3, p. 391, Jun. 1981, ISSN: 0363-8715.

[23] C. C. Watson, D. Newport, and M. E. Casey, "A Single Scatter Simulation Technique for Scatter Correction in 3D PET", in *Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, ser. Computational Imaging and Vision, P. Grangeat and J.-L. Amans, Eds., Dordrecht: Springer Netherlands,

1996, pp. 255–268, ISBN: 978-94-015-8749-5. DOI: 10.1007/978-94-015-8749-5_18.

[24] *Performance Measurements of Positron Emission Tomographs (PET)*, https://www.nema.org/standards/view/Performance-Measurements-of-Positron-Emission-Tomographs, Jun. 2018.

[25] T. Tomitani, "Image Reconstruction and Noise Evaluation in Photon Time-of-Flight Assisted Positron Emission Tomography", *IEEE Transactions on Nuclear Science*, vol. 28, no. 6, pp. 4581–4589, Dec. 1981, ISSN: 1558-1578. DOI: 10.1109/TNS.1981.4335769.

[26] S. Jan, G. Santin, D. Strul, *et al.*, "GATE: A simulation toolkit for PET and SPECT", *Physics in Medicine and Biology*, vol. 49, no. 19, pp. 4543–4561, Oct. 2004, ISSN: 0031-9155. DOI: 10.1088/0031-9155/49/19/007.

[27] J. Allison, K. Amako, J. Apostolakis, *et al.*, "Recent developments in Geant4", *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 835, pp. 186–225, Nov. 2016, ISSN: 0168-9002. DOI: 10.1016/j.nima.2016.06.125.

[28] B. A. Spencer, E. Berg, J. P. Schmall, *et al.*, "Performance Evaluation of the uEXPLORER Total-Body PET/CT Scanner Based on NEMA NU 2-2018 with Additional Tests to Characterize PET Scanners with a Long Axial Field of View", *Journal of Nuclear Medicine*, vol. 62, no. 6, pp. 861–870, Jun. 2021, ISSN: 0161-5505, 2159-662X. DOI: 10.2967/jnumed.120.250597.

[29] B. Dai, M. E. Daube-Witherspoon, S. McDonald, *et al.*, "Performance evaluation of the PennPET explorer with expanded axial coverage", *Physics in Medicine & Biology*, vol. 68, no. 9, p. 095 007, Apr. 2023, ISSN: 0031-9155. DOI: 10.1088/1361-6560/acc722.

[30] S. Surti, A. R. Pantel, and J. S. Karp, "Total Body PET: Why, How, What for?", *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 4, no. 3, pp. 283–292, May 2020, ISSN: 2469-7303. DOI: 10.1109/TRPMS.2020.2985403.

[31] P. Moskal, P. Kowalski, R. Y. Shopa, *et al.*, "Simulating NEMA characteristics of the modular total-body J-PET scanner—an economic total-body PET from plastic scintillators", *Physics in*

Medicine & Biology, vol. 66, no. 17, p. 175 015, Aug. 2021, ISSN: 0031-9155. DOI: 10.1088/1361-6560/ac16bd.

[32] L. P. Kaelbling, M. L. Littman, and A. W. Moore, "Reinforcement Learning: A Survey", *Journal of Artificial Intelligence Research*, vol. 4, pp. 237–285, May 1996, ISSN: 1076-9757. DOI: 10.1613/jair.301.

[33] W. S. Mcculloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity. 1943.", *Bulletin of Mathematical Biology*, vol. 52, no. 2, pp. 99–115, 1990. DOI: 10.1007/BF02459570.

[34] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach* (Pearson Series in Artificial Intelligence), 4th ed. Pearson, 2020, ISBN: 978-0-13-461099-3.

[35] F. Rosenblatt, "The perceptron: A probabilistic model for information storage and organization in the brain.", *Psychological Review*, vol. 65, no. 6, 1958, ISSN: 1939-1471. DOI: 10.1037/h0042519.

[36] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks", *Proceedings of Machine Learning Research*, pp. 315–323, Jun. 2011.

[37] A. L. Maas, A. Y. Hannun, and A. Y. Ng, "Rectifier nonlinearities improve neural network acoustic models", *ICML Workshop on Deep Learning for Audio, Speech, and Language Processing*, vol. 30, 2013.

[38] Y. Lecun, "Generalization and network design strategies", *Connectionism in perspective*, vol. 19, pp. 143–155, 1989.

[39] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.

[40] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation", *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9351, pp. 234–241, 2015, ISSN: 16113349. DOI: 10.1007/978-3-319-24574-4_28.

[41] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3431–3440, 2015.

[42] M. Pizzichemi, G. Stringhini, T. Niknejad, *et al.*, "A new method for depth of interaction determination in PET detectors", *Physics in Medicine and Biology*, vol. 61, no. 12, pp. 4679–4698, Jun. 2016, ISSN: 1361-6560. DOI: 10.1088/0031-9155/61/12/4679.

[43] A. Zatcepin, M. Pizzichemi, A. Polesel, *et al.*, "Improving depth-of-interaction resolution in pixellated PET detectors using neural networks", *Physics in Medicine and Biology*, vol. 65, no. 17, p. 175 017, Aug. 2020, ISSN: 1361-6560. DOI: 10.1088/1361-6560/ab9efc.

[44] P. Bruyndonckx, S. Leonard, S. Tavernier, *et al.*, "Neural network-based position estimators for PET detectors using monolithic LSO blocks", *IEEE Transactions on Nuclear Science*, vol. 51, no. 5, pp. 2520–2525, Oct. 2004, ISSN: 1558-1578. DOI: 10.1109/TNS.2004.835782.

[45] S. Tavernier, P. Bruyndonckx, S. Leonard, and O. Devroede, "A high-resolution PET detector based on continuous scintillators", *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, Proceedings of the 7th International Conference on Inorganic Scintillators and Their Use in Scientific Adn Industrial Applications, vol. 537, no. 1, pp. 321–325, Jan. 2005, ISSN: 0168-9002. DOI: 10.1016/j.nima.2004.08.035.

[46] P. Bruyndonckx, C. Lemaitre, D. J. van der Laan, *et al.*, "Evaluation of Machine Learning Algorithms for Localization of Photons in Undivided Scintillator Blocks for PET Detectors", *IEEE Transactions on Nuclear Science*, vol. 55, no. 3, pp. 918–924, Jun. 2008, ISSN: 1558-1578. DOI: 10.1109/TNS.2008.922811.

[47] Y. Wang, W. Zhu, X. Cheng, and D. Li, "3D position estimation using an artificial neural network for a continuous scintillator PET detector", *Physics in Medicine & Biology*, vol. 58, no. 5, p. 1375, Feb. 2013, ISSN: 0031-9155. DOI: 10.1088/0031-9155/58/5/1375.

[48] A. Iborra, A. J. González, A. González-Montoro, A. Bousse, and D. Visvikis, "Ensemble of neural networks for 3D position estimation in monolithic PET detectors", *Physics in Medicine & Biology*, vol. 64, no. 19, p. 195 010, Oct. 2019, ISSN: 0031-9155. DOI: 10.1088/1361-6560/ab3b86.

[49] M. Decuyper, M. Stockhoff, S. Vandenberghe, and R. Van Holen, "Artificial neural networks for positioning of gamma interactions in monolithic PET detectors", *Physics in Medicine & Biology*, vol. 66, no. 7, p. 075 001, Mar. 2021, ISSN: 0031-9155. DOI: 10.1088/1361-6560/abebfc.

[50] W. He, Y. Wang, X. Liang, *et al.*, "High-performance coded aperture gamma camera based on monolithic GAGG:Ce crystal", *The Review of Scientific Instruments*, vol. 92, no. 1, p. 013 106, Jan. 2021, ISSN: 1089-7623. DOI: 10.1063/5.0035991.

[51] J. Michaud, M. Tétrault, J. Beaudoin, *et al.*, "Sensitivity Increase Through a Neural Network Method for LOR Recovery of ICS Triple Coincidences in High-Resolution Pixelated- Detectors PET Scanners", *IEEE Transactions on Nuclear Science*, vol. 62, no. 1, pp. 82–94, Feb. 2015, ISSN: 1558-1578. DOI: 10.1109/TNS.2014.2372788.

[52] M. Bergeron, J. Cadorette, J. Beaudoin, *et al.*, "Performance Evaluation of the LabPET APD-Based Digital PET Scanner", *IEEE Transactions on Nuclear Science*, vol. 56, no. 1, pp. 10–16, Feb. 2009, ISSN: 1558-1578. DOI: 10.1109/TNS.2008.2010257.

[53] E. Berg and S. R. Cherry, "Using convolutional neural networks to estimate time-of-flight from PET detector waveforms", *Physics in Medicine & Biology*, vol. 63, no. 2, 02LT01, Jan. 2018, ISSN: 0031-9155. DOI: 10.1088/1361-6560/aa9dc5.

[54] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-Image Translation with Conditional Adversarial Networks", *arXiv:1611.07004 [cs]*, Nov. 2018. arXiv: 1611.07004 [cs].

[55] X. Huang, M.-Y. Liu, S. Belongie, and J. Kautz, "Multimodal Unsupervised Image-to-Image Translation", *arXiv:1804.04732 [cs, stat]*, Aug. 2018. arXiv: 1804.04732 [cs, stat].

[56] M.-Y. Liu, T. Breuel, and J. Kautz, "Unsupervised Image-to-Image Translation Networks", *arXiv:1703.00848 [cs]*, Jul. 2018. arXiv: 1703.00848 [cs].

[57] I. Häggström, C. R. Schmidtlein, G. Campanella, and T. J. Fuchs, "DeepPET: A deep encoder–decoder network for directly solving the PET image reconstruction inverse problem", *Medical Image Analysis*, vol. 54, pp. 253–262, May 2019, ISSN: 1361-8415. DOI: 10.1016/j.media.2019.03.013.

[58] W. P. Segars, G. Sturgeon, S. Mendonca, J. Grimes, and B. M. W. Tsui, "4D XCAT phantom for multimodality imaging research", *Medical Physics*, vol. 37, no. 9, pp. 4902–4915, 2010, ISSN: 2473-4209. DOI: 10.1118/1.3480985.

[59] Z. Hu, H. Xue, Q. Zhang, *et al.*, "DPIR-Net: Direct PET Image Reconstruction Based on the Wasserstein Generative Adversarial Network", *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 5, no. 1, pp. 35–43, Jan. 2021, ISSN: 2469-7303. DOI: 10.1109/TRPMS.2020.2995717.

[60] T. Würfl, M. Hoffmann, V. Christlein, *et al.*, "Deep Learning Computed Tomography: Learning Projection-Domain Weights From Image Domain in Limited Angle Problems", *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1454–1463, Jun. 2018, ISSN: 1558-254X. DOI: 10.1109/TMI.2018.2833499.

[61] K. Gregor and Y. LeCun, "Learning fast approximations of sparse coding", in *Proceedings of the 27th International Conference on International Conference on Machine Learning*, ser. ICML'10, Madison, WI, USA: Omnipress, Jun. 2010, pp. 399–406, ISBN: 978-1-60558-907-7.

[62] Y. Yang, J. Sun, H. Li, and Z. Xu, "Deep ADMM-Net for Compressive Sensing MRI", *Advances in Neural Information Processing Systems*, vol. 29, 2016.

[63] H. K. Aggarwal, M. P. Mani, and M. Jacob, "MoDL: Model-Based Deep Learning Architecture for Inverse Problems", *IEEE Transactions on Medical Imaging*, vol. 38, no. 2, pp. 394–405, Feb. 2019, ISSN: 1558-254X. DOI: 10.1109/TMI.2018.2865356.

[64] D. Wu, K. Kim, and Q. Li, "Computationally efficient deep neural network for computed tomography image reconstruction", *Medical Physics*, vol. 46, no. 11, pp. 4763–4776, 2019, ISSN: 2473-4209. DOI: 10.1002/mp.13627.

[65] H. Zhang, B. Liu, H. Yu, and B. Dong, "MetaInv-Net: Meta Inversion Network for Sparse View CT Image Reconstruction", *IEEE Transactions on Medical Imaging*, vol. 40, no. 2, pp. 621–634, Feb. 2021, ISSN: 1558-254X. DOI: 10.1109/TMI.2020.3033541.

[66] T. Eo, Y. Jun, T. Kim, J. Jang, H.-J. Lee, and D. Hwang, "KIKI-net: Cross-domain convolutional neural networks for reconstructing undersampled magnetic resonance images", *Magnetic Resonance in Medicine*, vol. 80, no. 5, pp. 2188–2201, 2018, ISSN: 1522-2594. DOI: 10.1002/mrm.27201.

[67] K. Hammernik, T. Klatzer, E. Kobler, *et al.*, "Learning a variational network for reconstruction of accelerated MRI data", *Magnetic Resonance in Medicine*, vol. 79, no. 6, pp. 3055–3071, 2018, ISSN: 1522-2594. DOI: 10.1002/mrm.26977.

[68] J. Schlemper, J. Caballero, J. V. Hajnal, A. N. Price, and D. Rueckert, "A Deep Cascade of Convolutional Neural Networks for Dynamic MR Image Reconstruction", *IEEE Transactions on Medical Imaging*, vol. 37, no. 2, pp. 491–503, Feb. 2018, ISSN: 1558-254X. DOI: 10.1109/TMI.2017.2760978.

[69] J. Zhang and B. Ghanem, "ISTA-Net: Interpretable Optimization-Inspired Deep Network for Image Compressive Sensing", in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Jun. 2018, pp. 1828–1837. DOI: 10.1109/CVPR.2018.00196.

[70] J. Cheng, H. Wang, L. Ying, and D. Liang, "Model Learning: Primal Dual Networks for Fast MR Imaging", in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*, D. Shen, T. Liu, T. M. Peters, *et al.*, Eds., ser. Lecture Notes in Computer Science, Cham: Springer International Publishing, 2019, pp. 21–29, ISBN: 978-3-030-32248-9. DOI: 10.1007/978-3-030-32248-9_3.

[71] C. Qin, J. Schlemper, J. Caballero, A. N. Price, J. V. Hajnal, and D. Rueckert, "Convolutional Recurrent Neural Networks for Dynamic MR Image Reconstruction", *IEEE Transactions on Medical Imaging*, vol. 38, no. 1, pp. 280–290, Jan. 2019, ISSN: 1558-254X. DOI: 10.1109/TMI.2018.2863670.

[72] K. Gong, D. Wu, K. Kim, *et al.*, "EMnet: An unrolled deep neural network for PET image reconstruction", in *Medical Imaging 2019: Physics of Medical Imaging*, vol. 10948, International Society for Optics and Photonics, Mar. 2019, p. 1 094 853. DOI: 10.1117/12.2513096.

[73] K. Gong, D. Wu, K. Kim, *et al.*, "MAPEM-Net: An unrolled neural network for Fully 3D PET image reconstruction", in *15th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, vol. 11072, International Society for Optics and Photonics, May 2019, 110720O. DOI: 10.1117/12.2534904.

[74] D. Liang, J. Cheng, Z. Ke, and L. Ying, "Deep MRI Reconstruction: Unrolled Optimization Algorithms Meet Neural Networks", *arXiv:1907.11711 [physics, stat]*, Jul. 2019. arXiv: 1907.11711 [physics, stat].

[75] V. Monga, Y. Li, and Y. C. Eldar, "Algorithm Unrolling: Interpretable, Efficient Deep Learning for Signal and Image Processing", *arXiv:1912.10557 [cs, eess]*, Aug. 2020. arXiv: 1912.10557 [cs, eess].

[76] A. Reader, G. Corda-D'Incan, A. Mehranian, C. da Costa-Luis, S. Ellis, and J. Schnabel, "Deep Learning for PET Image Reconstruction", *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. PP, pp. 1–1, Aug. 2020. DOI: 10.1109/TRPMS.2020.3014786.

[77] H. Zhang, L. Li, K. Qiao, *et al.*, "Image Prediction for Limited-angle Tomography via Deep Learning with Convolutional Neural Network", *arXiv:1607.08707 [physics]*, Jul. 2016. arXiv: 1607.08707 [physics].

[78] H. Chen, Y. Zhang, W. Zhang, *et al.*, "Low-dose CT via convolutional neural network", *Biomedical Optics Express*, vol. 8, no. 2, pp. 679–694, Jan. 2017, ISSN: 2156-7085. DOI: 10.1364/BOE.8.000679.

[79] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising", *IEEE Transactions on Image Processing*, vol. 26, no. 7, pp. 3142–3155, Jul. 2017, ISSN: 1941-0042. DOI: 10.1109/TIP.2017.2662206.

[80] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, "Deep Convolutional Neural Network for Inverse Problems in Imaging", *IEEE Transactions on Image Processing*, vol. 26, no. 9, pp. 4509–4522, Sep. 2017, ISSN: 1941-0042. DOI: 10.1109/TIP.2017.2713099.

[81] H. Chen, Y. Zhang, M. K. Kalra, *et al.*, "Low-Dose CT With a Residual Encoder-Decoder Convolutional Neural Network", *IEEE Transactions on Medical Imaging*, vol. 36, no. 12, pp. 2524–2535, Dec. 2017, ISSN: 1558-254X. DOI: 10.1109/TMI.2017.2715284.

[82] J. Xu, E. Gong, J. Pauly, and G. Zaharchuk, "200x Low-dose PET Reconstruction using Deep Learning", *arXiv:1712.04119 [cs]*, Dec. 2017. arXiv: 1712.04119 [cs].

[83] S. Kaplan and Y.-M. Zhu, "Full-Dose PET Image Estimation from Low-Dose PET Image Using Deep Learning: A Pilot Study", *Journal of Digital Imaging*, vol. 32, no. 5, pp. 773–778, Oct. 2019, ISSN: 1618-727X. DOI: 10.1007/s10278-018-0150-3.

[84] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition", in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2016, pp. 770–778. DOI: 10.1109/CVPR.2016.90.

[85] K. Gong, J. Guan, C.-C. Liu, and J. Qi, "PET Image Denoising Using a Deep Neural Network Through Fine Tuning", *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 3, no. 2, pp. 153–161, Mar. 2019, ISSN: 2469-7303. DOI: 10.1109/TRPMS.2018.2877644.

[86] J. Dong, J. Fu, and Z. He, "A deep learning reconstruction framework for X-ray computed tomography with incomplete data", *PLOS ONE*, vol. 14, no. 11, e0224426, Nov. 2019, ISSN: 1932-6203. DOI: 10.1371/journal.pone.0224426.

[87] I. Shiri, P. Sheikhzadeh, and M. R. Ay, "Deep-Fill: Deep Learning Based Sinogram Domain Gap Filling in Positron Emission Tomography", *arXiv:1906.07168 [physics]*, Jun. 2019. arXiv: 1906.07168 [physics].

[88] C.-C. Liu and H.-M. Huang, "Partial-ring PET image restoration using a deep learning based method", *Physics in Medicine & Biology*, vol. 64, no. 22, p. 225 014, Nov. 2019, ISSN: 0031-9155. DOI: 10.1088/1361-6560/ab4aa9.

[89] L. Xiang, Y. Qiao, D. Nie, *et al.*, "Deep auto-context convolutional neural networks for standard-dose PET image estimation from low-dose PET/MRI", *Neurocomputing*, vol. 267, pp. 406–416, Dec. 2017, ISSN: 0925-2312. DOI: 10.1016/j.neucom.2017.06.048.

[90] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Deep Image Prior", *arXiv:1711.10925 [cs, stat]*, May 2020. DOI: 10.1007/s11263-020-01303-4. arXiv: 1711.10925 [cs, stat].

[91] J. Cui, K. Gong, N. Guo, *et al.*, "PET image denoising using unsupervised deep learning", *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 46, Dec. 2019. DOI: 10.1007/s00259-019-04468-4.

[92] F. Hashimoto, H. Ohba, K. Ote, A. Teramoto, and H. Tsukada, "Dynamic PET Image Denoising Using Deep Convolutional Neural Networks Without Prior Training Datasets", *IEEE Access*, vol. 7, pp. 96 594–96 603, 2019, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2019.2929230.

[93] K. Gong, J. Guan, K. Kim, *et al.*, "Iterative PET Image Reconstruction Using Convolutional Neural Network Representation", *IEEE Transactions on Medical Imaging*, vol. 38, no. 3, pp. 675–685, Mar. 2019, ISSN: 1558-254X. DOI: 10.1109/TMI.2018.2869871.

[94] D. O. Baguer, J. Leuschner, and M. Schmidt, "Computed tomography reconstruction using deep image prior and learned reconstruction methods", *Inverse Problems*, vol. 36, no. 9, p. 094 004, Sep. 2020, ISSN: 0266-5611. DOI: 10.1088/1361-6420/aba415.

[95] S. E. Brunner and D. R. Schaart, "BGO as a hybrid scintillator / Cherenkov radiator for cost-effective time-of-flight PET", *Physics in Medicine & Biology*, vol. 62, no. 11, p. 4421, May 2017, ISSN: 0031-9155. DOI: 10.1088/1361-6560/aa6a49.

[96] S. Gundacker, R. M. Turtos, N. Kratochwil, *et al.*, "Experimental time resolution limits of modern SiPMs and TOF-PET detectors exploring different scintillators and Cherenkov emission", *Physics in Medicine &amp$\mathsemicolon$ Biology*, vol. 65, no. 2, p. 025 001, Jan. 2020, ISSN: 0031-9155. DOI: 10.1088/1361-6560/ab63b4.

[97] N. Efthimiou, N. Kratochwil, S. Gundacker, *et al.*, "TOF-PET Image Reconstruction With Multiple Timing Kernels Applied on Cherenkov Radiation in BGO", *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 5, no. 5, pp. 703–711, Sep. 2021, ISSN: 2469-7303. DOI: 10.1109/TRPMS.2020.3048642.

[98] N. Kratochwil, S. Gundacker, P. Lecoq, and E. Auffray, "Pushing Cherenkov PET with BGO via coincidence time resolution classification and correction", *Physics in Medicine & Biology*, vol. 65, no. 11, p. 115 004, Jun. 2020, ISSN: 0031-9155. DOI: 10.1088/1361-6560/ab87f9.

[99] A. Gonzalez-Montoro, A. J. Gonzalez, S. Pourashraf, *et al.*, "Evolution of PET Detectors and Event Positioning Algorithms Using Monolithic Scintillation Crystals", *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 5, no. 3, pp. 282–305, May 2021, ISSN: 2469-7303. DOI: 10.1109/TRPMS.2021.3059181.

[100] M. Stockhoff, M. Decuyper, R. Van Holen, and S. Vandenberghe, "High-resolution monolithic LYSO detector with 6-layer depth-of-interaction for clinical PET", *Physics in Medicine & Biology*, vol. 66, no. 15, p. 155 014, Jul. 2021, ISSN: 0031-9155. DOI: 10.1088/1361-6560/ac1459.

[101] P. Carra, M. G. Bisogni, E. Ciarrocchi, *et al.*, "A neural network-based algorithm for simultaneous event positioning and timestamping in monolithic scintillators", *Physics in Medicine & Biology*, vol. 67, no. 13, p. 135 001, Jun. 2022, ISSN: 0031-9155. DOI: 10.1088/1361-6560/ac72f2.

[102] E. Berg, E. Roncali, and S. R. Cherry, "Optimizing light transport in scintillation crystals for time-of-flight PET: An experimental and optical Monte Carlo simulation study", *Biomedical Optics Express*, vol. 6, no. 6, pp. 2220–2230, Jun. 2015, ISSN: 2156-7085. DOI: 10.1364/BOE.6.002220.

[103] D. Sarrut, M. Bała, M. Bardiès, *et al.*, "Advanced Monte Carlo simulations of emission tomography imaging systems with GATE", *Physics in Medicine & Biology*, vol. 66, no. 10, 10TR03, May 2021, ISSN: 0031-9155. DOI: 10.1088/1361-6560/abf276.

[104] S. Krishnamoorthy, E. Blankemeyer, P. Mollet, S. Surti, R. Van Holen, and J. S. Karp, "Performance evaluation of the MOLECUBES β-CUBE—a high spatial resolution and high sensitivity small animal PET scanner utilizing monolithic LYSO scintillation detectors", *Physics in Medicine & Biology*, vol. 63, no. 15, p. 155 013, Jul. 2018, ISSN: 0031-9155. DOI: 10.1088/1361-6560/aacec3.

[105] E. Roncali and S. R. Cherry, "Simulation of light transport in scintillators based on 3D characterization of crystal surfaces", *Physics in Medicine & Biology*, vol. 58, no. 7, p. 2185, Mar. 2013, ISSN: 0031-9155. DOI: 10.1088/0031-9155/58/7/2185.

[106] C. Trigila, G. Ariño-Estrada, S. I. Kwon, and E. Roncali, "The Accuracy of Cerenkov Photons Simulation in Geant4/Gate Depends on the Parameterization of Primary Electron Propagation", *Frontiers in Physics*, vol. 10, 2022, ISSN: 2296-424X.

[107] H. Shinotsuka, S. Tanuma, C. J. Powell, and D. R. Penn, "Calculations of electron inelastic mean free paths. X. Data for 41 elemental solids over the 50 eV to 200 keV range with the relativistic full Penn algorithm", *Surface and Interface Analysis*, vol. 47, no. 9, pp. 871–888, 2015, ISSN: 1096-9918. DOI: 10.1002/sia.5789.

[108] E. Dietz-Laursonn, *Peculiarities in the Simulation of Optical Physics with Geant4*, Dec. 2016. DOI: 10.48550/arXiv.1612.05162. arXiv: 1612.05162 [physics].

[109] C. Trigila, E. Moghe, and E. Roncali, "Technical Note: Standalone application to generate custom reflectance Look-Up Table for advanced optical Monte Carlo simulation in GATE/Geant4", *Medical Physics*, vol. 48, no. 6, pp. 2800–2808, 2021, ISSN: 2473-4209. DOI: 10.1002/mp.14863.

[110] "AFBR-S4N66P024M 2×1 NUV-MT Silicon Photomultiplier Array", Broadcom, Tech. Rep., 2022.

[111] D. Sánchez, S. Gómez, J. Mauricio, *et al.*, "HRFlexToT: A High Dynamic Range ASIC for Time-of-Flight Positron Emission Tomography", *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 6, no. 1, pp. 51–67, Jan. 2022, ISSN: 2469-7303. DOI: 10.1109/TRPMS.2021.3066426.

[112] P. Jarron, E. Auffray, S. Brunner, *et al.*, "Time based readout of a silicon photomultiplier (SiPM) for time of flight positron emission tomography (TOF-PET)", in *2009 IEEE Nuclear Science Symposium Conference Record (NSS/MIC)*, Oct. 2009, pp. 1212–1219. DOI: 10.1109/NSSMIC.2009.5402391.

[113] M. Decuyper, "Advancing medical imaging with artificial intelligence : PET acquisition enhancement and MRI-based brain tumour diagnosis", Ph.D. dissertation, Ghent University, 2021, ISBN: 9789463555364.

[114] V. Nadig, K. Herweg, M. M. C. Chou, *et al.*, "Timing advances of commercial divalent-ion co-doped LYSO:Ce and SiPMs in sub-100 ps time-of-flight positron emission tomography", *Physics in Medicine & Biology*, vol. 68, no. 7, p. 075 002, Mar. 2023, ISSN: 0031-9155. DOI: 10.1088/1361-6560/acbde4.

[115] T. Saleem, S. Ahmad, J.-B. Cizel, *et al.*, "Study experimental time resolution limits of recent ASICs at Weeroc with different SiPMs and scintillators", *Journal of Instrumentation*, vol. 18, no. 10, P10005, Oct. 2023, ISSN: 1748-0221. DOI: 10.1088/1748-0221/18/10/P10005.

[116] K. Doroud, M. C. S. Williams, A. Zichichi, and R. Zuyeuski, "Comparative timing measurements of LYSO and LFS-3 to achieve the best time resolution for TOF-PET", *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 793, pp. 57–61, Sep. 2015, ISSN: 0168-9002. DOI: 10.1016/j.nima.2015.04.056.

[117] R. Bugalho, A. D. Francesco, L. Ferramacho, *et al.*, "Experimental characterization of the TOFPET2 ASIC", *Journal of Instrumentation*, vol. 14, no. 03, P03029, Mar. 2019, ISSN: 1748-0221. DOI: 10.1088/1748-0221/14/03/P03029.

[118] S. Surti, "Update on Time-of-Flight PET Imaging", *Journal of Nuclear Medicine*, vol. 56, no. 1, pp. 98–105, Jan. 2015, ISSN: 0161-5505, 2159-662X. DOI: 10.2967/jnumed.114.145029.

[119] J. van Sluis, J. de Jong, J. Schaar, *et al.*, "Performance Characteristics of the Digital Biograph Vision PET/CT System", *Journal of Nuclear Medicine*, vol. 60, no. 7, pp. 1031–1036, Jul. 2019, ISSN: 0161-5505, 2159-662X. DOI: 10.2967/jnumed.118.215418.

[120] X. Li, W. Qi, M. Miyahara, and J. Kolthammer, "Performance Characterization of an SiPM-based Time-of-Flight Canon PET/CT Scanner", *Journal of Nuclear Medicine*, vol. 61, no. supplement 1, pp. 14–14, May 2020, ISSN: 0161-5505, 2159-662X.

[121] G. A. Prenosil, H. Sari, M. Fürstner, *et al.*, "Performance Characteristics of the Biograph Vision Quadra PET/CT system with long axial field of view using the NEMA NU 2-2018 Standard", *Journal of Nuclear Medicine: Official Publication, Society of Nuclear Medicine*, jnumed.121.261972, Jul. 2021, ISSN: 1535-5667. DOI: 10.2967/jnumed.121.261972.

[122] J. W. Cates and C. S. Levin, "Advances in coincidence time resolution for PET", *Physics in Medicine and Biology*, vol. 61, no. 6, pp. 2255–2264, Feb. 2016, ISSN: 0031-9155. DOI: 10.1088/0031-9155/61/6/2255.

[123] E. Lamprou, A. J. Gonzalez, F. Sanchez, and J. M. Benlloch, "Exploring TOF capabilities of PET detector blocks based on large monolithic crystals and analog SiPMs", *Physica Medica*, vol. 70, pp. 10–18, Feb. 2020, ISSN: 1120-1797. DOI: 10.1016/j.ejmp.2019.12.004.

[124] Y. Onishi, F. Hashimoto, K. Ote, and R. Ota, "Unbiased TOF estimation using leading-edge discriminator and convolutional neural network trained by single-source-position waveforms", *Physics in Medicine & Biology*, vol. 67, no. 4, 04NT01, Feb. 2022, ISSN: 0031-9155. DOI: 10.1088/1361-6560/ac508f.

[125] E. Roncali, M. Stockhoff, and S. R. Cherry, "An integrated model of scintillator-reflector properties for advanced simulations of optical transport", *Physics in Medicine & Biology*, vol. 62, no. 12, p. 4811, May 2017, ISSN: 0031-9155. DOI: 10.1088/1361-6560/aa6ca5.

[126] M. Stockhoff, S. Jan, A. Dubois, S. R. Cherry, and E. Roncali, "Advanced optical simulation of scintillation detectors in GATE V8.0: First implementation of a reflectance model based on measured data", *Physics in Medicine and Biology*, vol. 62, no. 12, pp. L1–L8, May 2017, ISSN: 0031-9155. DOI: 10.1088/1361-6560/aa7007.

[127] "Silicon Photomultipliers (SiPM), High PDE and Timing Resolution Sensors in a TSV Package", onsemi, Tech. Rep., Aug. 2021.

[128] M. Stockhoff, R. Van Holen, and S. Vandenberghe, "Optical simulation study on the spatial resolution of a thick monolithic PET detector", *Physics in Medicine & Biology*, vol. 64, no. 19, p. 195 003, Sep. 2019, ISSN: 0031-9155. DOI: 10.1088/1361-6560/ab3b83.

[129] H. T. van Dam, G. Borghi, S. Seifert, and D. R. Schaart, "Sub-200 ps CRT in monolithic scintillator PET detectors using digital SiPM arrays and maximum likelihood interaction time estimation", *Physics in Medicine and Biology*, vol. 58, no. 10, pp. 3243–3257, Apr. 2013, ISSN: 0031-9155. DOI: 10.1088/0031-9155/58/10/3243.

[130] C. Wang and J. Principe, "Training neural networks with additive noise in the desired signal", *IEEE Transactions on Neural Networks*, vol. 10, no. 6, pp. 1511–1517, Nov. 1999, ISSN: 1941-0093. DOI: 10.1109/72.809097.

[131] S. Vandenberghe, F. M. Muller, N. Withofs, *et al.*, "Walk-through flat panel total-body PET: A patient-centered design for high throughput imaging at lower cost using DOI-capable high-resolution monolithic detectors", *European Journal of Nuclear Medicine and Molecular Imaging*, Jul. 2023, ISSN: 1619-7089. DOI: 10.1007/s00259-023-06341-x.

[132] P. Carra, *Performance of monolithic BGO-based detector implementing a Neural-Network event decoding algorithm for TB-PET applications*, 2022.

[133] S. A. Zein, N. A. Karakatsanis, M. Conti, and S. A. Nehmeh, "Monte Carlo Simulation of the Siemens Biograph Vision PET With Extended Axial Field of View Using Sparse Detector Module Rings Configuration", *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 5, no. 3, pp. 331–342, May 2021, ISSN: 2469-7303. DOI: 10.1109/TRPMS.2020.3034676.

[134] T. Merlin, S. Stute, D. Benoit, *et al.*, "CASToR: A generic data organization and processing code framework for multi-modal and multi-dimensional tomographic reconstruction", *Physics in Medicine & Biology*, vol. 63, no. 18, p. 185 005, Sep. 2018, ISSN: 0031-9155. DOI: 10.1088/1361-6560/aadac1.

[135] K. Thielemans, C. Tsoumpas, S. Mustafovic, *et al.*, "STIR: Software for tomographic image reconstruction release 2", *Physics in Medicine & Biology*, vol. 57, no. 4, p. 867, Jan. 2012, ISSN: 0031-9155. DOI: 10.1088/0031-9155/57/4/867.

[136] F. Enríquez-Mier-y-Terán, G. Angelis, S. Meikle, and A. Kyme, "Implementation of a Virtual Ring for DOI-Based PET Reconstruction for Open-Field Mouse Brain PET", in *2022 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*, Nov. 2022, pp. 1–2. DOI: 10.1109/NSS/MIC44845.2022.10399265.

[137] M. Zhang, J. Zhou, Y. Yang, M. Rodríguez-Villafuerte, and J. Qi, "Efficient system modeling for a small animal PET scanner with tapered DOI detectors", *Physics in Medicine & Biology*, vol. 61,

no. 2, p. 461, Dec. 2015, ISSN: 0031-9155. DOI: 10.1088/0031-9155/61/2/461.

[138] R. Huesman, G. Klein, W. Moses, J. Qi, B. Reutter, and P. Virador, "List-mode maximum-likelihood reconstruction applied to positron emission mammography (PEM) with irregular sampling", *IEEE Transactions on Medical Imaging*, vol. 19, no. 5, pp. 532–537, May 2000, ISSN: 1558-254X. DOI: 10.1109/42.870263.

[139] J. C. Hart, "Sphere tracing: A geometric method for the antialiased ray tracing of implicit surfaces", *The Visual Computer*, vol. 12, no. 10, pp. 527–545, Dec. 1996, ISSN: 1432-2315. DOI: 10.1007/s003710050084.

[140] L. M. Popescu, R. M. Lewitt, S. Matej, and J. S. Karp, "PET energy-based scatter estimation and image reconstruction with energy-dependent corrections", *Physics in Medicine & Biology*, vol. 51, no. 11, p. 2919, May 2006, ISSN: 0031-9155. DOI: 10.1088/0031-9155/51/11/016.

[141] N. Efthimiou, J. S. Karp, and S. Surti, "Data-driven, energy-based method for estimation of scattered events in positron emission tomography", *Physics in Medicine & Biology*, vol. 67, no. 9, p. 095 010, Apr. 2022, ISSN: 0031-9155. DOI: 10.1088/1361-6560/ac62fc.

[142] X. Dong, T. Wang, Y. Lei, *et al.*, "Synthetic CT generation from non-attenuation corrected PET images for whole-body PET imaging", *Physics in Medicine & Biology*, vol. 64, no. 21, p. 215 016, Nov. 2019, ISSN: 0031-9155. DOI: 10.1088/1361-6560/ab4eb7.

[143] J. Li, Z. Qu, Y. Yang, F. Zhang, M. Li, and S. Hu, "TCGAN: A transformer-enhanced GAN for PET synthetic CT", *Biomedical Optics Express*, vol. 13, no. 11, pp. 6003–6018, Nov. 2022, ISSN: 2156-7085. DOI: 10.1364/BOE.467683.

[144] M. Boulanger, J.-C. Nunes, H. Chourak, *et al.*, "Deep learning methods to generate synthetic CT from MRI in radiotherapy: A literature review", *Physica Medica*, vol. 89, pp. 265–281, Sep. 2021, ISSN: 1120-1797. DOI: 10.1016/j.ejmp.2021.07.027.

[145] X. Dong, Y. Lei, T. Wang, *et al.*, "Deep learning-based attenuation correction in the absence of structural information for whole-body positron emission tomography imaging", *Physics in Medicine and*

Biology, vol. 65, no. 5, p. 055 011, Mar. 2020, ISSN: 1361-6560. DOI: 10.1088/1361-6560/ab652c.

[146] G. Wagenknecht, H.-J. Kaiser, F. M. Mottaghy, and H. Herzog, "MRI for attenuation correction in PET: Methods and challenges", *Magnetic Resonance Materials in Physics, Biology and Medicine*, vol. 26, no. 1, pp. 99–113, Feb. 2013, ISSN: 1352-8661. DOI: 10.1007/s10334-012-0353-4.

[147] W. Moses, "Time of flight in PET revisited", *IEEE Transactions on Nuclear Science*, vol. 50, no. 5, pp. 1325–1330, Oct. 2003, ISSN: 1558-1578. DOI: 10.1109/TNS.2003.817319.

[148] K. Lange and R. Carson, "EM reconstruction algorithms for emission and transmission tomography", *Journal of Computer Assisted Tomography*, vol. 8, no. 2, pp. 306–316, Apr. 1984, ISSN: 0363-8715.

[149] J. Nuyts, B. De Man, P. Dupont, M. Defrise, P. Suetens, and L. Mortelmans, "Iterative reconstruction for helical CT: A simulation study", *Physics in Medicine & Biology*, vol. 43, no. 4, p. 729, Apr. 1998, ISSN: 0031-9155. DOI: 10.1088/0031-9155/43/4/003.

[150] Y. Censor, D. E. Gustafson, A. Lent, and H. Tuy, "A New Approach to the Emission Computerized Tomography Problem: Simultaneous Calculation of Attenuation and Activity Coefficients", *IEEE Transactions on Nuclear Science*, vol. 26, no. 2, pp. 2775–2779, Apr. 1979, ISSN: 1558-1578. DOI: 10.1109/TNS.1979.4330535.

[151] M. Defrise, A. Rezaei, and J. Nuyts, "Time-of-flight PET data determine the attenuation sinogram up to a constant", *Physics in Medicine & Biology*, vol. 57, no. 4, p. 885, Jan. 2012, ISSN: 0031-9155. DOI: 10.1088/0031-9155/57/4/885.

[152] A. Rezaei, M. Defrise, G. Bal, *et al.*, "Simultaneous Reconstruction of Activity and Attenuation in Time-of-Flight PET", *IEEE Transactions on Medical Imaging*, vol. 31, no. 12, pp. 2224–2233, Dec. 2012, ISSN: 1558-254X. DOI: 10.1109/TMI.2012.2212719.

[153] M. Kalke and S. Siltanen, "Sinogram Interpolation Method for Sparse-Angle Tomography", *Applied Mathematics*, vol. 5, no. 3, pp. 423–441, Feb. 2014. DOI: 10.4236/am.2014.53043.

[154] J. Karp, G. Muehllehner, and R. Lewitt, "Constrained Fourier space method for compensation of missing data in emission computed tomography", *IEEE Transactions on Medical Imaging*, vol. 7, no. 1, pp. 21–25, Mar. 1988, ISSN: 1558-254X. DOI: 10.1109/42.3925.

[155] J. Wang, J. Liang, J. Cheng, Y. Guo, and L. Zeng, "Deep learning based image reconstruction algorithm for limited-angle translational computed tomography", *PLOS ONE*, vol. 15, no. 1, e0226963, Jan. 2020, ISSN: 1932-6203. DOI: 10.1371/journal.pone.0226963.

[156] S. Barutcu, S. Aslan, A. K. Katsaggelos, and D. Gürsoy, "Limited-angle computed tomography with deep image and physics priors", *Scientific Reports*, vol. 11, no. 1, p. 17 740, Sep. 2021, ISSN: 2045-2322. DOI: 10.1038/s41598-021-97226-2.

[157] P. Henzler, V. Rasche, T. Ropinski, and T. Ritschel, "Single-image Tomography: 3D Volumes from 2D Cranial X-Rays", *Computer Graphics Forum*, vol. 37, no. 2, pp. 377–388, 2018, ISSN: 1467-8659. DOI: 10.1111/cgf.13369.

[158] X. Ying, H. Guo, K. Ma, J. Wu, Z. Weng, and Y. Zheng, *X2CT-GAN: Reconstructing CT from Biplanar X-Rays with Generative Adversarial Networks*, May 2019. DOI: 10.48550/arXiv.1905.06902. arXiv: 1905.06902 [cs, eess].

[159] E. Levitan and G. T. Herman, "A maximum a posteriori probability expectation maximization algorithm for image reconstruction in emission tomography", *IEEE transactions on medical imaging*, vol. 6, no. 3, pp. 185–192, 1987, ISSN: 0278-0062. DOI: 10.1109/TMI.1987.4307826.

[160] A. J. Reader, G. Corda, A. Mehranian, C. da Costa-Luis, S. Ellis, and J. A. Schnabel, "Deep Learning for PET Image Reconstruction", *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 5, no. 1, pp. 1–25, Jan. 2021, ISSN: 2469-7303. DOI: 10.1109/TRPMS.2020.3014786.

[161] H. Rothfuss, V. Panin, A. Moor, *et al.*, "LSO background radiation as a transmission source using time of flight", *Physics in Medicine & Biology*, vol. 59, no. 18, p. 5483, Aug. 2014, ISSN: 0031-9155. DOI: 10.1088/0031-9155/59/18/5483.

[162] H. Alva-Sánchez, A. Zepeda-Barrios, V. D. Díaz-Martínez, T. Murrieta-Rodríguez, A. Martínez-Dávalos, and M. Rodríguez-Villafuerte, "Understanding the intrinsic radioactivity energy spectrum from 176Lu in LYSO/LSO scintillation crystals", *Scientific Reports*, vol. 8, no. 1, p. 17 310, Nov. 2018, ISSN: 2045-2322. DOI: 10.1038/s41598-018-35684-x.