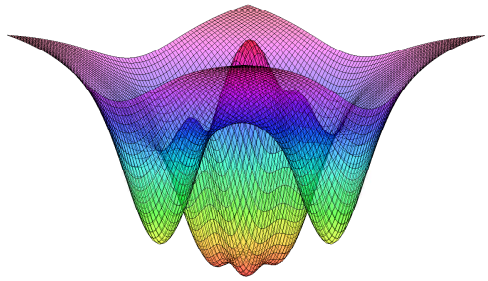
Vrije Universiteit Brussel

FACULTY OF SCIENCE
Department of Physics

Study of 3D position determination of the interaction point in monolithic scintillator blocks for PET



Thesis submitted in the fulfillment of the requirements for the award of the degree of Doctor in de wetenschappen (Doctor in Science) by

Zhi Li

Brussels 2011

Vrije Universiteit Brussel

FACULTY OF SCIENCE
Department of Physics

Study of 3D position determination of the interaction point in monolithic scintillator blocks for PET

Thesis submitted in the fulfillment of the requirements for the award of the degree of Doctor in de wetenschappen (Doctor in Science) by

Zhi Li

Promotors: Prof. dr. Stefaan Tavernier (Vrije Universiteit Brussel)
Prof. dr. ir. Gerd Vandersteen (Vrije Universiteit Brussel)

Vrije Universiteit Brussel

FACULTY OF SCIENCE
Department of Physics

Study of 3D position determination of the interaction point in monolithic scintillator blocks for PET

Thesis submitted in the fulfillment of the requirements for the award of the degree of Doctor in de wetenschappen (Doctor in Science) by

Zhi Li

Doctoral examination commission
Prof. dr. ir. Johan Schoukens (Chairman, Vrije Universiteit Brussel)
Prof. dr. Michel Defrise (Secretary, Vrije Universiteit Brussel)
Prof. dr. Stefaan Tavernier (Promotor, Vrije Universiteit Brussel)
Prof. dr. ir. Gerd Vandersteen (Promotor, Vrije Universiteit Brussel)
Prof. dr. Karl Ziemons (University of Aachen, Germany) (external)
Prof. dr. Jose Perez (Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas, Spain) (external)

Acknowledgment

I am really grateful to Prof. Stefaan Tavernier, who gave me the opportunity to work on this very interesting subject and provide me a good scientific environment. Many thanks to Prof. Johan Schoukens, building a bridge for me to the modeling world at his department ELEC in the Engineering faculty. The combined working experience in a physics and engineering environment was very stimulating.

I would like to express my sincere gratitude to my promotor, Prof. Gerd Vandersteen, for his supervision, fruitful discussions, always being available for my difficulties and proof-reading of this thesis.

No words can express my gratefulness to my first promoter, Prof Peter Bruyndonckx. It is him who taught me everything about the "PET world", trained me to make presentations, helped me developing the ability to solve problems, write scientific paper and a lot more.

I would like to express my appreciation to the other members in my Ph.D committee: Prof Michel Defrise, Prof Karl Ziemons and Prof Jose Perez for reviewing my dissertation.

Thanks a lot to Marleen Goeman, Ann Pintelon and Jenny Lievens for their efficient administrative management during my stay at the IIHE and ELEC.

I am also grateful to my college Mateusz Wedrowski for all the measurements he did for me.

Furthermore, I would like to express my appreciation to all the members from IIHE and ELEC, for helping or discussing all the "small problems" and all the good times we spend together. It was very pleasant to work with you all.

Finally, I would like to give my special thanks to my parents and my husband for their understanding, endless patience and encouragement. Without them, it would be impossible for me to finish this thesis.

Contents

Preface

Since X-rays were discovered by Wilhelm Conrad Röntgen in 1895, various diagnostic tools have been developed for non-surgically looking inside the body. Today, doctors and researcher have a number of imaging modalities at their disposal such as X-ray computed tomography (CT), Magnetic resonance imaging (MRI), ultrasound (US), Diffuse Optical Tomography, Single Photon Emission Computed Tomography (SPECT) and Positron Emission Tomography (PET). The image formation in all these modalities have one thing in common: they need some form of energy interacting with the human body. The source of energy can be either internal or external. Sometimes the energy source is naturally present, other times it needs to be externally stimulated or administered. Figure 1 shows a classification of the most common modalities as a function of the energy source.

In diagnostic medicine, the imaging modalities are used to determine the cause of a disease based on abnormalities showing up. Pathologies of diseases can be understood as the morphological and physiological consequences of dysregulated molecular pathways in a living being. Hence, in order to understand diseases and try to detect them as early as possible (i.e. before the pathologies themselves are visible or when they are still very small), it is necessary to study biological processes on a molecular level. Ideally, a molecular imaging method is characterized by a high sensitivity, minimal background signal and a high resolution in both the spatial and time domain. Unfortunately, such a technology does not exist. The imaging modalities at hand all have their strengths and weaknesses. Depending on the problem one faces, different modalities have to be used. A comparison of the afore mentioned modalities with regards to detectability (i.e. what is the minimal concentration this modality can pick up) and spatial resolution is given in figure 2. The high resolution modalities like MRI and X-ray CT are characterized by inferior sensitivity and are therefore of less interest for molecular imaging

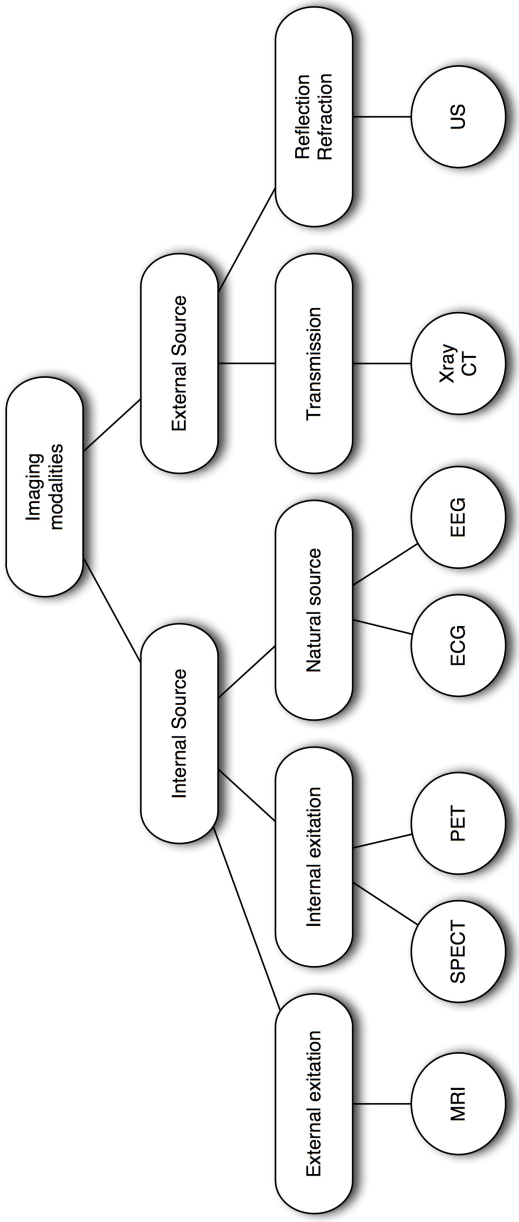


Figure 1 – Classification of the most common imaging modalities according to their energy source.

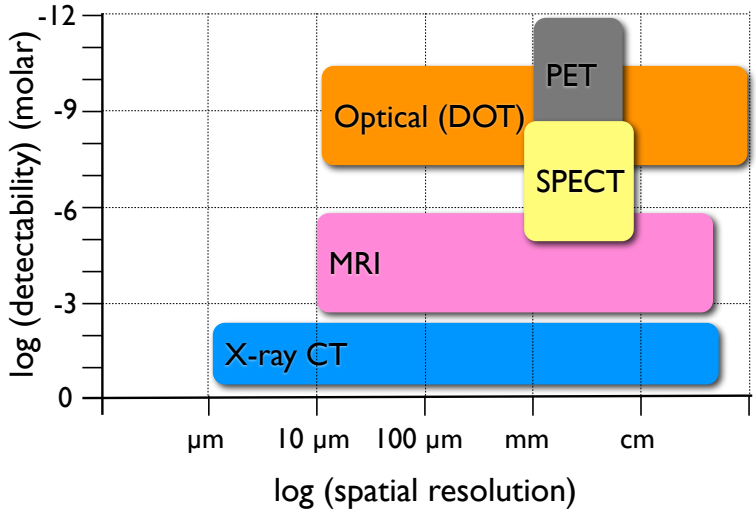


Figure 2 – Comparison of various imaging modalities with respect to spatial resolution and detectability. [62]

applications. They yield valuable structural (i.e. anatomical) and physiological information. On the other hand, optical imaging techniques and nuclear imaging modalities such as SPECT and PET are capable of detecting tracer concentration in the sub-nanomolar range. The images obtained by the different modalities are usually complementary, showing a different kind of information. As an example, figure 3 shows images of the brain using six modalities.

The work presented in this thesis concerns the development of novel instrumentation for PET imaging. PET, like its sister modality SPECT, is nuclear medicine's version of radiology. It is based on the administration of a chemical compound containing a radioactive isotope, also called tracer or radio pharmaceutical. The substance is given to the patient orally, by injection, or by inhalation. Once the compound has distributed itself according to the physiological processes active in the patients body, a radiation detector is used to make projection images of the gamma rays emitted during the decay of the agent. In other words, the radiation from the isotope in tracer acts as a beacon to report on the position of the molecule under study as it goes through a biochemical pathway. Different types of radioactive tracers are used, depending on the illness or organ being examined. In

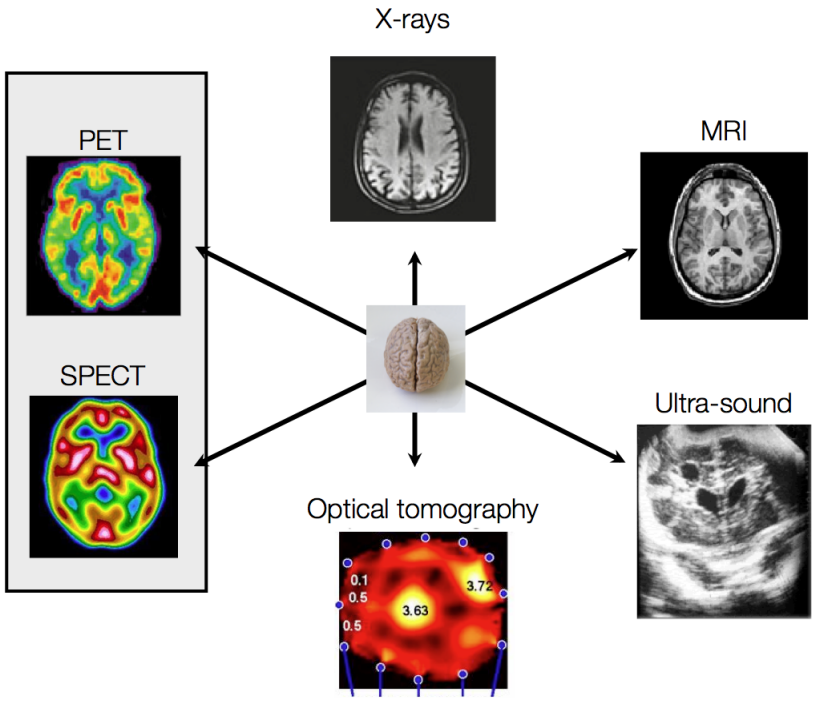


Figure 3 – Brain images from six different modalities showing different kinds of information.

the case of cancer diagnosis, for example, the radioactive tracer is often a glucose solution labeled with fluorine-18 called FDG (Fluoro-deoxyglucose). The uptake and metabolism of glucose in malignant cells is faster than in healthy cells because cancerous cells divide faster and therefore use more energy. Given that FDG can move into a cell but cannot move out again once it takes part in the biological processes inside the cell, the distribution of the decayed radioisotopes is a good reflection of the distribution of the cells. A PET scan can then map those areas of the body where glucose uptake is higher than normal and identify the location of the tumors.

Besides detecting tumors and determining where cancer has spread in the body, PET can also be used for many other diagnostic applications such as e.g. evaluating brain abnormalities or examining the blood flow to the heart. It is also an important part in treating many diseases. In addition to determining the extent of a cancer in the body, FDG scans are also useful for evaluating a patients

response to treatment. By comparing how far the cancer extended before and after a certain treatment, doctors can determine whether the treatment has been effective.

Positron emission tomography is also useful as a research tool. For example, it was used to map brain functions and study the heart. PET is now often used in oncology research, particularly in pre-clinical testing on laboratory animal. One advantage of PET is that researchers cannot only obtain detailed information, they can also obtain this information at different times in the course of prolonged experiments. This allows them to use fewer animals and minimize inter-animal variability in their work. PET is used in many other types of research as well. For example, to speed up the development of new drugs, the pharmacological properties are studied by injecting labeled drugs in animals and performing PET scans to evaluate how the drugs are metabolized and excreted. Another domain where PET is an important research aid, is the development of disease models and evaluating them on laboratory animals.

As a molecular imaging modality, PET images are often combined with anatomical images. The first multimodality imaging device was introduced by the combination of PET and CT scanners in the late 90s [7]. Although the PET-CT combination is a valuable tool in medicine, it also has some limitations and drawbacks:

- It does not allow simultaneous imaging of both modalities, which makes the fusion more difficult due to possible patient movement.
- CT adds to the total radiation dose the patient is exposed to.
- CT has low contrast for soft tissue, e.g. in the abdominal region.

To overcome these limitations, a lot of recent research efforts are put into the development of a combined PET-MRI scanner. This combination does not have any of the mentioned drawbacks, but is technologically much more difficult to realize.

A very important parameter in the design of PET scanners, is the accuracy and efficiency with which the gamma photons emitted by the isotope in the tracers can be localized by the detectors in the scanner. This ultimately determines the resolution and the quality of the image showing the distribution of the tracers. The goal of this thesis is to improve the performance of an upcoming detector concept for PET. This new detector configuration replaces the commonly used pixelated

detector approach with one based on monolithic scintillator detectors. Several algorithms have already been evaluated to make use of the signals extracted from this detector. However, they all require an extensive calibration procedure before the detectors can be used. This thesis describes a novel method we developed to employ these monolithic PET detectors **without** prior calibration by fitting the detector signals to a physics-based model and extracting the parameters of interest which are necessary to reconstruct PET images.

The first chapter introduces PET, the involved physical processes and the consequent limitations. It also introduces most of the PET related acronyms used. The detector technologies employed in PET are detailed in Chapter 2. This chapter also compares the commonly used pixelated PET detector with the monolithic ones we envisage to use. The use of monolithic PET scintillators in prototype PET detectors has also led to the development of various algorithms required to extract useful information from the detector signals. Chapter 3 will give an overview of existing algorithms developed by other research groups. The new approach we propose is described in Chapter 4. Its performance on simulated and experimental data is shown in Chapter 5. Comparison of this method with the others will be given in Chapter 6. Finally, the main conclusions from this thesis are summarized in the last Chapter.

Chapter 1

Positron Emission Tomography

1.1 PET imaging process

The process of PET imaging (figure 1.1) starts by injecting a radio-active tracer which contains a surplus of protons into the patent's body. The radio-active tracer stabilizes by the decay of a proton into a neutron. During this process, a positron and a neutrino are emitted. The latter is not detected. Depending on its initial kinetic energy, the positron will travel from a few tenths of a mm to a few millimeters in the environment, loosing its kinetic energy in multiple collisions. When it has lost most of its kinetic energy, it annihilates with an electron. The positron and electron disappear and are converted into two nearly back-to-back 511 keV photons. These are then detected by a ring of detectors surrounding the patient. The position of the annihilation is supposed to have occurred on the line connecting the couple of detected 511 keV photons. This line is called a line-of-response (LOR). By detecting many 511 keV photon pairs along all possible LORs going through the patient, the spatial and temporal distribution of tracers can be visualized using appropriate tomographic image reconstruction algorithms [19, 23, 31].

The process of PET imaging consist of four steps : (1) injection and uptake of a radio tracer, (2) emission of a 511 keV photon pair as a result of the annihilation of a positron emitted by the isotope in the tracer, (3) coincidence detection (i.e. quasi-simultaneous detection) of the emitted photon pairs and (4) tomographic image reconstruction of the radio-pharmaceutical distribution using the information of the registered coincidences. The following sections will explain the different steps involved in PET imaging in somewhat more detail.

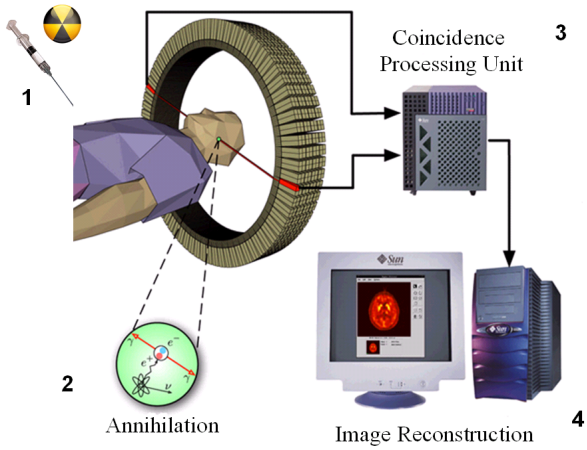


Figure 1.1 – PET imaging consists of four steps : (1) injection and uptake of a tracer, (2) decay of the radioisotope, followed by the emission and annihilation of a positron, (3) coincidence detection of the 511 keV photon pairs resulting from the positron annihilation, (4) reconstructing a 3D image using the positions of the registered photon pairs.

1.1.1 PET Tracers

PET, in contrast to SPECT, uses positron-emitting isotopes such as ^{18}F, ^{11}C, ^{13}N, ^{15}O. These elements occur naturally in many compounds of biological interest and can thus be easily incorporated in useful radio-pharmaceuticals. The molecule onto which the radioisotope is attached, is chosen as a function of the physiological process at interest. Four frequently used tracers are shown in table 1.1.

The disadvantage of the isotopes used in PET is that they require a cyclotron for their production. The short half-life time of the isotopes demand that the cyclotron should be either on-site or close by.

Radioisotope	Tracer	Application	Half-life
^{18}F	FDG	Brain function, tumors	109.8 min
^{15}O	^{15}O water	Blood flow, brain study	2.03 min
^{11}C	Carbon monoxide	Cerebral blood	20.4 min
^{13}N	^{13}N ammonia	Heart study	9.98 min

Table 1.1 – Examples of positron emitting radionuclides and their applications, from [65].

1.1.2 Positron emission and annihilation

During the decay of a PET tracer, a proton p in the nucleus of the radio isotope transforms into a neutron n by emitting a positron e^+ (or β^+), together with a neutrino ν_e which is not detected:

$$p \rightarrow n + e^+ + \nu_e \tag{1.1}$$

Once a positron is emitted, its kinetic energy is rapidly reduced by multiple Coulomb interactions, exciting and ionizing atoms on its trajectory. After traveling some distance until its energy reaches thermal equilibrium with the surrounding medium, it combines with an electron in the process of annihilation. Typical positron ranges vary from a few tenth of a mm to a few mm depending on their initial kinetic energy. The distribution of their ranges is narrowly peaked but still has a significant tail (figure 1.2). Table 1.2 gives the FWHM and FWTM of the positron range in water for a number of PET isotopes.

Radioisotope	FWHM (mm)	FWTM (mm)
^{18}F	0.102	1.03
^{15}O	0.501	4.14
^{11}C	0.188	1.86
^{13}N	0.082	2.53

Table 1.2 – FWHM and FWTM of the positron range for a number of PET isotopes in water. Data from [42].

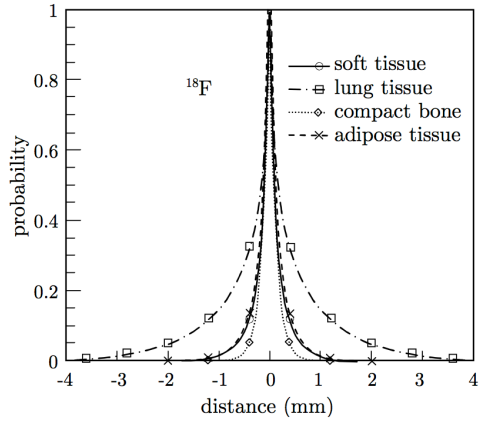


Figure 1.2 – Positron range distribution of ^{18}F, from [63]

When the original positron has lost its kinetic energy, it annihilates with an electron of the surrounding tissue and both particles are converted into two opposite, nearly back-to-back directed gamma photons. Due to the fact that the positron and electron annihilate approximately in rest, the energy of the resulting photons comes from the mass energy of the two particles, which can be computed by Einstein's mass-energy equivalence as:

$$E = mc^2 = m_e c^2 + m_p c^2 \tag{1.2}$$

where m_e and m_p are the mass of the electron and positron, c is the speed of light. According to this equation, the released energy is 1.022 MeV. Because the total momentum of the electron and positron at the moment of annihilation is practically zero, the two annihilation photons fly away in opposite directions and carry each 50% of the released energy, i.e. 511 keV (figure 1.3). The energy falls in the gamma-ray region of the electromagnetic spectrum. This makes these photons energetic enough to have a good chance to escape from the body but still be detected relatively easily by the detectors in the PET scanner (see Chapter 2).

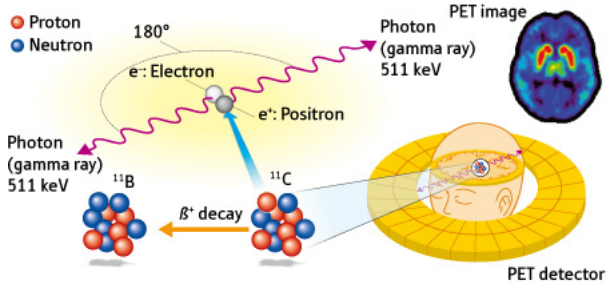


Figure 1.3 – The positron emitted during the decay of the tracers, annihilates with an electron of the surrounding tissue after it has lost its kinetic energy. The mass energy of the electron-positron pair is converted into two 511 keV photons leaving the annihilation scene in opposite directions.

1.1.3 Coincidence detection

The quasi-simultaneous detection of the two 511 keV annihilation photons by a ring of detectors surrounding the patient, is called a coincidence. We then know that the annihilation has happened somewhere along the LOR joining the two points where the photons were detected. Because the point of annihilation is very

close to the point of positron emission, one usually assumes that the LOR contains the position of the tracers that initially emitted the positron.

A certain percentage of the gamma rays that interact within the patient, either by photo electric absorption or Compton scattering (see Appendix A). As a result, four types of events can occur: (1) true events, (2) scattered events, (3) random events and (4) single events (figure 1.4). A *true event* is where both 511 keV photons leave the patient without any interaction and both are detected by the ring of detectors surrounding the patient. The accuracy with which the detectors can determine the exact arrival time of a photon is called *time resolution* (see section 2.3.4). Because the time resolution of PET detectors is not perfect, the PET coincidence electronics must foresee a *time window* $\triangle t$ within which two annihilation photons can be accepted. The length of the time window needed to accept most of the true coincidences depends on the time resolution, i.e. a better time resolution leads to shorter time windows. In current PET systems, the time window is in the ns range, i.e. usually less than 10 ns.

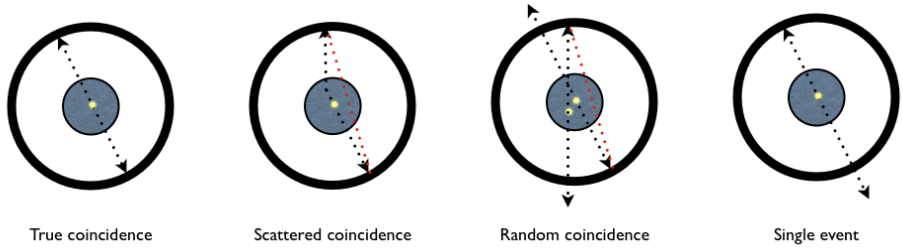


Figure 1.4 – Four types of events can occur in PET : (a) True coincidence event, (b) scattered coincidence event, (c) random coincidences event and (d) a single event. Types (b) and (c) yield incorrect LOR information and contribute to a relatively uniform background image that result in a loss of contrast. (d) The single event type detects only one of annihilation photons.

If one or both of the photons undergo a Compton interaction and are scattered in such a way that they still end up being detected, we call it a *scattered event*. During the scattering process, the 511 keV photons lose part of their energy. Hence, if the detectors in the PET scanner are capable of accurately measuring the energy of incoming photons, the scattered photons can be rejected. In a real system, the detectors usually employ a threshold of around 350 keV. Only if the measured energy is above this threshold, the photon is accepted. Imposing such a threshold also has a drawback. Because 511 keV photons impinging on the

PET detectors don't always deposit their full energy in the detector, they are also regarded as scattered photons and hence discarded if they don't cross the imposed energy threshold. This lowers the sensitivity of the PET system. In clinical PET, the scatter-to-true coincidence ratio ranges from 0.2 to 0.5 for brain imaging and from 0.4 to 2 for human body imaging [16].

Random events occur when two 511 keV photons, originating from two different annihilations (i.e. different tracer molecules) are detected as a coincidence, i.e. within the time window $\triangle t$ of the PET system. These accidental events don't carry any information about the tracer distribution in the patient and only add to the background noise in the image, reducing the image contrast. The probability of detecting random events increases when the length of the time window gets longer. The rate of random events detected by two detectors can be expressed as a function of the number of individual photons arriving on those two detectors and the time window in the following way :

$$R_{random} = R_1(1 - e^{-R_2 \triangle t}) \tag{1.3}$$

where R_1 and R_2 are the rates of individual photons, also called single events, detected by detector 1 and 2 respectively. If the average arrival time between two single events on detector 2 is much larger than $\triangle t$, then $\frac{1}{R_2} \gg \triangle t$ or $R_2 \times \triangle t \ll 1$. In this case equation (1.3) can be approximated as

$$R_{random} = R_1 R_2 \triangle t \tag{1.4}$$

Because the rate of individual photons R_1 and R_2 hitting the two detectors increases with the amount of tracer activity in the patient, it follow from (1.4) that the random rate R_{random} in a PET scanner increases quadratically with the amount of tracer administered. The true event rate increases only linear with the amount of tracer. Hence increasing the amount of tracer will also increase the random event ratio, i.e. the ratio of the random event rate over the true event rate. Equation (1.4) also shows that it is vital to keep the time window $\triangle t$ as small as possible to minimize random events. The minimum time window that can be operated in a PET scanner, depends on the time resolution of the detectors used. In actual PET scanners, the ratio of random-to-true coincidence rates is about 0.1 to 0.2 for brain imaging [16]. The randoms can be corrected for on a statistical basis, i.e. it is impossible to identify random events on an individual basis but the average number of randoms between a pair of detectors in a PET scanner can be determined. If the electronics in a PET system foresees

a counter for each detector, the random rate between each detector pair can be computed using equation (1.4). The method which is usually implemented to remove the contribution of random events, is called the *delayed coincidence technique*. Whenever a coincidence window is opened, a second coincidence window is opened after some delay that is much bigger than the length of the time window itself. If a coincidence between the two detectors is detected within the delayed time window, it must have been a random coincidence. Hence after a scan, for each detector pair the system has measured the total number of coincidences and an estimate of the average number of random events between those two detectors. Subtracting the two number removes the number of random events.

Finally, it is also possible that just one of the two annihilation photons is detected (figure 1.4 (d)), either because the second photon misses the detector or is not detected by the detectors. These events are rejected electronically because no second photon was detected during the length of the time window that was opened when the first photon arrived. Hence, the single events don't influence the image quality but they add extra dead time to the system during which it is not responsive to other good events.

1.1.4 Image reconstruction

The last step in the PET imaging process is the image reconstruction using the measured coincidences. A mathematical treatment of the algorithms involved is beyond the scope of this work, but intuitively one can see that the LORs of the coincidences generated by a point source, intersect at that point. A mathematical way of drawing an LOR, is called *backprojecting*. If an image can be considered as a matrix of pixels, then backprojection consists of putting the number of coincidences measured along a given LOR (i.e a detector pair) in all the pixels that are intersected by that LOR (figure 1.5). If this is repeated for all possible LOR's, the backprojected imaged of a point source gets the shape of a $\frac{1}{r}$ distribution around the position of the point source, where r is the distance to the point source (figure 1.6). In other words, the backprojected image is a convolution of the true image and a $\frac{1}{r}$ function. Deconvoluting the $\frac{1}{r}$ function from the backprojected image, yields the real image. This deconvolution is usually done in the spatial frequency domain by multiplying with the appropriate filter function, i.e. the Fourier transform of the $\frac{1}{r}$ function. Hence the reconstruction algorithms is called Filtered-BackProjection algorithm or FBP.

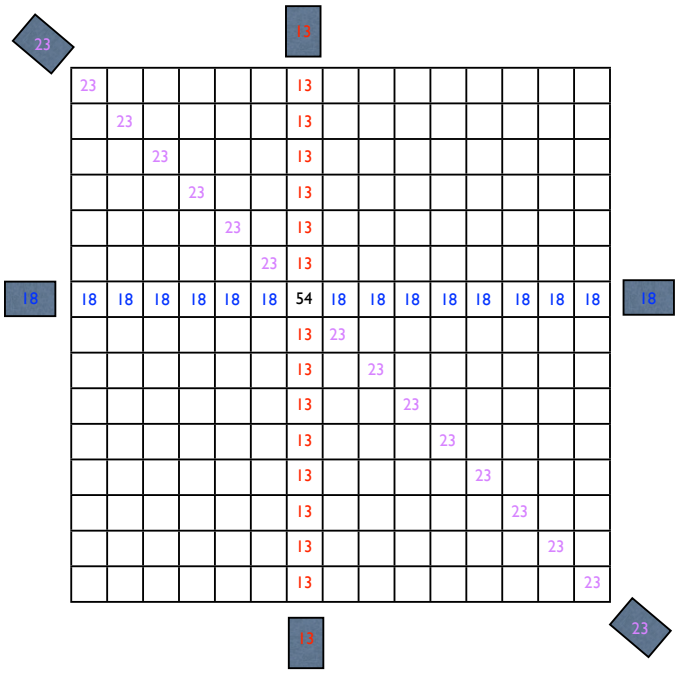


Figure 1.5 – Backprojection consists of putting the number of photons measured along an LOR into all image pixels crossed by that LOR.

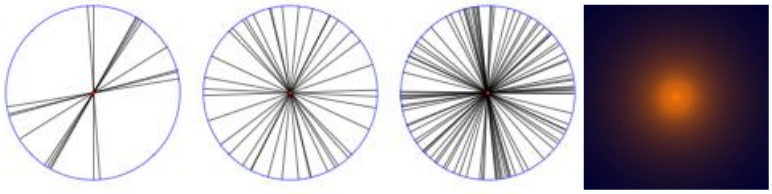


Figure 1.6 – The backprojected LORs all go through the point where the annihilation photons originated. When more and more LORs are backprojected, the image becomes a $1/r$ distribution.

1.2 Fundamental resolution limits of PET

When using the backprojection process during the image reconstruction, one assumes that the tracers which emitted the positron lies on the LOR defined by the two annihilation photons. There are two physical processes that make this assumption not completely valid and hence introduce some error.

1.2.1 Positron range

The first contribution to the resolution limit is given by the positron range, already mentioned in section 1.1.2. When a positron is emitted during the decay of a tracer, the positron travels a distance based on its kinetic energy, and hence the LOR actually passes through the annihilation point instead of the positron emission point. This range degrades the spatial resolution of a PET system and introduces some blurring into the image. For example, the positron range distribution of F-18 in water has a 0.1 mm full-width at half-maximum (FWHM) and O-15 has 0.5 mm FWHM [22]. One way to physically reduce this effect is using a strong magnetic field [8]. The transaxial resolution of the scanner will be improved since the positron range is reduced in that plane (figure 1.7), while the axial resolution is kept unaltered. This is an attractive method since a PET scanner can be combined with MRI to have a combined structural and functional multimodality imaging platform.

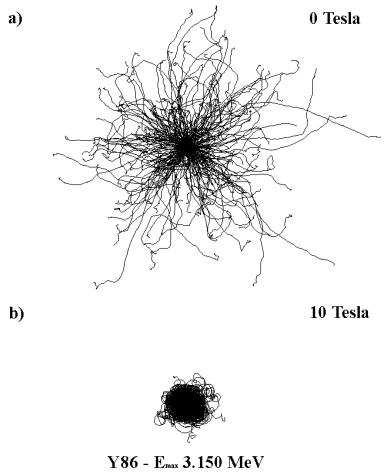


Figure 1.7 – Influence of the magnetic field on the positron range, for the radioisotope 86Y (E_{max}=3.15 MeV) in water, illustrated by the GEANT4 simulation toolkit. a) 0 Tesla and b) 10 Tesla field orthogonal to the paper plan. The average distance between the emission and annihilation point is reduced in presence of the magnetic field. [8]

1.2.2 Photon non-collinearity

The second fundamental physical limit on the spatial resolution in PET is the non-collinearity of the pair 511 KeV annihilation photons, i.e the flight paths of the two gamma rays have a small angular deviation from 180^o (figure 1.8). This is a result of the fact that at the time of annihilation, the electron is not perfectly in rest. The residual momentum of the electron-positron system can introduce a small deviation from the 180° angle between the two annihilation photons. The distribution of the angular deviation is Gaussian with a FWHM of 0.5°[16].

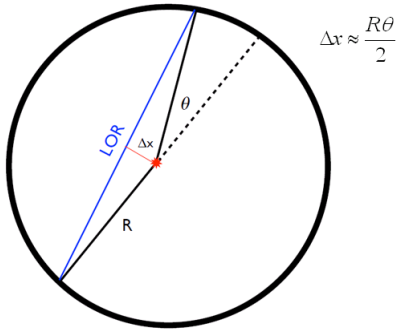


Figure 1.8 – Relationship between the acolinearity error Δx, the scanner radius R and the deviation from colinearity θ.

The amount of blurring introduced by the non-collinearity is given by the shortest distance Δx between the LOR and the annihilation point (figure 1.8) :

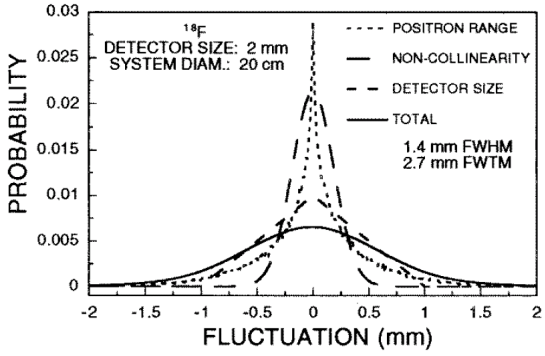
$$\Delta x \approx \frac{R \times \theta}{2}$$

where R is the scanner radius, and θ is the acolinearity between the two 511keV photons. The FWHM of Δx is then about $0.0022D$, where D is the diameter. Hence, for a whole body PET system with a diameter of 0.8 meter, the acolinearity introduces a Gaussian blurring with a FWHM of about 1.74 mm. A way to minimize this effect is by reducing the diameter of the ring detector as much as possible, which is more difficult for a whole-body PET scanner than for a brain or small animal scanner.

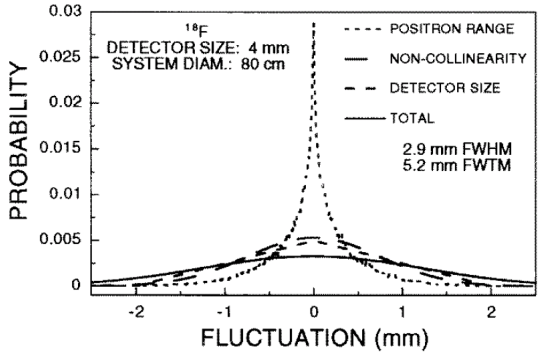
1.2.3 Combination of the resolution factors

The positron range and non-collinearity contribute to the system spatial resolution. Figure 1.9 (a) shows the combination of these factors on the spatial resolution for a 2mm crystal width and 20 cm system diameter, which is typical for animal imaging system designs, and figure 1.9 (b) illustrates the case for a 4 mm crystal width and 80 cm diameter system diameter, which is a type of human imaging system design. The positron source used in this example is ^{18}F and the detector is of a discrete type, which has a spatial resolution equal to the half of the width of the detector in the central of the field of view (more discussion in section 2.4.1).

For the small animal system, the total spatial resolution combined by the three factors are 1.4 mm FWHM and 2.7 full-width at tenth-maximum (FWTM). Thus, the resolution blurring due to positron range and non-collinearity adds about $(1.4 - 1)/1 = 40\%$ to the system resolution, relative to the detector resolution. In this situation, the ^{18}F positron range and the detector size are the dominating factors in the system spatial resolution. For the large human system, the total spatial resolutions are 2.9 mm FWHM and 5.2mm FWTM. Here the non-collinearity and detector size dominate.



(a)



(b)

Figure 1.9 – Spatial resolution blurring factors and their combination for ^{18}F with (*a*) 20 cm system diameter and 2 mm wide detectors, and (*b*) 80 cm system diameter with 4 mm detectors. Figure from [42].

Chapter 2

PET detectors

The detection of 511 keV photons emitted as a result of a electron-positron annihilation following the radio-tracer decay is one of the most important low-level stages in nuclear imaging modalities such as PET. The ability to precisely determine the coordinates of scintillation events implies a low uncertainty on the data passed to the reconstruction algorithm and thus results in high quality tomographic images.

Equally important is the detection efficiency of the annihilated photons. Neglecting photons is undesired in PET scans since higher efficiencies allows reducing the injected radio-tracer dose and the patients exposure to radiation. In addition, PET imaging is based on coincidence detection. If the sensitivity of its detectors is reduced by a factor p, then the sensitivity of the PET scanner is reduced by a factor p^2.

This chapter introduces the basic terms and physics involved in the detection process. It also introduces the two classes of scintillation based detectors for PET: pixelated detectors found in the traditional designs and monolithic PET detectors which are the subject of this thesis. The important performance parameters will be discussed for both classes. An overview of existing algorithms to obtain the 3D interaction position in monolithic detecters will be given in the next chapter while a novel method to obtain the 3D interaction position is described in Chapter 4.

2.1 Scintillators in PET

An ideal gamma ray detector would convert the energy of a 511 keV gamma directly into an electrical signal. This could theoretically be done using solid state

detectors such as silicon based detectors (e.g. APD) or Cadmium Zinc Telluride (CZT) detectors. However, an APD is made of low Z[1] materials, and hence the stopping power is rather low. This makes it ineffective as a direct gamma ray detector. As for CZT, the time resolution is limited by the rather slow speed of the charge collection and hence the thickness of the detector is restricted [43].

State of art detectors for 511keV gamma rays used in PET rely on scintillator. The detector efficiency is increased by transforming the energy of the 511 keV photon into an electrical signal in an indirect way by first converting them into many low energy photons (in order of a few eV) which can then be easily detected by classical photo detectors.

2.1.1 Brief description of the scintillation mechanism

The first step in this conversion process is done by scintillators. Scintillator materials exist in two different types: organic and inorganic scintillators. Organic scintillators are composed of organic molecules, consisting for the most part of low Z atoms and have therefore a long radiation length. Organic scintillators are mainly used to track charged particles. PET scanners therefore make only use of inorganic scintillators, which are composed of ionic crystals that contain a large percentage of atoms with a high Z. Hence, they can efficiently stop gamma rays.

In a pure crystal, there are two energy bands available for electrons: the valence band and the conduction band (figure 2.1 left). The energy range inbetween the valence band and the conduction band is called the band gap or forbidden gap, and is void of electrons in case of pure crystals. If an electron from the valence band receives enough energy ($>$ band gap energy), it can cross the band gap and reach the conduction band.

The scintillation mechanism in inorganic scintillators can be divided in three steps:

1. Excitation
2. Thermal relaxation
3. Photo emission

The excitation occurs when a gamma photon interacts with an inorganic scintillator, creating a hole in the valence band of the scintillator and an energetic electron.

[1] Z stands for the atomic number of a material

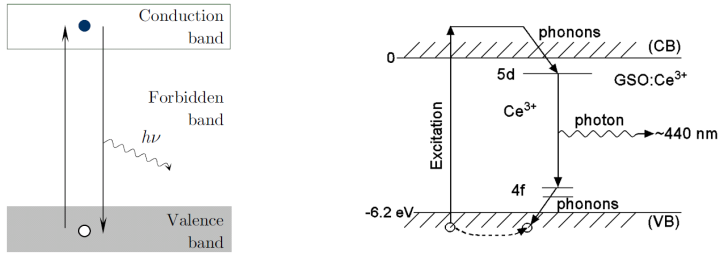


Figure 2.1 – (left) Bands in a pure crystal, (right) GSO doped with Ce is an example of a scintillator containing an activator. It has energy levels within the forbidden gap. Decay from an electron between these levels results in the emission of a photon with a wavelength below the absorption wavelength of the scintillator.

The electron losses its energy through multiple inelastic scattering, leading to a multiplication of the excitations.

The thermal relaxation deals with the thermalization of the electrons and holes through the emission of phonons (i.e. heat energy). At the end of this thermalization, the electrons are at the bottom of the conduction band and the holes are at the top of the valence band [58].

The electrons in the conduction band could then recombine again with holes in the valence band, resulting in the emission of a photon. This process is not very efficient since the photons can be re-absorbed in the scintillator to create a new electron-hole pair, i.e. the emission wavelength = absorption wavelength. In other words, the scintillator is not transparent to its own scintillation light. In addition, the energy corresponding to the width of the band gap corresponds typically to photons that have a wavelength above the visible or soft UV light. This makes them hard to detect by the photo detector.

To overcome the absorption problem in the crystal, activators are added to the scintillator material. These activators have energy levels within the band gap (figure 2.1 right). During the thermal relaxation phase, the holes ionize the activation sites. The electrons migrate until they drop into an ionized activation site, leaving the activator in an excited state. These excited activator states decay through the emission of a photon to the lower energy level just above the valence band. Due to a process called *stokes shift* [68], the energy of the emitted photon is lower than the energy required to excite the activators. Therefore, the crystal is transparent to the emitted scintillation photons. Since the energy difference between the two energy levels involved in the decay is smaller, the

emitted scintillation photons have a longer wavelength, situating them in the UV or visible light part of the spectrum.

The decay of the excited activator states is not instantaneous, i.e. the excited states have a certain lifetime. Consequently, the scintillation light is not produced in one infinitely short flash, but the intensity rises sharply to a maximum and then decays slowly depending on the lifetime, i.e.

$$I(t) = I_0 e^{-\frac{t}{\tau}} \tag{2.1}$$

where τ is the decay time. These decay times vary from a few tens of ns to a few hundred ns depending on the scintillator.

Only a fraction of the total energy deposited by the initial gamma photon is converted into low energy scintillation photons. Each of the three steps involved in the scintillation process has a certain efficiency to transfer the energy from one step to the next. The product of the three efficiencies determines the amount of light eventually produced by the scintillator.

There is another type of inorganic scintillator, which does not use luminescence impurities. Bismuth Germanate ($Bi_4Ge_3O_{12}$ or BGO) is an example of this kind of "pure" inorganic scintillator. It uses an optical transition of the Bi^{3+} ion that is a major constituted of the crystal. The optical absorption and emission spectra of Bi^{3+}states have a relatively large shift. As a result, the crystal remains transparent to most of its own emission.

2.1.2 The Properties of Scintillators

The ideal scintillation material should have the following properties [9, 39, 59]:

1. High density ρ and high atomic number Z. When Z increases, the average mean free path of the 511 keV photons decreases, resulting in a higher probability to interact in the scintillation crystal. In addition, the probability to undergo a photo electric absorption (photo fraction) instead of a Compton scattering also increases as Z increases (see figure 2.2). This is important because one wants the 511 keV photon to deposit all its energy in the crystal.

2. High light yield or luminosity (i.e. the number of scintillation photons produced per unit energy deposited in the scintillation crystal) to minimize the statistical error on the conversion step.

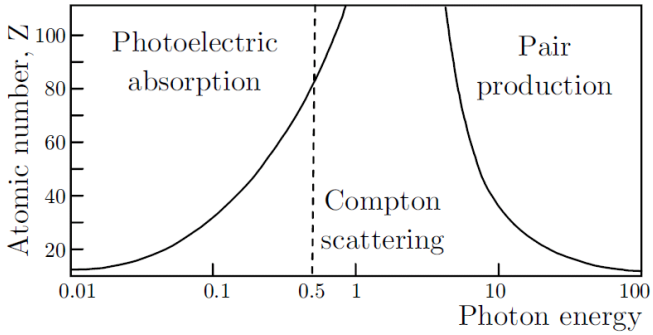


Figure 2.2 – Diagram showing the most probable photon interaction as a function of its energy and the atomic number Z of the material, from [16].

3. The conversion should be linear, such that the total energy of the low energy scintillation photons is proportional to the energy deposited by the incoming photon.

4. The wavelength of the scintillation light should match the wavelength where the photo detector is most sensitive.

5. The scintillator should be fast, i.e. the decay time of the induced luminescence should be short. This factor is important to achieve a good time resolution needed to minimize the time window. Short scintillation pulses also result in a smaller dead time of the detector, i.e. the PET system is capable of handling higher rate of events.

6. The material should be of good optical quality and subject to manufacturing sizes large enough to be of interest as a practical detector. This is important because the scintillator should be made thick enough to efficiently stop the incoming 511 keV photons.

7. Its index of refraction should be near that of glass (1.5) to permit efficient coupling of the scintillation light to a photomultiplier tube or other light sensor. Large mismatches in index of refraction results in significant internal reflection at the scintillator/photodetector boundary and reduces the light transmission to the photodetector.

8. The medium should be transparent to the wavelength of its own emission for good light collection.

9. No technological constraints, e.g. hygroscopic or fragile.

No scintillator can satisfy all the requirements mentioned above. The available scintillators are sub-optimal for different applications. Some properties of common inorganic scintillators used in PET are listed in table 2.1. A widely used inorganic scintillator in the early time of PET is NaI(Tl). It was popular because of its excellent light yield and the fact that it can be machined easily into large crystal volumes. The main drawbacks of NaI(Tl) are that it has a long decay time (230 ns), and a low density resulting in long absorption length. In addition, it is hygroscopic and fragile which makes the handling of the crystal less practical. It has therefore been gradually replaced by BGO which first appeared in 1973. BGO has a high density ($7.13 g/cm^3$) and a large atomic number Z=83, which gives it significantly better stopping power. But it has a relatively low light yield, i.e. about 10-20% of NaI(Tl) and also a long decay time (300ns).

Scintillator	NaI(T1)	BGO	LSO	LuAP
Wavelength of peak (nm)	415	480	420	365
Refraction index	1.85	2.15	1.82	1.94
Decay time (ns)	230	300	47	17
Mean free path (cm)	2.91	1.04	1.14	1.05
Density ρ(g/cm^3)	3.67	7.13	7.4	8.34
Abs. Light Yield in Photons/MeV	38000	8200	26000	17000

Table 2.1 – Properties of common inorganic scintillators used in PET. Data from [39, 15].

In 1991, $Lu_2(SiO_4)O : Ce^+$ (LSO) was first described as a scintillator. LSO has a much larger light yield (75% of NaI(Tl)), a fast decay time of 47 ns and an emission spectrum conveniently peaked at 420 nm. At the same time it has nearly the same high stopping power as BGO. Hence LSO has become one of the most suitable scintillator materials for PET. However, it contains the long lived radioactive isotope ^{176}Lu, which decays via β^- emission with half-life of 3.1^{10} years. This leads to a background count rate of approximately 240 counts/s per cm^3 of LSO [30] continuously distributed over the full energy spectrum up to 596keV [49]. But since PET is based on coincidence detection, this natural background is not a prohibitive factor.

Another interesting scintillation material is $LuAlO_3$ (LuAP), which has a very short decay constant of 17 ns but its strong self-absorption of the scintillation light limits the crystal thickness.

2.2 PET photo detector

When talking about PET detectors based on scintillators, we should make a distinction between the high-energy annihilation photons (511 keV) that are absorbed by the scintillator and the burst of low-energy optical photons that are subsequently emitted by the scintillator and converted into an electric current by a photo detector, as shown in figure 2.3. It is the number of visible light photons generated in the scintillator that determines the amplitude of the electrical pulse. Those electrical pulses will be later used to establish where the 511 keV interacted in the scintillators.

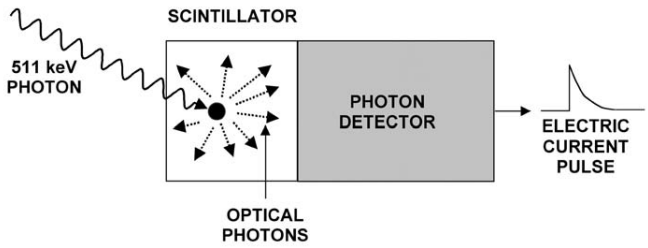


Figure 2.3 – Process of converting the enrgy of a 511 keV gamma into a detectable electrical pulse, from [59].

2.2.1 The properties of the photodetector

The photo detectors are designed to convert the low energy scintillation light into electrical signals. In general, it involves generating a detectable electrical signal proportional to the number of incident scintillation photons. The process involved can be separated in three steps. First, primary photo electrons or electron-hole (e-h) pairs are generated by incident optical photons. Then, one or more multiplicative bombardment steps (i.e an avalanche multiplication process) amplifies the primary electrons to a detectable level of secondary electrons. Finally, the electrical signal is formed by the collection of these secondary electrons.

Some important characteristics of a photo detector are:

1. Quantum efficiency (QE): the percentage of impinging photons that are converted into primary photo electrons. This is usually a strong function of the photon wavelength λ.

2. Collection efficiency (CE): the percentage of primary photo electrons that are captured for the amplification process.

3. Gain (G): the mean multiplication factor of the amplification process.

4. Excess noise factor (ENF): describes the uncertainty in the produced signal due to gain fluctuations σ_G.

5. Equivalent noise charge (ENC): the amount of signal charge, usually expressed in number of electrons, which injected at the input of the electronic amplifier, generates an output signal whose amplitude is equal to the r.m.s. output noise voltage.

6. Time jitter: the time fluctuation on the difference between the moment a photon enters the photo detector and the moment of the output pulse. The achievable lower bound on the time jitter is a function of the signal-to-noise ratio (SNR).

A good photo detector should have a high QE and CE, a high gain with low gain fluctuations (i.e. low ENF), a low ENC and a low time jitter.

The vast majority of commercially available PET scanners use photomultiplier tubes (PMT), which have a low ENF and a high gain. The latter makes the readout noise negligible. A limitation of the PMTs is their low quantum efficiency (around 25%). Moreover, a PMT is very sensitive to magnetic fields and thus it is impossible to operate a PMT-based PET in combination with a MRI scanner.

An interesting alternative to PMTs are the avalanche photo diodes (APD). They have a higher quantum efficiency for blue-UV light (around 70-80 %) which is emitted by most scintillators and are insensitive to magnetic fields. They are also very light and not as bulky as PMTs. This makes it possible to create very compact detectors and allow detector geometries not possible with PMT, e.g mounting the APD on top of the scintillation crystal without much interference of the incoming 511 keV photons. As will be shown later, this configuration opens possibilities for parallax correction (section 2.4.1) or better positioning in monolithic scintillators (section 5.2.8). The weak points of APD's compared to PMT's are their need for very low noise pre-amplifiers due to the low internal gain, worse timing accuracy, a higher sensitivity to changes in temperature and a low voltage needed to operate the APDs.

Type	QE	CE	ENF	G	Time jitter
PMT	0.15-0.25	0.9	1.3	10^6	100ps
APD	0.7	1.0	1.75	100	a few ns
GM-APD	0.15-0.3	1.0	1	10^6	90 ps

Table 2.2 – General comparison of photo detectors. Data from [72, 3, 13].

In recent years, a new type of photo detector called Geiger-mode avalanche photo diodes (GM-APD) or silicon photo multipliers (SiPMT), has generated quite some interest. It has a high gain like PMTs and hence a very good signal to noise ratio and it can also be used in magnetic fields. The first commercial samples are coming onto the market now and they look very promising for future medical imaging applications. Table 2.2 shows a comparison of typical values of the properties from these three types of photo detector. A brief description of their internal mechanisms will be given in the following subsections.

2.2.2 Photomultiplier tube

A simplified construction of a PMT is shown in figure 2.4. The scintillation light enters a PMT through a transparent window. On the back side of this window is a thin photo sensitive layer, called the photo cathode. When they interact, incident scintillation photons create low-energy electrons with an efficiency given by the QE of the PMT. The charge produced by the photo cathode contains only few hundred electrons and this is not sufficient to serve as a usable electrical signal. Therefore, PMTs are equipped with a series of metal dynodes in a vacuum tube to multiply the electrons. The photo electrons generated from the photo cathode are focused and accelerated by an electric field to the first dynode. Each of these electrons liberates additional electrons in the first dynodes, and a small fraction of these liberated electrons can escape from the surface of the dynode and reach to the second dynode. This process is repeated until the last electrode, called the anode. The anode is at ground potential, while the photo cathode is at some large negative potential, typically -500 V to -2000V . Each dynode is kept at some large negative potential, increasing from the previous one in steps of typically 150 V. These voltages are obtained with a resistor chain. The total multiplication factor of the electrons can reach up to 10^6 or more. The electrical output signal is taken from the anode and is proportional to the light received by the PMT.

Presently, the QE of PMTs is limited to about 30%. The CE represents the percentage of photo electrons that are captured by the first dynode and has a typical value around 90%. The total PMT gain is the product of the electron multiplication factors at each of the dynodes. The total PMT gain is usually in order of 10^6. The signal g at anode, expressed in electrons, can then be defined as

$$g = N_\gamma \cdot QE \cdot CE \cdot G \tag{2.2}$$

where N_γ is the number of incident scintillation photons.

Each dynode amplification process introduces a widening of the spread in the distribution of the electron gain. This spread is represented by the ENF, i.e.

$$ENF = 1 + \frac{1}{\delta_1} + \frac{1}{\delta_1 \cdot \delta_2} + \cdots + \frac{1}{\delta_1 \cdot \delta_2 \cdots \delta_n} \tag{2.3}$$

where δ_i is the gain of the i^{th} dynodes [9]. A typical value of the ENF in a PMT is about 1.3.

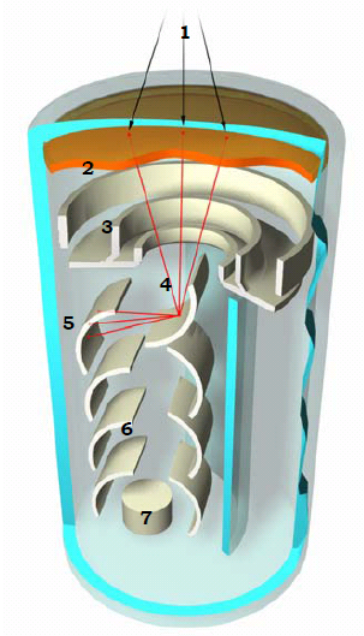


Figure 2.4 – Scheme of a photomultiplier tube: 1 - incident light, 2 - semi-transparent photo cathode, 3 - focusing electrodes, 4 - typical photo electron trajectories, 5 - vacuum tube (glass), 6 - dynodes, 7 - anode. Figure from [9].

2.2.3 Avalanche photo diode

A simple silicon photo diode is based on the formation of a PN semiconductor junctions. The PN junction principle is presented in figure 2.5. N-type silicon is silicon that has been chemically combined (doped) with phosphorus gas to make

it conductive. A silicon atom has four electrons in its outer shell and bonds tightly with four surrounding silicon atoms creating a crystal matrix with eight electrons in the outer shells. However, phosphorus has five electrons, and when combined, the fifth electron becomes a "free" electron that moves easily within the crystal. Because the charge carriers are electrons, n-type refers to a negative charge. In contrast, p-type silicon is silicon doped with boron gas. Boron has only three electrons in its outer shell and can bond with only three of the four surrounding silicon atoms. This leaves one silicon atom with a vacant location in its outer shell, called a "hole," that readily accepts an electron. Because the charge carriers are holes, p-type silicon is said to have a positive charge. The PN junction is the interface at which p-type silicon and n-type silicon make contact with each other. The density difference of electrons and holes makes them diffuse to the adjacent region of the other type. After losing its free electrons, the n-type region becomes positively charged, while similar in the p-type region, a negative charge will be created. Thus an electric field is built up. Once the electric field becomes strong enough, the diffusion of the electrons and holes will be significantly reduced. This region with reduced number of charge carriers is called the depletion region.

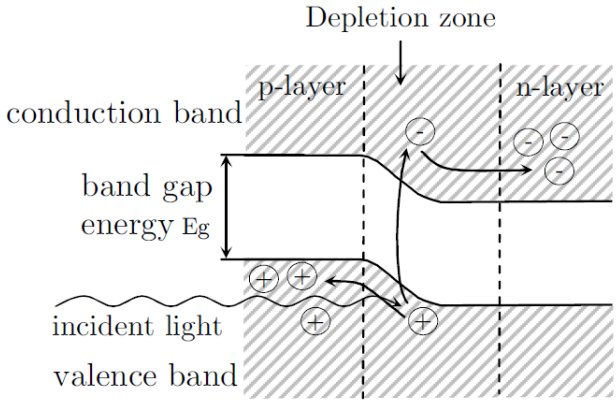


Figure 2.5 – P-N junction principle.

When a scintillation photon, which typically carries about 3-4 eV of energy, falls into the depletion region and interacts, it has sufficient energy to create electron-hole pairs since the band gap energy of a semiconductor is approximately 1-2 eV. The created charges will be pushed to the side by the electric field in the depletion zone and collected, contributing to a current signal.

Using external metal contacts on the n-type and p-type silicon side, one can enhance or reduce the internal potential. The former case is called reverse biasing and increases the electric field in the depletion area, which efficiently sweeps out any electron-hole pairs created by absorbing incident photons.

A normal silicon photo diode only converts some of the impinging photons to electrons, which means it generates a signal which has no internal gain. A small signal leads to a reduced signal-to-noise ratio, which makes the simple photo diodes not suitable for PET. An improved version of a photo diode is the avalanche photo diode. By adding a big reverse bias voltage over the P-N junction, an avalanche multiplication process can be initiated. One common choice of configuration of avalanche photo diodes is shown in figure 2.6. A photon enters through a very thin p+-layer, interacts in the silicon area, and the electron is drawn to the right by a low field region, where it enters a high-voltage multiplication region. Here the electron can obtain sufficient energy to start an avalanche of electron-hole pair multiplication. A typical gain of 100 can be achieved in this stage, resulting in a small but usable signal.

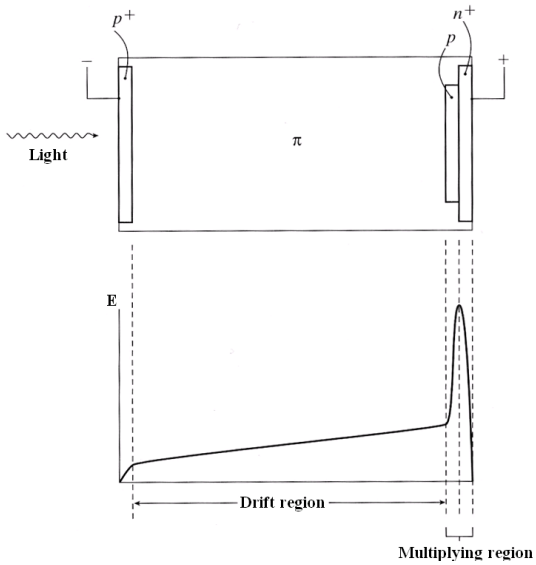


Figure 2.6 – A configuration of avalanche photo diodes. Figure from [39].

A typical gain in an APD is only about 100 because the reversed bias voltage should be carefully added in order to make a good control of the fields. This is because only the electrons should contribute to the signal amplification in the

avalanche process. The holes, on the other hand, should be extracted to the opposite direction faster than they can multiply themselves. Hole multiplication also produces additional free electrons and leads to a runaway process. It will result in a very high gain of about 10^6, but the signal is constant rather than being proportional to the impinging number of photons. Hence, it cannot be used to indicate the amount of the detected scintillation light. In practice, the APD voltage applied is lower than the voltage at which this breakdown occurs. The gain factor is an exponential function of the reverse voltage, and it can only be in the order of a few hundred. Beside that, the gain is extremly sensitive to the voltage and it is also a strong function of temperature. It decreases by a few percent per degree as the temperature is increased [39].

Depending on the wavelength of the impinging photon, the QE in APDs usually ranges from 60% to 80%, which is much higher compared to PMTs. Figure 2.7 shows an example of the QE as a function of the wavelength of the incident photons for the Hamamatsu S8550 APD, which is the detector used in this thesis. The peak photon wavelength value produced by LSO is 420nm. At this wave length, the Hamamatsu S8550 APD has a QE of about 70%. The ENF for an APD is worse than for a PMT because of the larger statistical fluctuations that occur during the avalanche multiplication process. A typical value of the ENF for APDs is around 2.

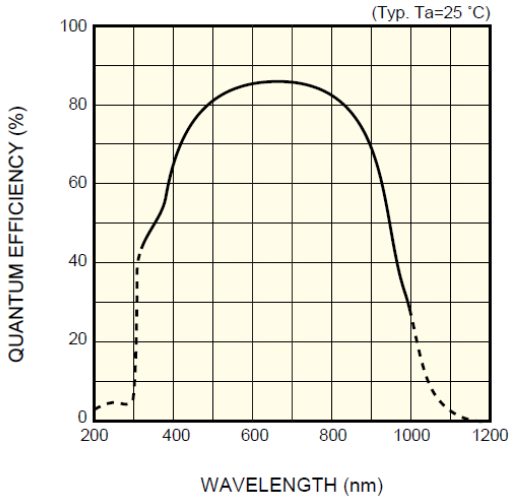


Figure 2.7 – Wavelength dependence of the QE for the S8550 Hamamatsu APD. Figure from [29].

2.2.4 Geiger-mode avalanche photo diode

The gain as a function of the reverse bias voltage for an ordinary photo diode, a normal APD (operated in linear mode) and Geiger-Mode APD are compared in figure 2.8. When the APD is biased above the breakdown voltage, the electric field in the multiplication region is so high that, when an impact ionization chain is initiated, not only electrons but also the holes multiply faster than they can be extracted. In this way gains of about 10^6 can be reached. APDs working in this bias voltage regime are called Geiger-Mode APDs. They act as a digital device, switching from an off-state to a conducting state in reaction to the detection of a single photon.

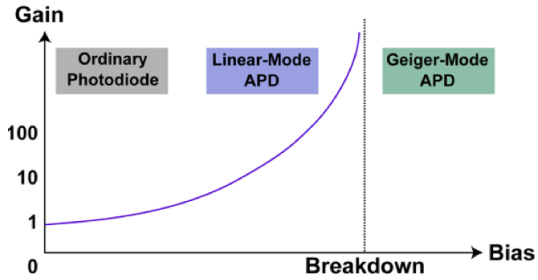


Figure 2.8 – Typical gains for an ordinary photo diode (gain = 1), a normal APDs biased below the breakdown voltage and a Geiger-mode APDs biased above the breakdown voltage

GM-APDs are known under several different names. Due to their PMT-like characteristics (i.e. much higher gain than APDs and better timing capabilities), they are also commonly called Silicon Photo multipliers (SiPMT). And because they have the ability to count single photons, they are sometimes also known as multi-pixel photon counters (MPPC).

Because a GM-APD is a binary device, a large single pixel cannot determine how many scintillation photons imping at one given time. The idea of using Geiger-Mode APDs to detect many scintillation photons, is to subdivide the sensitive area of the GM-APD in an abundance of very small micro cells in the order of 25×25 to 100×100 μm^2 each (figure 2.9 left). Every micro cell can detect one single impinging scintillation photon. The output signal from the Geiger-mode APD pixel is the total sum of the outputs from all the micro pixels , i.e. a signal proportional to the number of detected scintillation photons (figure 2.9 right) which on its turn is related to the amount of energy deposited by a 511 keV photon

in the scintillation crystal. In other words, a GM-APD works in photon counting mode. As long as the number of micro cells is much larger than the number of impinging scintillation photons, it will respond linearly with the magnitude of the scintillation light signal.

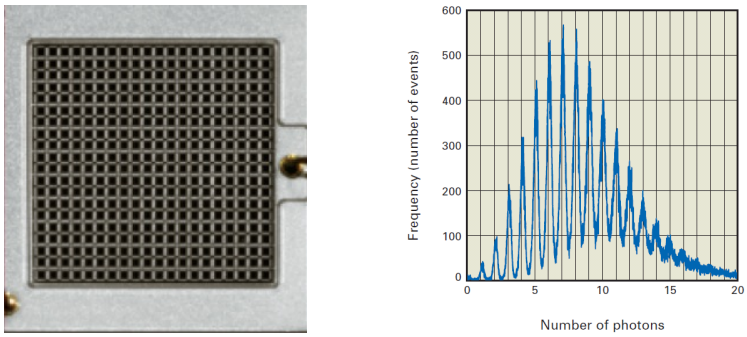


Figure 2.9 – (left) picture of a 1x1 mm Hamamatsu GM-APD pixel with 20x20 micro cells, (right) Pulse height spectrum using a Hamamatsu S 10362-11-025U. Figure from [29].

Because GM-APD are separated into micro cells, there are dead spaces between the micro cells within one GM-APD. Therefor, instead of using the term QE for each detector channel, GM-APD are usually characterized by the Photon Detection Efficiency (PDE). The PDE is a measurement of what percentage of the incident photons was detected, taking into account the geometric fill factor (i.e. percentage of a pixel that is actually sensitive) and the avalanche initiation probability:

$$PDE = QE \times f_{geom} \times P_A \tag{2.4}$$

where QE is the quantum efficiency, f_{geom} is the geometric fill factor, P_A is the Avalanche initiation probability. The latter two are defined as

$$f_{geom} = \frac{Effective\,Cell\,Size}{Total\,Pixel\,Size} \tag{2.5}$$

$$P_A = \frac{Number\,of\,excited\,pixels}{Number\,of\,pixels\,hit} \tag{2.6}$$

When an avalanche occurs in one micro cell, it is dead during some time. This period is called the recovery time and is typically of the order of a few tens of ns. During that period no other impinging photons can be detected in that

cell. Hence, the number of triggered pixels has a nonlinear relationship with the number of incident photons:

$$N_{fired} = N_{total} \times \left[1 - exp\left(\frac{-N_{photon} \times PDE}{N_{total}}\right)\right] \tag{2.7}$$

where N_{fired} is the number pixels that fired, N_{total} is the total number of micro cells in a GM-APD pixel and N_{photon} is the number of incident scintillation photons.

Increasing the number of micro cells also increases the total dead space in the GM-APD pixel and hence reduces f_{geom}. For a 1600 micro cells/mm^2 Hamamastu GM-APD, the PDE is 20%, while for 400 micro cells/mm^2 it is 35%. This is significantly lower compared to the QE of an ordinary APD. Therefore the size of the micro cells is a trade-off between the PDE and the linearity at high light fluxes.

Finally, it should also be noted that because GM-APD are APD-like detectors, they can also be used in magnetic fields and allow compact detector geometries.

2.3 Detector performance

The quality of a PET image is determined by a number of factors such as the fundamental resolution limits discussed in section 1.2, the detector components and image reconstruction algorithms. In this section, we will discuss several key parameters defining the performance of PET detectors and their impact on PET imaging. These parameters are energy resolution, sensitivity, spatial resolution and time resolution.

2.3.1 Energy resolution

The energy resolution tells us something about how precisely the detector measures the energy of the gamma ray. Usually one wants only the events in the photo peak, corresponding to gamma rays that deposited all their energy. This ensures that these gamma rays were not scattered within the patient body and thus resulting in wrong LORs. It is usually accomplished by imposing an energy threshold on the detection of a impinging photon. Only when the measured energy is higher than this threshold, the event will be accepted. A finite energy resolution results in a loss of 'good events' when the energy threshold is put too high because some of the photons that deposited their full energy will have mea-

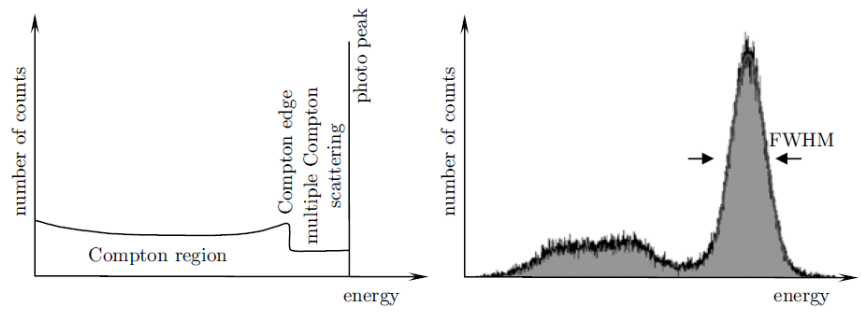


Figure 2.10 – (left) Energy spectrum of an ideal detector, (right) Spectrum of 511 keV photons measured with a $20 \times 10 \times 10$ mm LSO scintillator block coupled to a Hamamatsu S8550 APD array. Figure from [40].

sured energy below the threshold level. Hence, a good energy resolution (i.e. a narrower photo peak) allows to more efficiently remove incident photons that did not deposit 511 keV in the detector.

When a 511 keV photon interacts with an ideal detector, a spectrum as depicted in figure 2.10 left will be generated. The *photo peak*, corresponding to the full energy of the impinging photons, is a perfect δ-peak. Events in this peak have either undergone a single photo-electric absorption, or one or more Compton interactions followed by a photo-electric absorption. In both cases, the gamma ray deposits all its energy into the detector. Events that experience only one Compton scattering and then leave the detector are in the lower energy part of the spectrum called the *Compton region*. The upper boundary of the Compton region shows up as a sharp edge called the *Compton edge*. This is because during Compton scattering (see Appendix A) only a part of the gamma ray energy is transferred via the recoil electron to the detector. In case a 511 keV gamma scatters over an angle θ, the amount of the energy E_{re} that the recoil electron obtains is given by (2.8):

$$E_{re} = 511keV - E_{sc}\ with\ E_{sc} = \frac{511keV}{2 - cos\theta} \tag{2.8}$$

where E_{sc} is the remaining energy of the photon after being scattered. Since the scattering angle of the gamma photon can range from 0° to 180°, the energy of the recoil electron can vary between 0 keV and 340.7 keV, i.e. the gamma photon cannot lose all its energy in a single Compton interaction. The energy lost during a 180° backscatter determines the position of the Compton edge in figure 2.10.

The valley between the Compton edge and the photo peak contains events that have undergone multiple Compton scatterings but no photo electric absorption, i.e they left the detector carrying away some residual energy.

In a real detector, one can never obtain the ideal spectrum but rather something shown in figure 2.10 right. In this spectrum, the photo peak is widened due to the finite energy resolution of the detector. The energy resolution is defined as the relative FWHM of the photo peak, i.e.

$$\triangle E(\%) = \frac{FWHM}{E} \tag{2.9}$$

The FWHM for a Gaussian distribution is 2.35 times standard deviation.

The energy resolution is worsened by three effects: the statistics of the light generation in scintillator; the detector related contributions (e.g. QE, CE, ENF) and readout electronic noise.

Scintillator contribution

The amount of light generated in the scintillator when a gamma photon deposits its energy is not constant. In addition, the amount of scintillation light does not always vary exactly linearly with the amount of energy put in the crystal (figure 2.11). Hence the total amount of light produced for a given total amount of energy deposited depends on how this energy was deposited. Also, non-uniformity of the scintillation crystal and reflective coatings make the amount of light exiting the crystal depended on the position of the gamma photon interaction.

Let us assume that the amount of scintillation photons generated N_o is drawn from a distribution with mean $\overline{N}_o$ and variance $\sigma^2_{N_o}$. These optical photons then have some probability η that they impinge onto the photo detector array. Therefore, the amount of scintillation photons N_γ impinging on the photo detector array is generated by a binomial selection of a random process. The mean $\overline{N}_\gamma$ and variance $\sigma^2_{N_\gamma}$ are then [5]:

$$\overline{N}_\gamma = \eta \overline{N}_o \tag{2.10}$$

$$\sigma^2_{N_\gamma} = \eta \overline{N}_o + \eta^2(\sigma^2_{N_o} - \overline{N}_o) \tag{2.11}$$

Since the energy resolution is measured by summing the signals from all photo detector pixels together, the probability for the binomial selection η can be approximated by the total solid angle subtended by the photo detectors as seen from

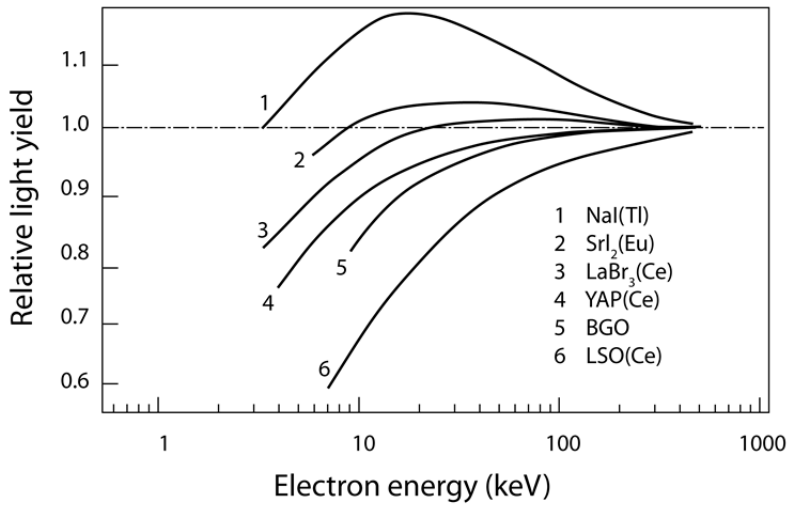


Figure 2.11 – The non-linear light yield behavior of scintillators. Data for this figure are taken from [57].

the source. In addition it also includes effects due to the transport of the optical photons in the scintillation crystal.

One interesting fact from (2.11) is that if η is small, $\sigma^2_{N_\gamma} \approx \eta \overline{N}_o$, i.e. a binomial selection with a small probability of a random process is a Poisson process. This observation will be used later in section 3.2.1 and section 5.2.1. However, for the discussion on the energy resolution where we look at the light collected by the complete photo detector array, η is too big to be ignored.

Detector related noise contribution

The conversion of the N_γ scintillation photons into charge carriers (i.e. photo electrons in case of a PMT or electron-hole pairs in case of an APD/GM-APD) and the amplification of those charge carriers introduces extra uncertainties on the measured signal amplitude. The QE and CE both define probabilities involved in converting the scintillation photons into charge carriers and the collection of these charge carriers for further amplification. Hence, the number of primary charge carriers N_{cc} is again the result of a binomial selection process. The mean and variance on the resulting distribution is given by:

$$\overline{N}_{cc} = QE \times CE \times \overline{N}_{\gamma} \tag{2.12}$$

$$\sigma^2_{N_{cc}} = QE \times CE \times (1 - QE \times CE) \times \overline{N}_{\gamma} + (QE \times CE)^2 \times \sigma^2_{N_{\gamma}} \tag{2.13}$$

The gain G of the photo detectors introduces further fluctuations. After multiplication by an average gain factor $\overline{G}$, the mean of the signal g emerging from the photo detector is then given by

$$\overline{g} = \overline{N}_{cc} \times \overline{G} = \overline{N}_{\gamma} \times QE \times CE \times \overline{G} \tag{2.14}$$

while the variance on the photo detector signal is given by [5]:

$$\sigma^2_g = \overline{G}^2 \times ((1 + \frac{\sigma^2_G}{\overline{G}^2}) \times \overline{N}_{cc} + \sigma^2_{N_{cc}} - \overline{N}_{cc}) \tag{2.15}$$

$$= \overline{G}^2 \times QE \times CE \times \overline{N}_{\gamma} \times \left(ENF + QE \times CE \times \left(\frac{\sigma^2_{N_{\gamma}}}{\overline{N}_{\gamma}} - 1 \right) \right) \tag{2.16}$$

where $ENF = 1 + \frac{\sigma^2_G}{\overline{G}^2}$ is the excess noise factor.

Additional electronic read-out noise contribution

Electronic noise generates an additional spread on the final signal emerging from the detector pixel which can be modeled as an additive noise term with zero mean and variance σ^2_E. The σ_E is usually expressed as an equivalent noise charge (ENC) referred to the input of the preamplifier reading out the photo detector. So, the final energy resolution can be written as

$$\triangle E = \frac{2.35 \times \sigma}{E} \tag{2.17}$$

$$\frac{\sigma}{E} = \frac{\sqrt{\sigma^2_g + \sigma^2_E}}{\overline{g}}$$

$$= \frac{\sqrt{\overline{G}^2 \times QE \times CE \times \overline{N}_{\gamma} \times \left(ENF + QE \times CE \times \left(\frac{\sigma^2_{N_{\gamma}}}{\overline{N}_{\gamma}} - 1 \right) \right) + ENC^2}}{\overline{G} \times QE \times CE \times \overline{N}_{\gamma}}$$

$$= \sqrt{\left(\frac{\sigma^2_{N_{\gamma}}}{\overline{N}^2_{\gamma}} - \frac{1}{\overline{N}_{\gamma}} \right) + \frac{ENF}{QE \times CE \times \overline{N}_{\gamma}} + \left(\frac{ENC}{\overline{G} \times QE \times CE \times \overline{N}_{\gamma}} \right)^2}$$

The first part under the square root is the influence of the scintillator, while the second part is the effect due to the photo detector and the last part is due to the electronic noise. Since the gain G from a PMT or GM-APD is very large (10^6) compared to the electronic noice, the last term can therefore be ignored in those two detectors. Further discussion and the value of those parameters in (2.17) used in our setup will be given in Chapter 5.

2.3.2 Sensitivity

The image quality also depends on the amount of coincidence events collected. Injecting a higher tracer dose to the patient or prolonging the acquisition time will obviously result in the detection of more coincidences. Nevertheless, these two ways are limited in practice due to patient comfort and safety. Therefore, system sensitivity is an essential factor for a high quality image. Sensitivity is defined as the ratio of the number of detected coincidences over the number of emitted photons pairs for a given tracer distribution. It tells us how efficiently a scanner can detect the 511 keV photon pairs emitted from the patient.

The scanner sensitivity η can be divided in a number of components: the 511 keV gamma ray detection efficiency ε, the solid angle coverage of the detectors $\frac{\Omega}{4\pi}$, the time and energy windows applied. The sensitivity can hence be expressed as:

$$\eta = \Phi\varepsilon^2 \frac{\Omega}{4\pi}\varphi \tag{2.18}$$

where Φ is the probability that coincident events are detected within a time window and the energy window (typically set at 350-650 keV) given the source emission rate. The detection efficiency ε of an individual detector is defined as

$$\varepsilon = 1 - e^{-\frac{d}{\lambda}} \tag{2.19}$$

where λ is the attenuation length of the 511 keV gamma ray in the detector material, d is the thickness of the detector. According to the Beer-Lambert law, $e^{-\frac{d}{\lambda}}$ is the probability that a gamma ray can pass though a distance d without interaction. The detection efficiency ε is squared as a result of the coincidence detection. A thicker detector therefore leads to a better sensitivity for the same detector material.

To enlarge the solid angle coverage $\frac{\Omega}{4\pi}$ of the detectors, PET scanner are most often designed as multiple rings of detector modules surrounding the patient.

Inbetween the detector modules, there are usually wedge shaped gaps, resulting in a loss of efficiency. The detector modules themselves can be made up of continuous or pixellated crystals. In the later case, an additional factor, called the packing fraction φ, needs to be added in (2.18), because there are small gaps for reflective material in between the pixelated detector elements. This fraction is about 10% for a modem PET scanner [43].

2.3.3 Spatial resolution

The aim of a PET scanner is to measure the decay positions of the tracer. The accuracy of the final PET image depends on the spatial resolution of the detectors and is usually expressed in terms of FWHM and FWTM of the point spread function (PSF), i.e. the response in the image to a perfect point source. The system spatial resolution of PET consists of several components:

1. fundamental spatial resolution limits due to the positron range and non-collinearity of the annihilation gamma rays (see section 1.2)

2. error on the determination of the LOR positions due to the scanner geometry, detector properties and the algorithm being used to determine the position from the detector signals

3. smoothing and possible artifacts introduced during the tomographic image reconstruction process.

Depending on the purpose, a different range of spatial resolutions are required from a PET scanner. A scanner for small animal desires a system resolution below 1.5 mm, while a brain or breast PET usually needs 2-4 mm resolution and for a whole body human PET, 4-6 mm resolution is needed. Here in this thesis, we will only discuss the spatial resolution component introduced by the detector.

Intrinsic and coincidence resolutions

The contribution of the detector is often defined in terms of intrinsic resolution. The intrinsic resolution is the accuracy with which the interaction point of a 511 keV gamma on a detector can be determined. If the detector is not capable to measure the interaction depth (DOI), then the interaction point is always placed at a fixed depth. In this case, the DOI is usually fixed at a depth equal to the mean free path length of a 511 keV gamma in the scintillator material.

The positioning error is given by the minimum distance between the LOR, defined by the positions where two annihilation photons are localized, and the true electron-positron annihilation point. We will call this error the coincidence resolution to distinguish it from the intrinsic detector resolution.

The distance Δx between the LOR and the true annihilation point (i.e. coincidence error) is related to the errors $\triangle x_1$ and $\triangle x_2$ on the interaction point of both 511 keV photons and the distances d_1 and d_2 from the source to the detectors (figure 2.12).

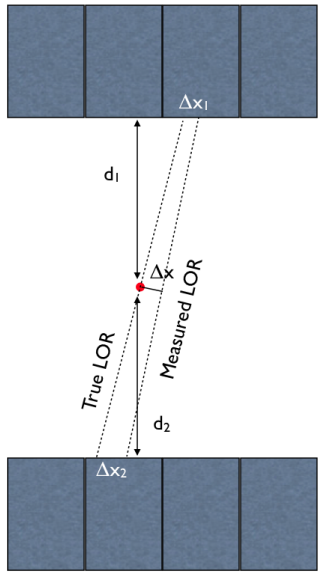


Figure 2.12 – The coincidence error $\triangle x$ not only depends on the 511 keV photon position estimation errors $\triangle x_1$ and $\triangle x_2$, but also on the distances d_1 and d_2 of the annihilation point to the detectors

The probability $P(\Delta x)$ to obtain a coincidence error Δx for an annihilation point in the center of the field of view (FOV) is given by

$$P_{coinc}(\Delta x) = \int_{-\infty}^{+\infty} d(\Delta x_1) d(\Delta x_2)\, P_{intr}(\Delta x_1) P_{intr}(\Delta x_2) \delta(\Delta x - \frac{\Delta x_1 + \Delta x_2}{2}) \tag{2.20}$$

$$= 2 \int_{-\infty}^{+\infty} d(\Delta x_2)\, P_{intr}(2\Delta x - \Delta x_2)\, P_{intr}(\Delta x_2) \tag{2.21}$$

where $P(\Delta x_1)$ and $P(\Delta x_2)$ are the probabilities that the two annihilation photons were localized with an intrinsic error of respectively Δx_1 and Δx_2. The Dirac delta function in the last term of the integrand in equation (2.20) gives the relationship between the intrinsic errors measured by the two detectors and the coincidence error at the center, i.e. at the center the coincidence error is the mean of the intrinsic errors. If the annihilation point was not in the center, this term has to be changed accordingly. If we rewrite equation (2.20) as (2.21), the coincidence error distribution can be seen as two times the convolution of the intrinsic error distribution with itself evaluated at a value of $2\Delta x$, i.e.

$$P_{coinc}(\Delta x) = 2(P_{intr} \otimes P_{intr})(2\Delta x) \tag{2.22}$$

Parallax error

When two annihilation photons, originating from a source close to the center of the scanner, are detected by two radially opposite detectors, the maximum uncertainty is determined by the crystal size (detector pair A-B in figure 2.13 left). The depth at which the two gamma photons interacted in the crystal doesn't matter when the photons imping perpendicular. The inability to determine the DOI of a gamma photon creates an additional contribution to the coincidence resolution when the incidence angle is not perpendicular (detector pair C-D in figure 2.13 left). This happens when the annihilation occurred outside the center of the FOV. The origin of this parallax error is due to the significant penetration power of the gamma photons into the scintillator. When only the (X,Y) coordinate of the interaction point can be determined, any point along the line in the Z-direction (i.e. radially away from the scanner center) passing through the estimated interaction point could be considered as a starting point of the LOR. At non-perpendicular incidence, the coincidence resolution is degraded as the incidence angle increases, i.e. as the radial distance of the annihilation point to the center increases. In other words, the resolution of a scanner is optimal in the center and degrades towards the edges. This phenomena is called the parallax error.

If the detector can determine the DOI with some level of accuracy, the segment of the line along the Z direction containing possible interaction points is shortened. This decreases the parallax component to the coincidence resolution (detector pair C-D in figure 2.13 right).

Unfortunately most detectors fail to provide full three-dimensional information about the interaction position within the scintillation crystal.

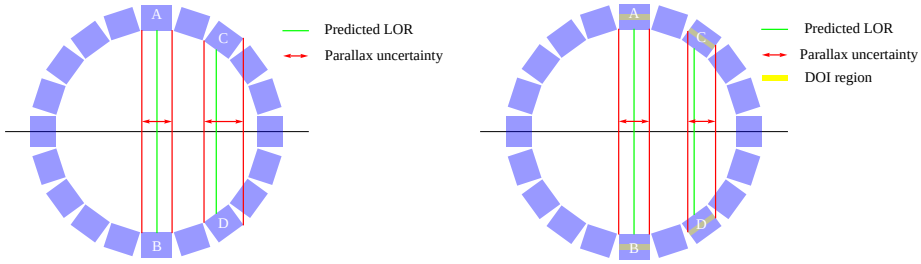


Figure 2.13 – (left) Parallax error if no DOI measurement is available, (right) Parallax error with DOI measurement.

2.3.4 Time resolution & time of flight PET

Time resolution is the fluctuation on the time difference between the moment that the 511keV gamma ray interacts with the scintillator and the time the impinging event is recorded by the data acquisition of the detector. It depends on the rise time of the scintillation light, the jitter on the transit time of the charge carriers in the photodetector, as well as the signal to noise ratio of the amplified signal emerging from the detector. The time resolution determines the minimal length of the coincidence time window that can be chosen without significant loss of coincidence events.

A good time resolution can also improve the signal to noise ratio in the reconstructed image. If the information on the time difference between the arrivals of the two coincidence photons is not used, each location along the LOR is attributed with the same probability of having emitted the pair (figure 2.14 left).

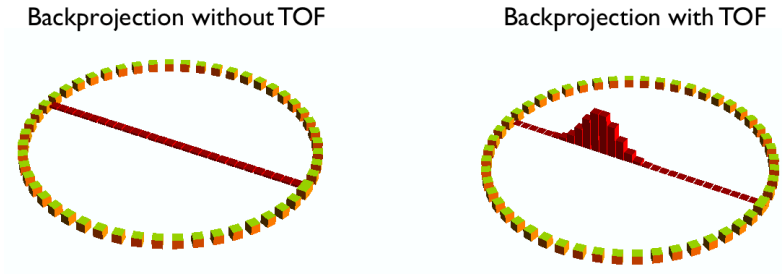


Figure 2.14 – Illustration of TOF PET : (left) uniform-probability weighting of annihilation site in standard PET, (right) use of TOF information to constrain the location of annihilation site during image reconstruction.

In theory, if a PET scanner would be able to record the precise interaction times of the photons involved in a coincidence, the annihilation point along the LOR can be deduced exactly from the time difference Δt and speed of the light in the object c. Hence no image reconstruction would be required any more. The location of the annihilation event presented by the distance to the midpoint between the two detectors, can be calculated as:

$$\Delta d = \frac{\Delta t \times c}{2} \tag{2.23}$$

Suppose we want to achieve a spatial resolution of about 5 mm without the need for image reconstruction, then the photon arrival times would have to be recorded with a precision of approximately 33 ps. This is currently impossible to achieve using present technology.

If the PET system has a time resolution that can record the time difference on the arrival of the 2 annihilation photons with reasonable accuracy, then this can help to reduce the signal-to-noise ratio (SNR) in the reconstructed image by given each location along the LOR a probability of containing the annihilation point (figure 2.14 right). The probability is Gaussian distributed around a position determined by the measured time difference and a standard deviation equal to the time resolution. This is called "time-of-flight" PET or TOF-PET. The improvement of the SNR of TOF-PET images compared to non-TOF-PET images is given by [15]:

$$SNR_{TOF} \approx \sqrt{\frac{2D}{c\triangle t}} \times SNR_{non-TOF} \tag{2.24}$$

where D is the diameter of the object being imaged, c is the speed of light in the object, and $\triangle t$ is the time resolution of the system. From (2.24) if follows that the SNR enhancement is higher for thicker patients. Instead of improving the image SNR, one can also use TOF PET to keep the same SNR while either shortening the scan time or injecting a lower dose.

In the early 1990s', TOF became a popular research topic a first time. But at that time, the only existing fast scintillator crystal was BaF_2. However, it emitted only 15% of its scintillation light with a fast decay time of 0.8 ns. In addition, this crystal has a low Z. This means that to have enough sensitivity, very long crystals are needed, and that will introduce large parallax errors.

In recent years, TOF-PET became more practical due to the discovery of fast scintillators with good stopping power such as LSO and LYSO, and the

development of speedy electronics. A recent commercial LYSO TOF-PET system (Philips Gemini TF) has achieved a 600 ps time resolution, i.e. the uncertainty on the location of the annihilation point along the LOR is reduced to about 9 cm. The newer GM-APD photodetector also seem very promising for TOF-PET. Recent research [72] has shown that a time resolution of 100 ps can be achieved with $3 \times 3mm^2$ Hamamatsu GM-APD S10362-33-050c coupled to a $3 \times 3 \times 5mm^3$ $LaBr_3 : Ce5\%$ crystal, and 172 ps FWHM when a $3 \times 3 \times 5mm^3$ LSO crystals is read out with same photodetector. As for monolithic crystals, because the scintillation light has to be shared over multiple photo detectors, the signal-to-noise ratio for each photodetector is lower than for pixelated crystal. However, a promising time resolution of 225 ps FWHM for a monolithic crystal with size $16.2 \times 18 \times 10mm^3 LaBr_3 : Ce5\%$ crystal coupled to a Hamamatsu GM-APD S11064-050P (X1) 4×4 array, each pixel size of $3 \times 3mm^2$, has been reported recently by [72].

2.3.5 Requirement of PET scanners

Based on the target application area, each kind of PET scanner requires some different kind of performance [44].

For a pre-clinical (e.g. small animal) PET, it is important to have a very high spatial resolution due to the size of the small animal. For instance, it would need $700\mu m$ spatial resolution for mouse imaging to achieve the same image quality as we currently obtain in human bodies [44]. Very high time resolution is currently not essential since for such a small diameter scanner, no adequately useful information can be provided by TOF.

A whole-body human PET is different from pre-clinical PET. Current systems typically have discontinuous scintillators with pixel sizes of $4 \times 4mm^2$. The dose limitation and big radius of the detector ring, makes that each pixel detects a limited amount of the gamma rays. Trying to reduce the pixels size will reduce the statistics for the detectors even further. Therefore the SNR will become too bad to have a good quality image. Thus, the most interesting development of the whole body system is not on spatial resolution but on TOF.

A Brain-PET (or other specialized imaging systems) normally requires both high spatial resolution ($< 3mm^3$) and very good sensitivity, because many brain-PET system need to produce good dynamic imaging. One of the strategies to improve the sensitivity is to make the ring smaller. However this also introduces

more parallax errors which require DOI measurement capabilities in the system. So far, commercial brain PET systems rarely provide DOI information [44]. The reason is that the existing methods that will be discussed in section 2.4.1 are not cost-effective to produce.

2.4 Pixelated and monolithic scintillator detectors

The detector modules for PET can be divided in two broad categories: those that make use of pixelated scintillators and those that use undivided (or monolithic) scintillators.

2.4.1 Pixelated detectors

Localization principle

Most clinical PETs use discrete, pixelated crystals. In an ideal pixelated detector module, each small individual crystal is coupled one-to-one to a corresponding photodetector. In this case, the center of the crystal pixel that produced a signal is used as the position where that photon was detected, i.e. no positioning algorithm is needed. From a practical point of view, such a solution is not viable in large systems since it would require a huge number of small photo detectors to read out all the scintillator pixels. It would too expensive and too complicated. For this reason, the PET block detectors were introduced (figure 2.15).

These detector modules consist of only few big PMTs (typically 4) which read out the scintillation light of an array of smaller detector elements. The array of pixels are created by making partial cuts with varying lengths into a block of scintillator material. The pixels are separated by reflective material. The depth of the saw cuts is empirically done in order to share scintillation light over all four PMTs. The four PMT signals can then be used to determine which individual pixel has generated the light due to a gamma ray interaction. In practice, this will be done in two steps. First, the X and Y coordinates are calculated for each gamma ray interaction based on:

$$X = \frac{S_A + S_B - S_C - S_D}{S_A + S_B + S_C + S_D} \tag{2.25}$$

and

$$Y = \frac{S_A + S_C - S_B - S_D}{S_A + S_B + S_C + S_D} \tag{2.26}$$

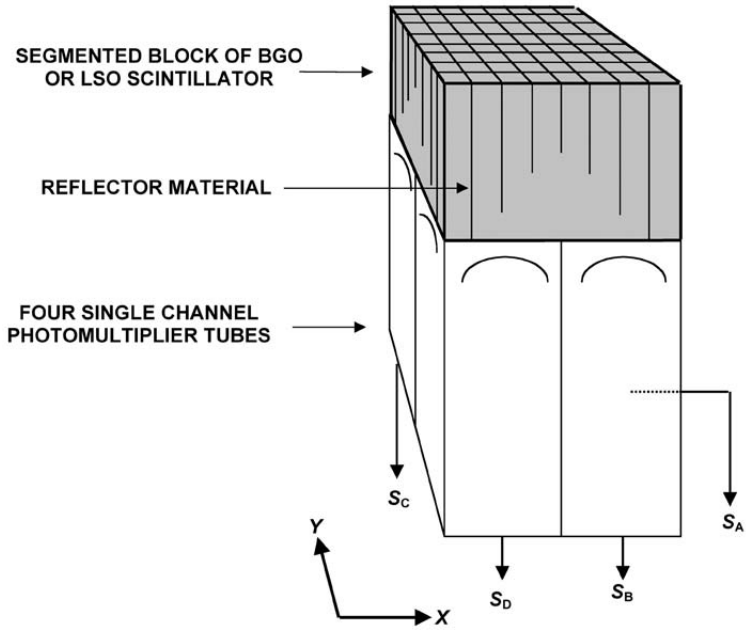


Figure 2.15 – Block detector. Figure from [59].

where S_A, S_B, S_C and S_D are the four PMT signals as shown in figure 2.15. These X and Y position are not the coordinates assigned to the gamma ray interaction point, but somehow reflect which small crystal generated the light. This relationship is not completely linear across the whole detector surface. Therefore, an empirical look up table is created to relate the (X, Y) to a given individual pixel. Figure 2.16 shows an examples of such a look up table for a 8*6 crystal array.

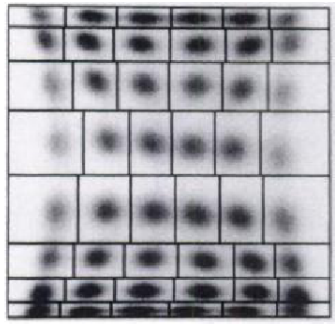


Figure 2.16 – LUT relating the computed (x,y) position with a scintillator pixel

If the position falls near the boundary of an area in the look up table, it is possible that the event will be attributed to the wrong pixel due to statistical fluctuations. This is called a coding error in the block detector. If the gamma photon underwent a Compton interaction, and is subsequently absorbed in a photo electric interaction, two light sources are produced. Since the light from the two pixels can spread once it passed the end of the cuts, it will be interpreted by the position algorithm as one event with a position somewhere inbetween the two interactions. This obviously also results in a wrong position.

Intrinsic and coincidence resolution

The pixelated crystal detectors only have discrete position information, i.e. the (X,Y) coordinate is always fixed to the center of the pixels that was selected using the look up table. The maximum intrinsic positioning error is thus $w/2$, where w is the width of the individual pixel. The probability $P(\Delta x_1)$ to obtain a given intrinsic error Δx_1 is uniformly distributed over the face of the pixel, i.e.

$$\begin{aligned} P(\Delta x_{1,2}) &= \frac{1}{w} \quad if\ |x_{1,2}| < \frac{w}{2} \\ &= 0 \quad if\ |x_{1,2}| > \frac{w}{2} \end{aligned} \tag{2.27}$$

Using (2.22) and (2.27), the coincidence resolution can be computed for an annihilation point in the central of FOV:

$$\begin{aligned} P_{coinc}(\Delta x) &= \frac{2}{w^2}\left[w - |2\Delta x|\right] \ if\ |\Delta x| < \frac{w}{2} \\ &= 0 \qquad if\ |\Delta x| > \frac{w}{2} \end{aligned} \tag{2.28}$$

Equation (2.28) shows that the coincidence error distribution at the center of the FOV has a triangular shape with a FWHM equal to half of the pixel width w (in the middle of figure 2.17). It can be shown that as the source moves closer to one of the detector, the coincidence error distribution gets a trapezoidal shape and the FWHM also increases. When the source position is close to one of the detector, the response function becomes a rectangular function, and the FWHM is equal to the size of the detector w (figure 2.17).

Parallax correction

In recent years, many possible approaches have been proposed to solve the parallax problem in pixelated detectors. These solutions need to add extra hardware in

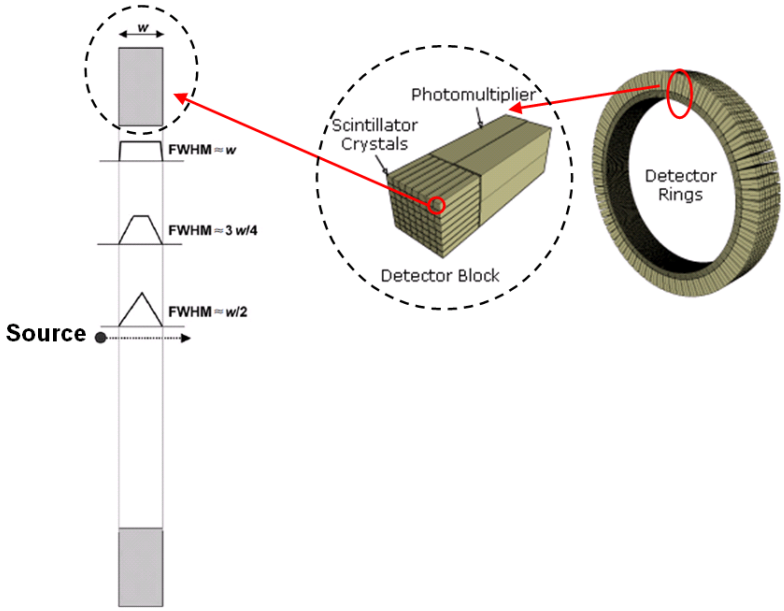


Figure 2.17 – Coincidence point spread function for a system based on individual scintillator pixels. A point source is moved between two detectors. When the source is at center of the FOV, the coincidence point spread function has a triangle shape with a FWHM equal to half the pixel detector size w. For off-center emission points, the shape becomes trapezoidal and ultimately the FWHM is the same as w.

order to provide DOI information. One method, known as a phoswich, uses two layers of scintillator materials, where each of them can be distinguished by their different decay times (the top of figure 2.18). Therefore, by looking at the decay time of the pulses, a discontinuous two-level (top half and bottom half) depth encoding is available. The DOI resolution is, however, limited by the thickness of the crystal layer.

Another method, to determine the DOI is to use a second photo detector at the top end (the bottom of figure 2.18). For long narrow pixelated crystal, the photons generated by gamma ray interaction will reflect many times on the edges before they reach an exit surface at the top or the bottom. At each reflection, there is some probability for the light to leave the crystal. In addition, photons which travel a longer distance will have a higher chance to be reabsorbed. Hence,

the ratio of the amount of detected scintillation photons at the top and bottom reflects the ratio of the distance from the interaction point to both ends and hence the DOI can be estimated. The drawback of this method is that it increases the cost and complexity of the scanner by doubling the number of photo detectors and associated electronics. In addition, bulky PMTs can't be used since the photo detector on the top side of the pixel array blocks the incoming 511 keV gammas. Hence, only APD or GM-APD detectors are possible candidates.

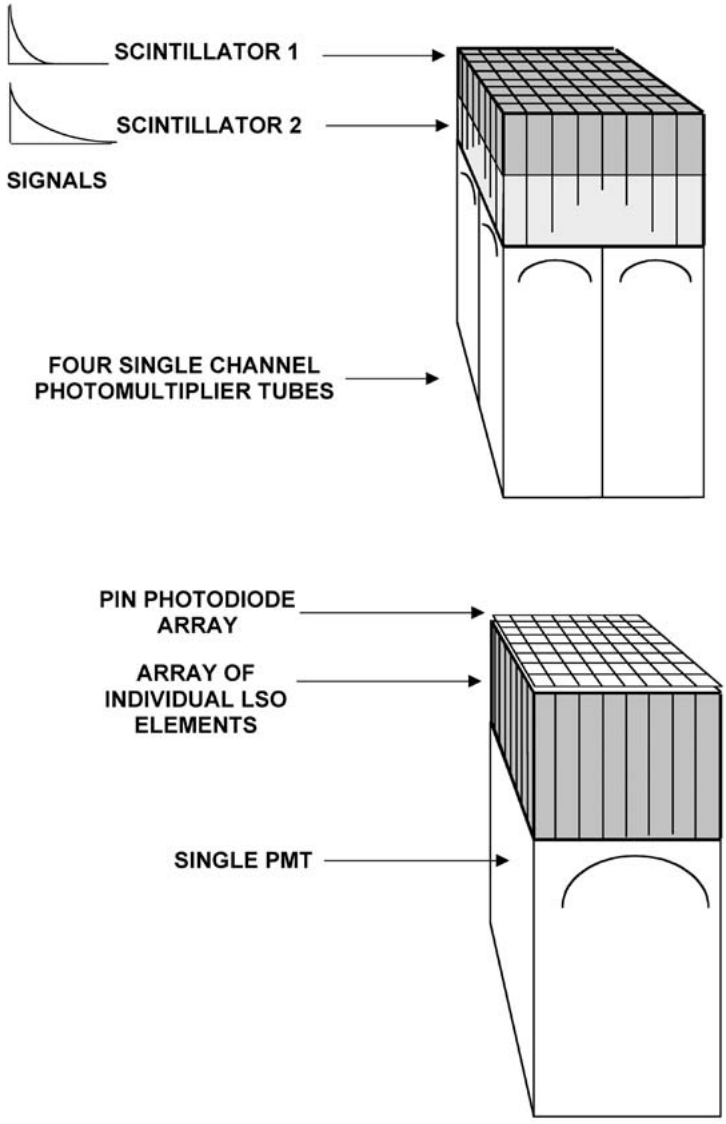


Figure 2.18 – (top) Phoswhich method to obtain a 1 bit DOI information, (bottom) the use of a second photo detector matrix on the top side of the crystal matrix to determine the DOI based on the ratio of the light measured on both sides. Figure from [59].

2.4.2 Monolithic scintillator detectors

Localization principle

Monolithic scintillator based detectors [34, 36, 47, 69, 71, 74] consist of an undivided block of scintillator material that is read out by an array of photo detector pixels (e.g. multi anode PMT, APD array or GM-APD array). The difference with pixelated detectors is that the scintillation photons can travel completely free inside the scintillator block. As a result, the shape of the scintillation light distribution measured by the photo detector array is related to the interaction position (figure 2.19). The center of the light distribution is related to the (X,Y) position whereas the width of the light distribution carries some information about the interaction depth of the gamma photon.

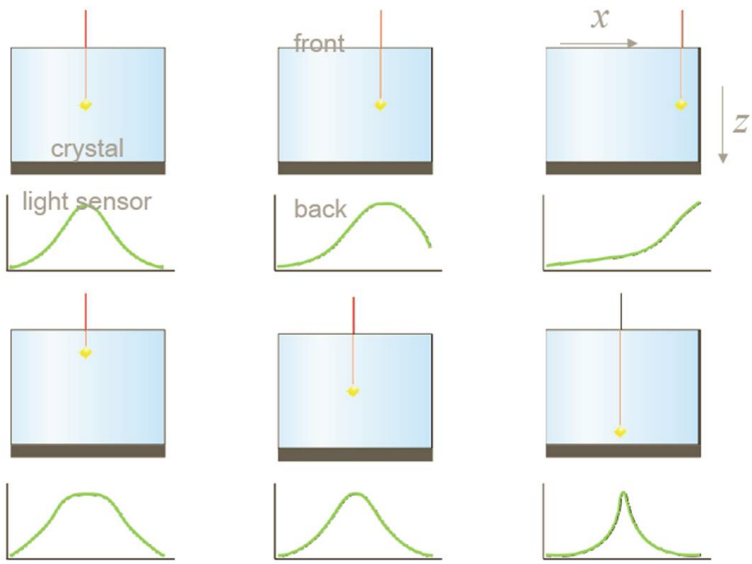


Figure 2.19 – The interaction position of a gamma photon in a monolithic scintillator block is related the shape of the scintillation light distribution measured at an exit surface. Figure from [74].

Contrary to the discrete case where the intrinsic resolution is determined by the pixel size, it is the position algorithm (i.e. how to extract the position information from the array of photodetector signals) and the signal to noise ratio that determine the achievable spatial resolution. In addition, the estimated coordinate are continuous in nature. Finally, the DOI can be provided using no extra

hardware since the information is already encoded in the measurements.

The position algorithm for intrinsic spatial resolution in monolithic crystal is the central topic of this thesis, which will be discussed in the following chapters.

Coincidence response function

We can again use (2.22) to find the distribution of the coincidence errors. If we assume that the intrinsic error distribution is Gaussian distributed with a $\sigma_{intr} = {FWHM_{intr}}/{2.35}$, then it follows that the coincidence error distribution at the center of the FOV is again Gaussian distributed with a standard deviation $\sigma_{coinc} = {\sigma_{intr}}/{\sqrt{2}}$. Similar like discrete detectors in 2.4.1, the coincidence resolution increases as the emission point source moves away from the FOV center, as shown in figure 2.20.

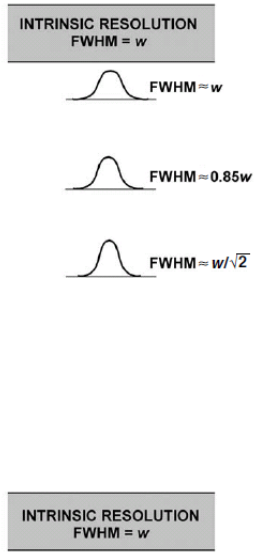


Figure 2.20 – For a system based on monolithic scintillators, the coincidence point spread functions are Gaussian shaped. In the center of the FOV, the coincidence resolution is a factor $\sqrt{2}$ smaller than the intrinsic detector resolution. Again, the coincidence resolution increases as the emission point source moves away from the center. Figure from [59].

Parallax error correction

In continuous crystals the shape of the detected light signal has a relationship with the DOI. Contrary to the pixelated detectors, no extra measurement hardware is necessary. This make the use of monolithic scintillator blocks easier and cheaper to implement. On the other hand, the algorithms that need to extract the DOI from the measured signals are more complicated. This will also be detailed in the following chapters.

Chapter 3

Position Estimation in Monolithic Crystals

In order to use monolithic scintillator crystals in a PET detector module, one needs a way to extract the interaction position from the scintillation light distribution as it exits the block and is detected by a photo-sensitive detector array. A first requirement is that the pixels in the photo detector array sample the 2D light distribution with enough precision, i.e. the pixel size and pitch should be small enough. To this extend, one is usually limited to the position sensitive photo detectors available on the market. Pixellated photo detectors with pixels of 2x2 mm or smaller can now be easily found.

Numerous methods to obtain spatial information from a monolithic scintillator block have been presented in the past, starting with the simple 2D centroid method, i.e Anger logic [2]. However, the bias and resolution of the estimator increase a lot at the edge of the crystal. To overcome this problem, statistical estimation of the interaction position using 2D Maximum Likelihood (ML) was first introduced by [28] and developed further by [6]. Since 511keV gamma rays have a rather high penetrating power, the crystal is usually made thick and therefore additional knowledge about the coordinate of the interaction point along the third dimension (i.e. depth-of-interaction or DOI information) is required to correct for parallax errors (see section 2.3.3 on parallax error).

Some groups have made an effort to obtain DOI information in monolithic crystal by e.g. adding an analog resistor network to determine the spread of the light, which reflects the DOI [41], or by separating the crystal into two to four layers to achieve discrete DOI information [52]. Other groups tried to find the

3D interaction position without using additional DOI measurements, but only use the signals of the detected light themselves to find DOI directly or indirectly, e.g. using an ML approach to localize the 3D interaction position [34, 47] or using artificial neural networks (ANN) trained to determinate the 2D surface entry position in the scintillator block [11]. The latter doesn't require DOI information since the DOI coordinate is fixed to the surface.

This chapter will give a brief introduction to some of the algorithms which only use the embedded information from the signals of the detected light themselves, i.e. which don't require additional hardware. Figure 3.1 shows an overview of those algorithms. The algorithms can be divided in two main categories: those which need some prior training (blue) and those which extract all the necessary information immediately from the measured signals themselves (red). A direct comparison of their performance is difficult since it also depends on the size and type of scintillation crystal and on the photo sensitive detector. In the literature, the described algorithms were usually applied to one specific combination of crystal size and photo detector. For the sake of argument, we will therefore assume that the photo detector consists of an 8x8 array of pixels found in many position sensitive PMTs or combination of APD arrays.

The algorithm based on a simplified physical model (yellow) will be detailed in Chapter 4 since it is the main subject of this thesis. The other numbered algorithms in figure 3.1 will be introduced in this chapter.

3.1 Anger logic

A classic positioning method in monolithic crystals is the so-called Anger logic[2] (No. 8 in figure 3.1). This method is a 2D-positioning algorithm that computes a weighted center-of-gravity (COG) of the measured light distribution, i.e.

$$\begin{aligned} \hat{x} &= \frac{\sum_{m=1}^{64}(w_m^x \times g_m)}{\sum_{m=1}^{64} g_m} \\ \hat{y} &= \frac{\sum_{m=1}^{64}(w_m^y \times g_m)}{\sum_{m=1}^{64} g_m} \end{aligned} \tag{3.1}$$

where $\hat{x}$ and $\hat{y}$ are the estimated position in X and Y direction, w_m^x and w_m^y are fixed detector dependent weights for the m^{th} pixel, and g_m are the signal values from the corresponding pixels. A trivial choice for the weights w_m^x and w_m^y are the coordinates x_m and y_m of the m^{th} pixel.

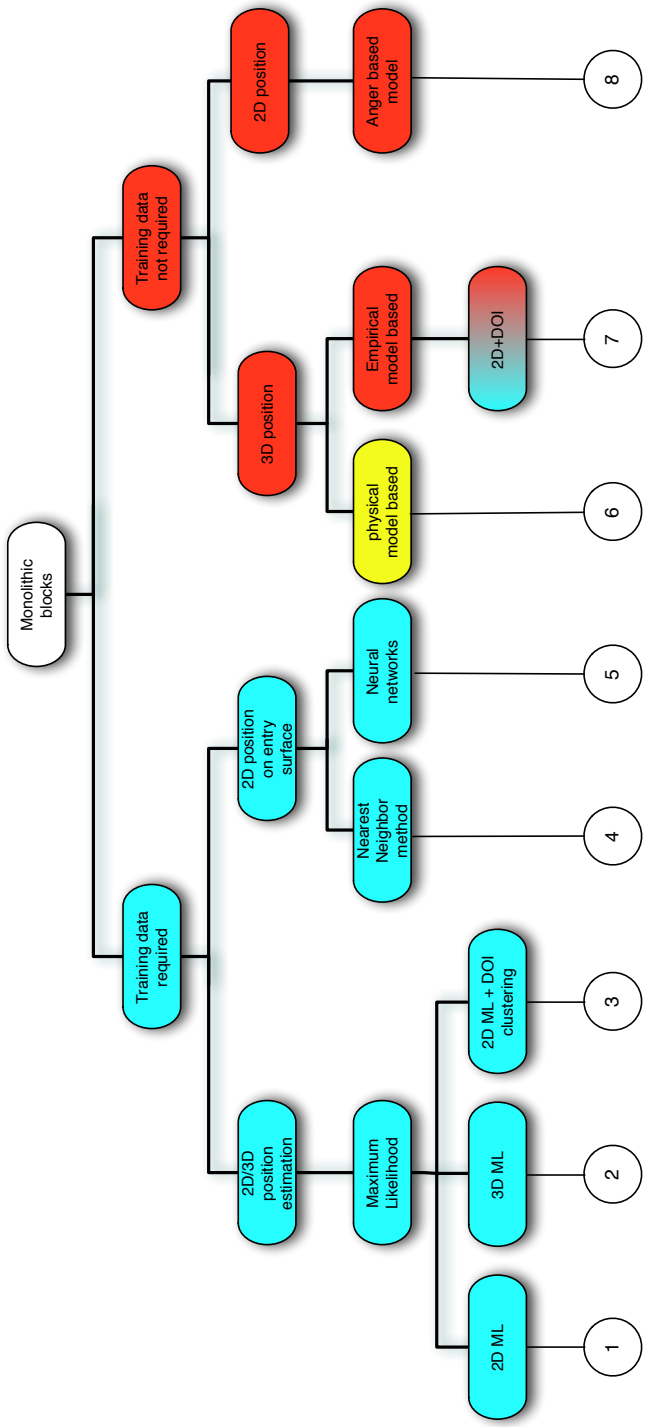


Figure 3.1 – Schematic diagram showing the different algorithms to find the position of an impinging photon using the embedded information from the signals.

The Anger logic can be computed fast and is easily implemented in analog hardware. One of the primary drawbacks of the basic Anger algorithm is the fact that it suffers from large non-linearity effects when approaching the edge of the monolithic crystal. As a result, photons impinging near the edge cannot be localized properly and should therefore not be used. This lowers the sensitivity of the PET system.

Another problem with the classic Anger logic is that it treats all signals equally. In practice, pixels far away from the interaction point of the 511 keV photon in the monolithic scintillator received only a very small fraction of the light. The signal-to-noise ratio (SNR) of the electrical pulses produced by these pixels is rather low. Some signals might even be dominated by noise. Treating all the signals equally results in worse spatial resolution. Using a threshold preamplifier with each of detector pixels can improve the situation, i.e. only signals above the threshold value will contribute to the position calculation. In addition the weight of pixels producing large signals can be increased by using nonlinear amplifiers.

Finally, the last major limitation of the Anger logic based positioning method, is that it cannot supply any DOI information. Hence, Anger logic is incapable of doing parallax correction.

3.2 Maximum likelihood algorithms

A number of research groups have been developing algorithms that take the statistical nature of the measured photo detector signal into account by using Maximum Likelihood (ML) based algorithms [5, 4, 33, 34, 45, 46]. We start by introducing the 2D-ML method [4] (No. 1 in figure 3.1), where it is assumed that the detector is thin and thus the response as a function of depth does not significantly vary.

3.2.1 2D Maximum likelihood position estimation

It was pointed out in section 2.3.1 that the light generated in a scintillator from a gamma interaction is a non-Poisson process. For events which deposited an energy E, the amount of light produced $N_o(E)$ is distributed with a mean $\overline{N}_o(E)$ and variance $\sigma^2_{N_o}(E)$. Each of the photons then faces a binomial selection process: it can either reach the m^{th} photodetector or not. Suppose the fraction of optical photons emitted from position $\mathbf{R}(x, y)$ that will reach the m^{th} photodetector is $\eta_m(\mathbf{R})$, then the amount of scintillation photons N_γ impinging on the m^{th} photo detector

$$N_\gamma(E) = N_o(E) \times \eta_m(\mathbf{R}) \tag{3.2}$$

is a Poisson random variable, due to the fact that a low-efficiency binomial selection of a non-Poisson random variable is a Poisson random variable, as it was shown in (2.11).

Once a photon enters the m^{th} photo detector, it will produce a photo electron (or electron-hole pair in case of an APD) with a probability given by the quantum efficiency QE. This process is Poisson process because a Poisson random variable after a binomial selection is still a Poisson random variable [5] with mean

$$\overline{N}_m(\mathbf{R}, E) = QE \cdot \eta_m(\mathbf{R}) \cdot \overline{N}_o(E) \tag{3.3}$$

The probability of generating N_m photo electrons in the m^{th} photodetector will hence be [5]:

$$pr(N_m|\mathbf{R}, E) = \frac{\left[\overline{N}_m(\mathbf{R}, E)\right]^{N_m}}{N^m!} exp\left[-\overline{N}_m(\mathbf{R}, E)\right] \tag{3.4}$$

Assuming the number of photo electrons generated in each detector are independent to each other, the multivariate probability of the number of photo electrons, $\{N_m, m = 1, \cdots, M\}$ in all M photo detectors is given by

$$pr(\{N_m\}|\mathbf{R}, E) = \prod_{m=1}^{M} \frac{\left[\overline{N}_m(\mathbf{R}, E)\right]^{N_m}}{N^m!} exp\left[-\overline{N}_m(\mathbf{R}, E)\right] \tag{3.5}$$

These photo electrons in a PMT (or electron-hole pairs in an APD) will then be multiplied by the gain of the photo detector. If we ignore the noise generated in the photo detector (ENF) and readout electronics (ENC), which is an acceptable assumption for PMT detectors, the distribution of the observed signals $\mathbf{g} = \{g_m, m = 1 \cdots M\}$ of each photo detector can be seen as a multivariate Poisson distribution:

$$pr(\mathbf{g}|\mathbf{R}, E) = \prod_{m=1}^{M} \frac{\left[\overline{N}_m(\mathbf{R}, E)\right]^{N_m}}{N^m!} exp\left[-\overline{N}_m(\mathbf{R}, E)\right] \tag{3.6}$$

where

$$N_m \equiv u\left({}^{g_m}/_{G_m}\right), \; and \, \overline{N}_m \equiv {}^{\langle g_m(\mathbf{R},E)\rangle}/_{G_m} \tag{3.7}$$

$\langle \cdots \rangle$ is an ensemble average, and $u(\cdots)$ is a rounding operator, and G_m is the channel gain. If, on the other hand, the amplifier and electronic noise are non-negligible, e.g in case of an APD detector, then an independently distributed (i.d.) normal model instead of a Poisson model is expected for each output signal g_m [47].

To characterize $pr(\mathbf{g}|\mathbf{R}, E)$, we need to determine G_m and $\overline{N}_m$. Calibrating the gains G_m is a separate task. Assuming the gains G_m have been calibrated, (3.6) then only depends on the $\overline{N}_m$.

Calibration of $\overline{N}_m$

The mean of each Poisson distribution $\overline{N}_m$ for the m^{th} photon detector and for a gamma ray impinging position (x, y) is obtained from by calibration data and stored in a look-up table in order to characterize the response of the detector for that position. To fully characterize the complete detector, the procedure is repeated for many positions distributed over the surface of the block. This set of positions is usually referred to as calibration positions. The mean response of each photo detector pixel for gamma impact positions that were not in the calibration set can be generated by interpolation of the look-up table. Hence the look up table should be populated with a sufficient number of calibration positions covering the whole crystal in order to have a good accuracy of interpolation.

Estimation of unknown events

When a new gamma photon impinges at an unknown position, one only knows the set of 64 output signals g_m from the detector. Given this set of measured pixel responses and the scintillator calibration characteristics $\overline{N}_m$, it is possible to compute the likelihood (using (3.6)) that they were generated by an interaction at any given position (x, y).

The position (x, y) that is most likely to have generated the measured response (i.e. the maximum value for the likelihood or maximum likelihood estimate), is then assumed to be the best estimate $(\hat{x}, \hat{y})$ for the true 2D impact position, i.e.

$$(\hat{x}, \hat{y})_{ML} = \underset{(x,y)}{arg\,max}\{pr(\mathbf{g}|\mathbf{R}, E)\} \tag{3.8}$$

However, calculating the likelihood for every position recorded in the look-up table and some interpolated positions and finding the maximum value among those likelihoods, is extremely time consuming. To avoid this exhaustive search

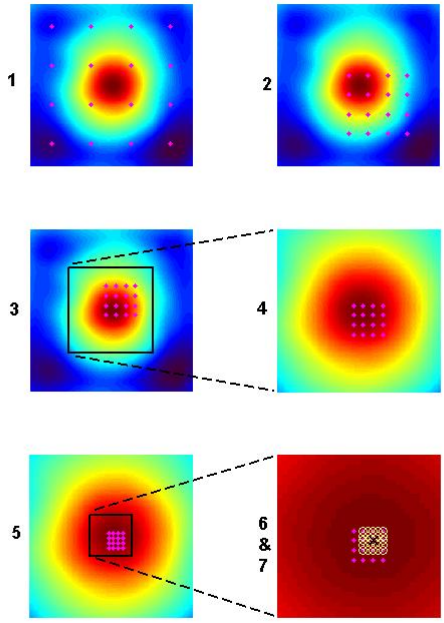


Figure 3.2 – 2D-ML contracting-grid search. Figure from [33].

procedure, a method called contracting-grid search is sometimes used to find an ML estimate [33]. Figure 3.2 shows an example of searching in 2 dimensions. In each iteration, the likelihoods of 4x4 positions are calculated. The position with the biggest likelihood is selected and a new set of 4x4 positions is chosen in a smaller area around this position. This procedure is then repeated a number of times. Assuming that e.g. 6 iterations are performed to achieve a good accuracy [24], then 16x6=96 likelihood calculations are required.

3.2.2 3D-ML position estimation

Multiple interactions within the crystal will introduce uncertainties on the signals. This influence can be ignored for thin crystals in the 2D-ML method. Since a 511 keV gamma ray is a rather high energy photon, the scintillation crystal used to detect it needs to be thick. The Compton scattered events might need to be modeled by additional parameters, which makes the model more complicated.

One way to deal with these events is use hypothesis testing to reject them [4]. Therefore, the 3D-ML method deals mostly with single interactions, multiple-interaction events should be filtered first. A method to perform this filtering is detailed in [34].

For a thick scintillator, the distribution of the scintillation photons detected by the pixellated photo detector also depends on the depth of interaction or DOI. Hence the distribution of the signals measured by a given pixel when 511 keV gammas are impinging at a fixed 2D entry position (x,y) is a continuous sum of Poisson distributions with a mean varying according to the depth of the gamma ray interaction. To estimate an interaction point including DOI, (3.6) is still valid but with $\mathbf{R}(x, y, z)$ as a function of three dimensions.

One way to present the relationship of how $\overline{N}_m$ changes as a function of depth of interaction is to choose an analytic form, such as a cubic polynomial [34]:

$$\overline{N}_m(\mathbf{R}, E, \Theta) = \sum_{n=0}^{3} c_{mn}(x, y) \cdot z^n \tag{3.9}$$

The parameters Θ describe the set of coefficients $c_{mn}(x, y)$ as a function of lateral position. Hence, for each of the given calibration positions (x, y), 4 parameters of the 3^{rd} order polynomial need to be determined per pixel. Interpolation is once more needed for un-calibrated position.

In reality, the depth of interaction (z) cannot be obtained directly. Therefore prior knowledge of pathlength distribution is used to overcome this problem. Figure 3.3 shows a gamma ray entering a monolithic block at known position $R_0(x_0, y_0)$ and along a known direction s. The pathlength l is ideally expected to be exponentially distributed. However, due to the fact that events are prefilterd to remove the multiple-interaction events in the scintillator [34], the filter is not perfect and the detector geometry etc. also influence the distribution, it is no longer exponentially. One has to rely on the simulation to determine the $pr(l|E, R_0, s)$ [34].

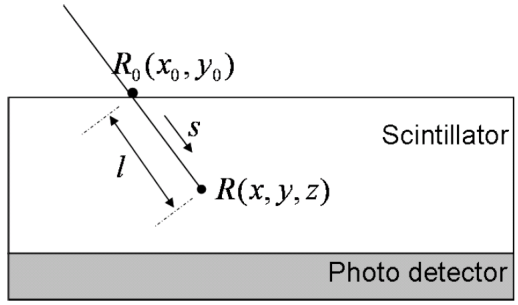


Figure 3.3 – A cross-sectional view of an obliquely incident gamma ray into a thick scintillation camera.

To obtain Θ using a given data set $\mathbf{\Gamma}$, which contains J events, one can resort again to maximum likelihood using the prior knowledge of the pathlength distribution. The estimated Θ is then given by

$$\hat{\Theta}_{ML} = \underset{\Theta}{argmax} \left\{ pr\left(\mathbf{\Gamma}|\mathbf{\Theta}\right) \right\} \tag{3.10}$$

where ([34, 33])

$$pr(\mathbf{\Gamma}|\mathbf{\Theta}) = \prod_{j=1}^{J} \int_{0}^{\infty} dl \, pr(l|E_j, R_{0j}, s_j) \prod_{m=1}^{M} pr\left[\mathbf{g}|\overline{N}_m \left(R_{0j} + ls_j, E_j, \Theta\right)\right] \tag{3.11}$$

After calibrating the parameter set Θ and storing it into a look up table, determining the interaction position of a new event is similar to the procedure used in the 2D-ML case.

To calibrate the 4 times larger number of parameters, an enormous reference data set is needed. In addition, the 3D positioning algorithm now also needs to calculate 4^3 times instead of 4^2 likelihood functions in each iteration.

3.2.3 2D-ML+DOI clustering position estimation

Another group implemented a 2D-ML +DOI clustering position algorithm (No. 3 in figure 3.1) based on Gaussian signal distributions instead of Poisson. The idea of DOI clustering is that the light distribution varies according to DOI, i.e. events interacting in the same (x, y) position and in a similar DOI region will generate alike light distribution patterns. For each of the DOI region in one calibration

position (x, y) , two look-up tables for the mean and standard deviation of each m^{th} photon detector will be built. This algorithm therefore yields discontinues DOI information. A flow chart of this algorithm for 7 DOI regions (interpolated to 15 DOI regions) is shown in figure 3.4. More details can be found in [45, 47].

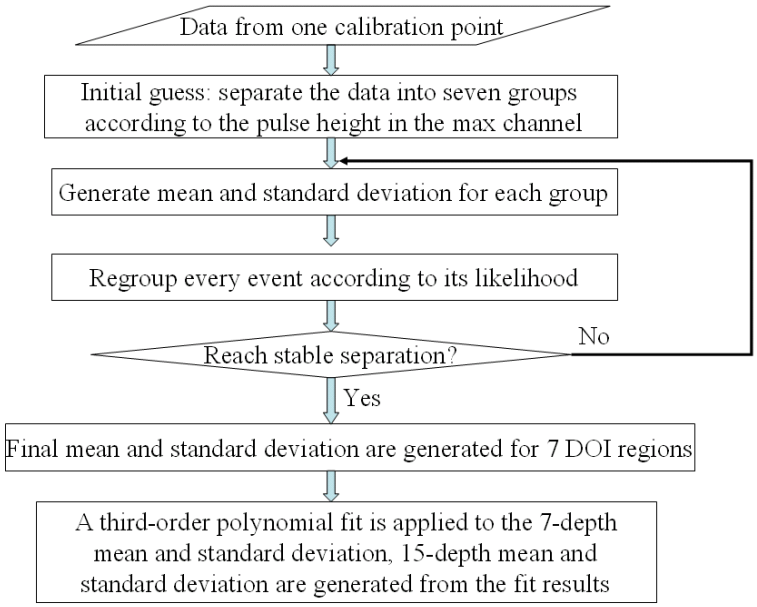


Figure 3.4 – Flow chart for 2D-ML+DOI clustering algorithm. From [45].

3.3 2D Entry Point Estimation

Instead of determining the 3D position (x, y, z) of an interaction point within the crystal, some groups use an alternative way by extracting the 2D entry point (x_0, y_0) when an annihilation photon enters the front surface of the crystal. As shown in figure 3.5 (left), an LOR connecting from the entry point A on the surface of a crystal is identical to an LOR connecting from the 3D interaction points B within the crystal. Hence, knowing the entrance position can also be used to correct for parallax error. However, there is a price to be paid for this apparent reduction in the number of coordinates to determine. As shown in figure 3.5 (right), photons can enter the crystal at the same point A, but generated different

light distributions because they interact at different points B_1 or B_2. Therefore, the relationship between the measured light distribution and the incidence position A also depends on the incidence angle θ and ψ . Because PET is based on the detection of two co-linear photons by a ring of detectors, this information is readily available to some extend. The incidence angle θ_1, ψ_1 and θ_2, ψ_2 of the two photons on the crystal blocks can be approximately determined by the line joining the center of those two crystal blocks. Hence, for each block pair combination that can yield valid coincidences in a PET scanner, a single 2D entry positioning algorithm can be used. If the size of the crystal is small compared to the radius of the PET ring, the difference between the true incidence angle and the approximated incidence angle is small. For example, if the blocks have a surface area or 2x2 cm and the PET system has a ring diameter of 40 cm (i.e a brain PET system), the maximum difference is 2.86^o.

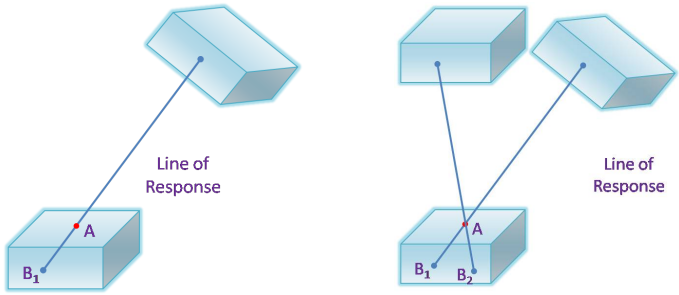


Figure 3.5 – An LOR can be defined by a pair of 3D coordinates of the interaction position (e.g. B_1or B_2) or by a pair of 2D coordinates of the points where the photons entered the crystal (point A).

3.3.1 L-nearest neighbor algorithm

One way to get the 2D entry point is by using a statistical nearest-neighbor algorithm [64] (No. 4 in figure 3.1). Reference data sets are collected for different surface entry points (x_0, y_0) and incidence angle θ (i.e. crystal block combination). At each position and angle, the light distributions **g** of a number of reference events are recorded in a large look up table. Before they are stored, the pixel values are normalized to make the light distributions energy independent, i.e.

$$g_m^{norm} = \frac{g_m}{\sum g_m} \tag{3.12}$$

The difference with the look up tables used in the ML-based methods is that the values of the 64 pixel signals are stored for each recorded event. Previously, only the mean (or the 4 parameters used for the cubic polynomial parameterization of the DOI dependent mean) of the signal distribution was stored for each detector pixel. Hence, the look up tables in the nearest neighbor algorithm are much larger.

When an annihilation photon enters the crystal at an unknown position, the reference data set corresponding to its incidence angle is selected. Then a least squares comparison is made between the set of 64 signal values of the unknown event and the signal sets of all the reference events in the selected reference data set. A subset containing L reference events producing the smallest least squares differences is retained. The most frequently occurring entry point (x_0, y_0) within this subset, is then taken as the best estimate $(\hat{x}_0, \hat{y}_0)$ of the impact position of the unknown event.

Although this method does not need to calibrate signal distribution parameters such as mean or variance, it still needs a large amount of reference data to produce good statistical estimates of the incidence position. Consequently, determining the position of an unknown event is slow because it requires a huge amount of least-squares comparisons.

3.3.2 Artificial neural network

Artificial neural networks or ANNs is another method that is used to estimate the 2D entry point (x_0, y_0) [10] (No. 5 in figure 3.1). This method is not simply comparing an unknown data set with reference data as in previous section. An ANN represents a *black box function*, which expresses the relationship between the measured photodetector signals $\mathbf{g}$ (ANN inputs) and the 2D entry point (x_0, y_0), (ANN output) for a certain incidence angle ϑ. This black box function is composed of so-called "neurons", acting as non-linear transfer functions. These neurons, divided over an input layer, one or more hidden layers and an output layer, are connected by links which carry a weight. When a value travels from one neuron to the next over a link, it is multiplied by the link weight. The weighted values arriving at a neuron over all the links connected to its input are summed and offset by a bias value that is a property of the receiving neuron, i.e. if the l^{th} neuron in a given layer is linked to K neurons in the previous layer, its input is given by

$$I_l = \sum_{k=1}^{K} w_{k,l} \cdot O_k + b_l \tag{3.13}$$

where $w_{k,l}$ is the link weight between the k^{th} neuron in the first layer and the neuron l, O_k is the output from the kthneuron and b_l is the bias of neuron l. The output of (3.13) is then passed through a nonlinear function such as e.g. a sigmoid transfer function:

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{3.14}$$

The resulting value is then sent over a new link to the next layer of neurons (figure 3.6). This process is executed for all neurons in the ANN except for the 'output' neuron where the nonlinear function in (3.14) is not applied. In case of the ANN with N inputs, and 2 hidden layers with resp. K and L neurons, the signal from the neuron in the output layer is given by

$$\sum_{l=1}^{L} w_{lo} \cdot \sigma \left(\sum_{k=1}^{K} w_{kl} \cdot \sigma \left(\sum_{n=1}^{N} w_{nk} \cdot I_n + b_k \right) + b_l \right) + b_o \tag{3.15}$$

where w_{nk} are the link weights from the n^{th} input neuron to the k^{th} neuron in the first hidden layer, $b_k$is the bias of the k^{th} neuron in the first hidden layer, w_{kl} are the link weights from the k^{th} neuron in the first hidden layer to the l^{th} neuron in the second hidden layer, b_l is the bias of the l^{th} neuron in the second layer, w_{lo} are the link weights from the l^{th} neuron in the second hidden layer to the output neuron and $b_o$is the bias of the output neuron.

Before an ANN can be used, it has to undergo supervised training to find the values of the link weights and biases. During the training process, the neural network is presented with a sufficiently large set of inputs and the corresponding outputs it should produce. The link weights and neuron offsets are adjusted such that the root mean square error (RMSE) between the produced outputs and the true outputs is minimal, i.e.

$$(w_{trained}, b_{trained}) = \underset{(w,b)}{argmin} \sqrt{\frac{\sum_{i=1}^{N} (ANN(input|w,b) - output)^2}{N}} \tag{3.16}$$

Similar to the L-nearest neighbor method, the pixel values (i.e. inputs to the neural network) are first normalized for energy using (3.12).

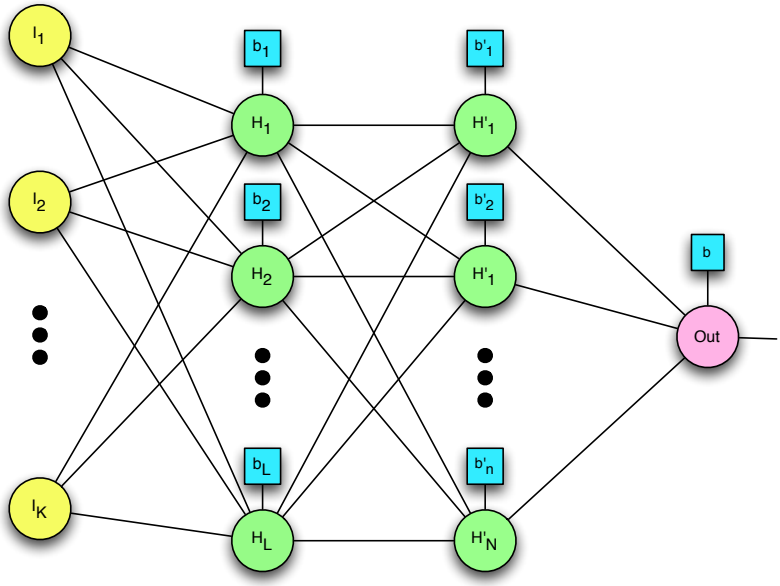


Figure 3.6 – Structure of a neural network with K input neurons (yellow), 2 hidden layers with respectively L and N neurons (green) and one output neuron (pink). Values are injected in the network through the input neurons and propagated through the network over the links connecting the neurons in the different layers. While they travel along a link, a scaling factor is applied. The inputs of the neurons sum the weighted values and add a neuron dependent bias. This weighted sum is then non-linearly transformed by a transfer function and the result is passed on to the next layer. The output neuron has usually a linear transfer function.

Usually, two separate Neural Networks are build: one for the X and one for the Y position. To further reduce the complexity of the ANN, the signals in the rows and columns of the photo detector array are summed. The rational behind this, is that changes in the x-coordinate are mainly reflected in the distribution of the signals over the different columns, while the signals in the different rows contain most of the information about the y interaction coordinate (figure 3.7). If the summing of the row and column signals is done in the amplifiers chip reading out the photo detector array, the number of electronic channels to digitize and process is also significantly reduced. In case of an 8x8 photo detector array, the number of channels is then reduced from 8x8=64 to 8+8=16, i.e. 4 times less electronic channels.

Similar to the nearest neighbor algorithm, a pair of ANNs needs to be trained for every crystal block combination in the PET system. The advantage however,

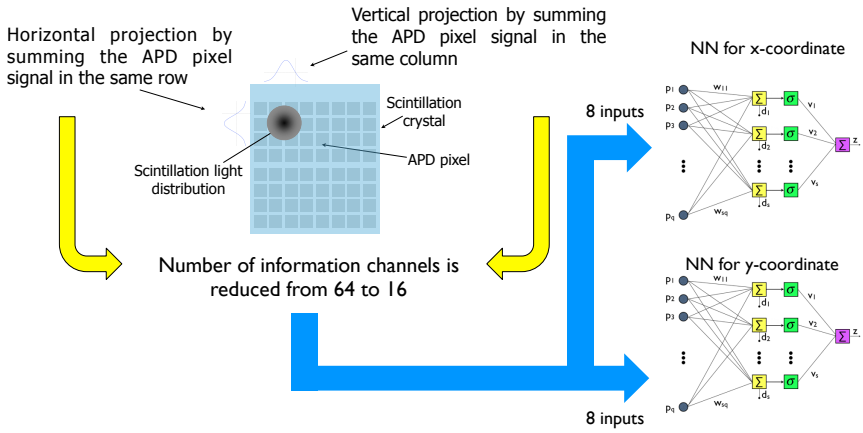


Figure 3.7 – To simplify the readout electronics and the topology of the ANNs, the rows and columns of the photo detector array can be summed. The summed columns are used as inputs for the ANN trained to compute the x-coordinate, while the summed rows are used as the inputs for the ANN trained to computed the y-coordinate.

is that once the ANN are trained, they can very quickly compute the incidence position of an unknown event. Processing speeds of up to 25 million events per second have been achieved on an average computer. If needed, neural networks can also be implemented in hardware using FPGA's.

3.4 Parametric modeling method

All the methods mentioned above, except for the simple centroid method, require the acquisition of reference data sets for calibration, comparison or training. This can cause a problem if the measurement conditions have changed since the time the reference data sets were collected. This could be because of e.g. changes in the gain of photo detector pixels due to temperature or high voltage fluctuations. To overcome this problem, one could try to extract the position directly from the measured scintillation light distribution without prior calibration.

One research group developed an empirical parametric model that is used to extract the 2D interaction position by fitting it to the measured light distribution

(No. 7 in figure 3.1). The model they used is given by [46]:

$$\overline{N}_m(x, y, a, b, \gamma) = a\left(1 + \frac{(x - x_i)^2 + (y - y_i)^2}{b^2\gamma}\right)^{-\gamma} \tag{3.17}$$

where $\overline{N}_m$ is the same as in (3.3), which is the mean of the Poisson variable representing the number of photon electrons generated by the m^{th} detector. x_i *and* y_i are the coordinates of the i^{th} pixel in the photo detector array. The parameter *'a'* relates to the peak height of the light distribution, while *'b'* depends on the width of the light distribution. The shape of the distribution can be changed using the parameter γ. If $\gamma \to 1$ the distribution becomes a Cauchy distribution and if $\gamma \to \infty$, the model will converge to the Gaussian shape [46].

Comparing to the ML method in section 3.2.1 which uses a look up table to represent all $\overline{N}_m$, this method uses a general function to describe the $\overline{N}_m$. Therefore no calibration in $x - y$ direction needs to be done.

The model is fitted to the light distribution using either MLE fitting or weighted least square fitting. Once the 2D estimated position $(\hat{x}, \hat{y})$ is obtained using the set of best fit parameters (a,b,γ), the DOI still needs to be determined separately because it was not directly included in the parametric model. In general, the parameters a, b and γ are a function of the DOI. This fact is used to set up an empirical relation between the DOI and the parameters a,b and γ [46]:

$$DOI(a, b, \gamma) = \beta_0 + \beta_1 a + \beta_2 b^2 + \beta_3 \gamma \tag{3.18}$$

Before equation 3.18 can be used, the coefficients β_0, β_1, β_2 and β_3 should be determined. This still has to be done through experimental calibration by shooting 511 keV events at a 45 degree angle onto the monolithic block. The DOI position of an interaction was obtained indirectly. It is equal to the difference $\triangle X$ between the coordinate x_0 where it enters the LSO block and the coordinate x where it interacts (figure 3.8). For each event, model 3.17 is fitted to the light distribution. The best-fit parameters (a,b,γ) are then recorded together with the approximate DOI position $\triangle X$. When a sufficient number of events are recorded, the coefficients β_0, β_1, β_2 and β_3 are obtained by fitting (3.18) to the measurement data.

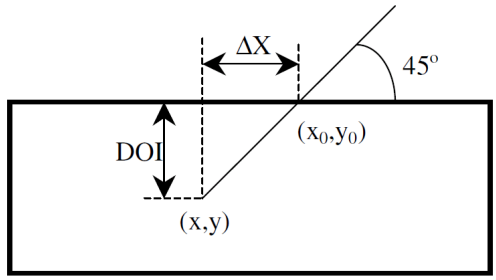


Figure 3.8 – When a photon impinges at a 45^o angle, the DOI is equal to the difference between the interaction X-coordinate and the X-coordinate where it entered the monolithic scintillator.

Because (3.17) is shift-invariant, it cannot take any edge effects (such as reflections) into account. Therefore, the monolithic scintillator blocks used in this method are painted black on the sides to minimize the edge effects. However, one of the consequences is that this will worsen the energy resolution due to lower scintillation photon statistics.

Chapter 4

NLS Modeling Algorithm

In this chapter, we will model the light distribution corresponding to a gamma ray interaction and use a nonlinear least square (NLS) method to estimate its position. The goal is to obtain a high spatial resolution in X, Y and DOI direction, only depending on the information embedded in the signals of individual events without any prior training or calibration.

The introduction of regression and Least Square fittings will be given first in this chapter, followed by the relation to Maximum Likelihood, Weighted Least Squares and Least Squares. The description of the problem and the way it has been modeled are in section 4.3 to 4.5. The minimization methods to estimate the model parameters and the confidence intervals of the estimated parameters will be discussed in the end of the chapter.

The evaluations of the method using simulations data and experiments will be shown in Chapter 5. Comparisons with Cramér-Rao lower bound, Maximum Likelihood and Neural Network Method will be performed in Chapter 6.

4.1 Introduction to regression and least square

Regression is the analysis of the relationship between a dependent variable y and one or more independent variables x. This relationship is expressed by a "model" with an adjustable set of parameters θ. Regression analysis can be used for prediction of the dependent variables or for parameter estimation. The target of prediction is to find a model to forecast future events based on known past events, i.e. to predict the dependent variable before they are measured. In this case, the goal is not to determine the best-fit values of parameters that define a model,

but to generate a curve that one can use to interpolate unknown values. As for parameter estimation, the parameters in the model that are of interest cannot be measured directly. They can, however, be estimated from the measurable independent and dependent variables. If the unknown relation between these variables can't be expressed in a linear way, then a nonlinear model should be used.

One strategy to find the model's best-fit parameters $\hat{\theta}$ is to use a least square method, which minimizes the sum of the squared distances between the measured points (dependent values) and the values of the model at the corresponding independent variable values, i.e.:

$$\hat{\theta} = \underset{\theta}{argmin} \sum_{i=1}^{N} (y_i - y(x_i|\theta))^2 \tag{4.1}$$

where (x_i, y_i) are the measurement data values, and where $y(x_i)$ is generated from the model evaluated using the parameters θ. N represents the number of measurements.

To guarantee that a least squares approach works well, there are a set of assumptions that the data should obey [55]:

1. The independent variable x is known exactly. Only the dependent variable y is perturbed with additive, zero mean noise n_{y_i}:

   $$y_i = y(x_i|\theta) + n_{y_i} \tag{4.2}$$

   If both x and y are perturbed with errors, the problem becomes Total Least Squares [32].

2. The additive perturbation y follows a known distribution, which is usually assumed to be zero-mean Gaussian.

3. The scatters (the variance) all the way along the curve are assumed uniform, which is called "homoskedasticity", i.e. the "errors" have constant standard deviation σ and are independent of x.

4. The observations are independent.

Assumption 3 is not always valid in practice. The standard deviation of the errors can vary along the graph, which is called "heteroskedastic". In this situation,

an inefficient[1] result will be obtained by minimizing the sum-of-squares because points with larger deviations from the curve will have a bigger impact on the sum-of-squares value. To treat all points equally, it is necessary to add an appropriate weight w_i to the different error terms in the sum-of-squares, i.e.

$$\hat{\theta} = \underset{\theta}{argmin} \sum_{i=1}^{N} w_i \left(y_i - y(x_i|\theta)\right)^2 \tag{4.3}$$

This is called a Weighted Least Squares (WLS). It is not always easy to choose the weights. If the noise variance σ_i^2 at each measurement i is known, and all measurements are independent, then it is common to use $w_i = \frac{1}{\sigma_i^2}$. Hence, the WLS becomes

$$\hat{\theta} = \underset{\theta}{argmin} \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \left(y_i - y(x_i|\theta)\right)^2 \tag{4.4}$$

If the perturbation are correlated, then the weights becomes $w_i = C^{-1}$, where $C = cov(n_y)$ is the covariance matrix of the measurement noise. However, a covariance matrix can not present the full stochastic characterization of the noise distortions, while the probability density functions (pdf) can. This brings us to the Maximum likelihood method.

The Maximum likelihood method is determined using the joined pdf for the data $(y_1, \cdots y_N)$ given the parameter set θ: $Pr\left(y_1, \cdots y_N|\theta\right)$. This pdf tells us how probable it is that the data set $\{y_1, \cdots, y_N\}$ is observed, given a set of parameter values. In reality, we are facing the inverse problem. Given some observed data, find the set of parameters in the model which have most likely generated this data set. This inverse function is called the likelihood function $\mathcal{L}\left(\theta|y_1, \cdots y_N\right) = Pr\left(\theta|y_1, \cdots y_N\right)$. The estimation of the parameters is (Bayes rule [60]):

$$\hat{\theta} = \underset{\theta}{argmax}\, \mathcal{L}\left(\theta|y_1, \cdots y_N\right) \tag{4.5}$$

$$= \underset{\theta}{argmax}\, Pr\left(y_1, \cdots y_N|\theta\right) \frac{Pr(\theta)}{Pr(y_1, \cdots y_N)} \tag{4.6}$$

[1] Efficiency: the asymptotic variance of the estimated parameter achieves the Cramér–Rao lower bound.

Because $Pr(y_1, \cdots y_N)$ is independent of the parameters θ, so (4.6) becomes

$$\hat{\theta} = \underset{\theta}{argmax}\, Pr\left(y_1, \cdots y_N|\theta\right) Pr(\theta) \tag{4.7}$$

This is a Bayes estimator, where $Pr(\theta)$ is the prior information about the parameter. In practice, this information is uaually not available. If we assume that the prior is uniform distributed, an ML estimation is:

$$\begin{aligned} \hat{\theta} &= \underset{\theta}{argmax}\, Pr\left(y_1, \cdots y_N|\theta\right) && (4.8) \\ &= \underset{\theta}{argmax}\, Pr\left(n_{y_1}, \cdots n_{y_N}|\theta\right) && (4.9) \end{aligned}$$

4.2 The relation of ML, WLS and LS

A maximum likelihood method incorporates most statistical properties of the measurements in a model. It can be shown that if an efficient estimator exists, then the ML estimator is efficient[60]. Least Squares can be seen as a Maximum Likelihood estimator when the error of the observations follow an independent and identically Gaussian distribution. The following few simple manipulations will show why this is the case.

Assume that each measurement data point y_i has a measurement error n_{y_i}, which are independently and Gaussian distributed around 0. The probability of the error distribution then equals:

$$f(n_{y_i}) = f(y_i(x_i|\theta)) = \frac{1}{\sqrt{2\pi\sigma_i^2}} exp\left(-\frac{1}{2\sigma_i^2}\left(y_i - y\left(x_i|\theta\right)\right)^2\right) \tag{4.10}$$

where σ_i is the standard deviation of the ith error distribution. Since n_{y_i} are independent of each other, the probability of the data set equals the product of the probabilities:

$$f\left(n_{y_1}, \cdots n_{y_N}\right) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma_i^2}} exp\left(-\frac{1}{2\sigma_i^2}\left(y_i - y\left(x_i|\theta\right)\right)^2\right) \tag{4.11}$$

Maximizing (4.11) is equivalent to minimizing the negative of its logarithm, which is

$$\left[\sum_{i=1}^{N}\left(\frac{1}{2}log(2\pi) + log(\sigma_i) + \frac{1}{2\sigma_i^2}\left(y_i - y\left(x_i|\theta\right)\right)^2\right)\right] \tag{4.12}$$

If we now assume that all σ_i are equal, then minimizing (4.12) is equivalent to minimizing (4.1). It can be concluded that least-squares fitting is equal to maximum likelihood estimation if the measurement errors are independent and normally distributed with constant standard deviation. This is true for both linear and nonlinear models.

If we don't impose the assumption that all σ_i are equal, then we obtain (4.4). In this situation, the Weighted Least Squares using the weight $w_i = \frac{1}{\sigma_i^2}$ is equivalent to Maximum Likelihood [61].

4.3 Description of the problem

The detector used in this study consists of a $20 \times 20 \times 10$ mm^3 Lutetium Oxyorthosilicate (Lu_2SiO_5 or LSO) scintillator block coupled to two Hamamatsu S8550 Avalanche Photo Diode (APD) arrays. The LSO crystal is polished on all sides and wrapped in Teflon (figure 4.1). It is coupled to the APD surface with Meltmount type 5870 (Cargille Laboratories Inc.). This optical coupling compound was chosen because it is of a non-curing (i.e. thermoplastic) type. Disassembling or reassembling of the LSO crystal and the APD arrays then becomes straightforward by heating the detector module up to 60^oC. The particular type of Meltmount was selected because it has a high transmission at wavelengths above 400 nm, i.e. more than 90% for a 1.5 mm thick layer [37]. Its refractive index $\mu_{meltmount}$ equals 1.582. The APD arrays have an epoxy entrance window with a μ_{epoxy}= 1.52, close to the refractive index of meltmount and consists of 8×4 pixels, each measuring 1.6×1.6 mm^2 (figure 4.2).

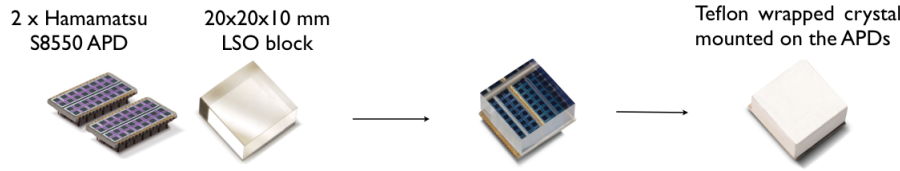


Figure 4.1 – Assembly stages of the detector components

We define the origin of the coordinate frame as the central point along the line inbetween the two APD arrays. The coordinates for the central point of the mth pixel are given by (x_m, y_m). For each gamma ray interaction, we measure the signal g_m in each of the 64 pixels. The problem can then be defined as

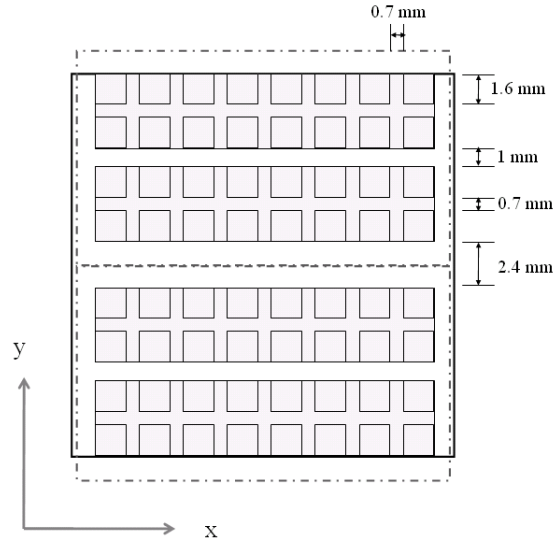


Figure 4.2 – Layout of the APD pixels in the two adjacent S8550 APD arrays (dash lines) used to read out the 20x20x10 mm LSO block (solid line).

finding a function $F(x_m, y_m|\boldsymbol{\theta})$ which predicts the signal detected by the APD pixels, given a set of adjustable parameters $\boldsymbol{\theta}$ that includes the 3D gamma ray interaction position (x, y, z) within the crystal. The (x_m, y_m) are the independent variable, g_m are the dependent variables. Finding the values for the parameter set θ such that $F(x_m, y_m|\boldsymbol{\theta})$ is as close as possible to the measured signals g_m is done by fitting the model F through the 64 data points using either a least squares procedure:

$$\hat{\boldsymbol{\theta}} = \underset{\theta}{argmin} \sum_{m=1}^{64} \left(g_m - F(x_m, y_m|\boldsymbol{\theta})\right)^2 \tag{4.13}$$

or a weighted least squares procedure, assuming that the number of photons seen by an APD pixel follows a Poisson distribution and ignoring electronic noise contributions to the signals:

$$\hat{\boldsymbol{\theta}} = \underset{\theta}{argmin} \sum_{m=1}^{64} \frac{\left(g_m - F(x_m, y_m|\boldsymbol{\theta})\right)^2}{F(x_m, y_m|\boldsymbol{\theta})} \tag{4.14}$$

Finding the model function F is the crucial point for all estimation methods. If no general function F exist, the least square method will not work. Then each data point needs to be treated separately as it has been done in section 3.2 for

the 3D ML method, where a look-up table had to be built for this reason. If a model function F can be found, fitting the parameters becomes much easier by minimizing the LS/WLS objective function using nonlinear iterative optimization methods (e.g. Levenberg-Marquardt) or using a boundary constrained search, which will be discussed in section 4.6.

4.4 Light distribution model $F(x_m, y_m|\boldsymbol{\theta})$

To find the general light distribution model F, we first have a look how the optical photons travel within the crystal. To simplify the problem, only photons which reflect maximum one time on each side of the surfaces are considered.

Scintillation photons produced at the interaction site inside an LSO block, wrapped in Teflon on 5 sides, can travel to the photo detector pixels along different paths (figure 4.3):

1. Optical photons are emitted directly towards the bottom surface of the LSO block.

2. Optical photons reach a pixel after internal reflection on a surface. In case of polished surfaces, these mirror-like reflections behave like if there is a virtual light source at a symmetric position on the other side of the surface.

3. Optical photons exit the block at the side or top, undergo diffuse reflection on the Teflon layer and re-enter the crystal at a random direction. These optical photons contribute to the background signal. They only add to the total energy measured (and hence to the energy resolution) but contain no spatial information anymore.

Given the above physical argumentations, we propose a model F that is composed of three parts:

1. A term f corresponding to the optical photons directly impinging on the pixel.

2. Additional virtual light sources (represented by an identical term f as in 2), mirrored around the surfaces to take internal reflections into account.

3. A constant C_{est} that is mainly due to the reflection of the optical photons on a diffuse reflector (e.g. Teflon) surrounding the scintillator block.

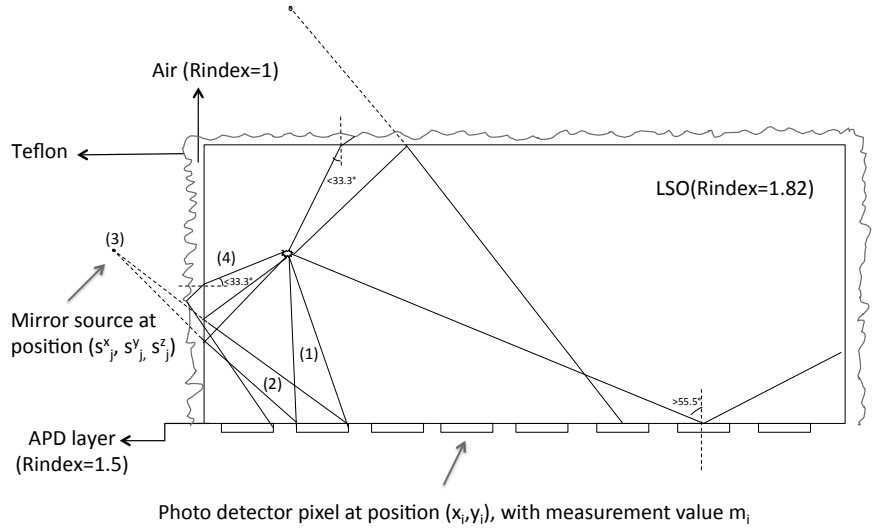


Figure 4.3 – Schematic drawing of the direct detected scintillation light (1) and the internally reflected light in a polished LSO crystal wrapped in Teflon on 5 sides (2). The latter is represented by virtual light sources mirrored around the edges (3). Optical photons can also reach a pixel after loosing all information about the point of origin in a diffuse reflection on the external Teflon layer (4).

The signals measured by the APDs are proportional to the number of scintillator photons and hence

$$F(x_m, y_m|\boldsymbol{\theta}) = C_{est} + f(x - x_m, y - y_m, z) + \sum_j f(s_j^x - x_m, s_j^y - y_m, s_j^z) \quad (4.15)$$

where s_j^x, s_j^y and s_j^z are the coordinates of the virtual source mirrored around the j^{th} side surface of the block.

4.5 The direct or virtual source term $f(x, y, z)$

One of the choices for the term f in model (4.15) is based on the assumption that the detected signal in a pixel is proportional to the solid angle Ω subtended

by that pixel seen from the source location. The signal produced by a pixel is computed by multiplying the number of photons arriving on that pixel with a constant factor A_0, which takes both the total amount of produced scintillation light, the APD gain and the amplifier gain into account, i.e.

$$f = A_0 \times \Omega \tag{4.16}$$

4.5.1 Exact solid angle based model

Considering the geometry in figure 4.4, the solid angle seen from a point source $S1$ located at a distance d straight up from one corner of a rectangular shaped pixel with length a and width b, is given by [27]

$$\Omega = arctan\left[\frac{ab}{d\sqrt{a^2+b^2+d^2}}\right] \tag{4.17}$$

When the projection of the source is located outside of the pixel, as $S2$ in figure 4.4, the solid angle can be written as a linear combination of four other solid angles, each of them similar to (4.17). The formula for the solid angle then becomes

$$\Omega_{S2-9456} = \Omega_{S2-1357} - \Omega_{S2-1348} - \Omega_{S2-1267} + \Omega_{S2-1298} \tag{4.18}$$

where $\Omega_{S2-9456}$ indicates a solid angle subtended by an area, whose four corners are points 9, 4, 5 and 6, as seen from the point source $S2$.

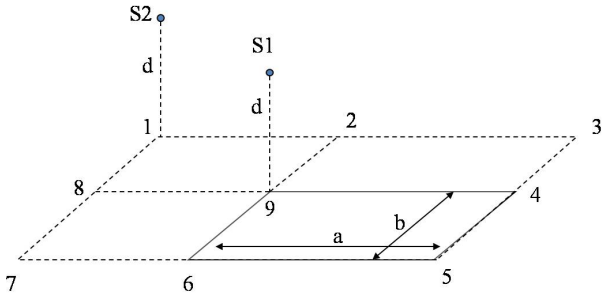


Figure 4.4 – The relative position of point sources at $S1$ and $S2$, positioned at a distance d above a rectangular detector of length a and width b.

The solid angle subtended by a pixel with dimensions (dx, dy) at position (x_m, y_m), as seen from the interaction position (x, y, z), can thus be written as

$$\begin{aligned}
\Omega &= arctan\left[\frac{(x-(x_m-{}^{dx}/_2))\times(y-(y_m-{}^{dy}/_2))}{z\times\sqrt{(x-(x_m-{}^{dx}/_2))^2+(y-(y_m-{}^{dy}/_2))^2+z^2}}\right] \\
&- arctan\left[\frac{(x-(x_m-{}^{dx}/_2))\times(y-(y_m+{}^{dy}/_2))}{z\times\sqrt{(x-(x_m-{}^{dx}/_2))^2+(y-(y_m+{}^{dy}/_2))^2+z^2}}\right] \\
&+ arctan\left[\frac{(x-(x_m+{}^{dx}/_2))\times(y-(y_m+{}^{dy}/_2))}{z\times\sqrt{(x-(x_m+{}^{dx}/_2))^2+(y-(y_m+{}^{dy}/_2))^2+z^2}}\right] \\
&- arctan\left[\frac{(x-(x_m+{}^{dx}/_2))\times(y-(y_m-{}^{dy}/_2))}{z\times\sqrt{(x-(x_m+{}^{dx}/_2))^2+(y-(y_m-{}^{dy}/_2))^2+z^2}}\right] \qquad (4.19)
\end{aligned}$$

4.5.2 Approximate solid angle based model

If we assume that the distance from the source to the center of each pixel is large compared to the size of the pixel, then we can write an approximate solid-angle function

$$\Omega = dx \times dy \times \frac{z}{\left((x-x_m)^2+(y-y_m)^2+z^2\right)^{3/2}} \qquad (4.20)$$

where dx and dy are the width and length of one pixel (both 1.6 mm in our geometry). The above condition is not satisfied when the interaction point is very close to the crystal bottom. To check the error made by this approximation, the solid angle subtended by two neighboring pixels at positions (0, 0) and (2.3, 0) was computed using (4.19) and (4.20) for two different light source positions A and B (figure 4.5). Light source B at position (-10, 0, 5) is far away from the pixels. In this case the numerical results from (4.19) and (4.20) were identical. On the other hand, light source A at position (0, 0, 1) is very close to the pixels. The numerical result for the approximate solid angle subtended by pixel (0, 0) is 59.6% more than the value found for the exact solid angle. But the difference between the results obtained from (4.19) and (4.20) for the neighboring pixel at (2.3, 0) is already significantly reduced to 9.9 %. Hence, only inaccurate approximate solid angles (4.20) will be found when the interaction is very close to the bottom part

of the LSO block and only for the pixel directly underneath. The other 63 pixels yield values that are accurate enough to produce an excellent overall accuracy.

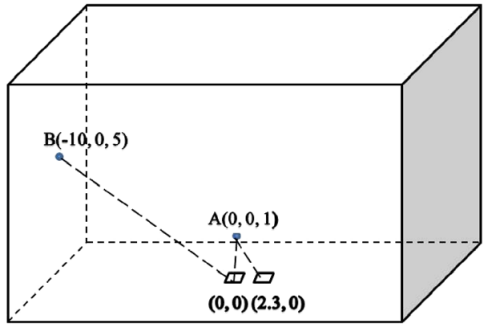


Figure 4.5 – Two example sources A, B and two pixels.

4.5.3 Extended approximate solid angle in the model

In (4.16), we made the following implicit approximations:

- Absorption and scattering of optical photons in the block along its path is negligible and when internal reflections occurs, they behave like 100% specular reflections.
- The internal reflections at the bottom surface are not considered, i.e. all optical photons arriving at the bottom surface leave the scintillator block.
- Each pixel always receives a contribution from all virtual sources

The last two approximations can be removed by extending the solid angle model to take internal reflection into account. Only optical photons emitted towards the bottom surface within a cone with an opening angle smaller than the critical angle of 55.5^o, determined by the LSO-Meltmount interface, can immediately leave the block. The Meltmount-APD interface can be neglected since their refractive indices are nearly the same. The previous two models did not take the effect of internal reflections at the bottom surface into account. This effect is modeled by adding a cut-off factor to the term $f(x, y, z)$, which becomes zero when the angle between the bottom surface normal and the line connecting the scintillation light source with the pixel center is larger than the critical angle of 55.5^o.

The two previous models also assumed that every APD pixel detects optical photons arriving after internal reflection on the side surfaces, i.e. the four virtual

sources are always present for a given pixel. This is not always the case. For some indirect optical paths, arriving at a given pixel after hitting a side surface, the angle with the normal of that side surface is less than the critical angle of 33.3^o for an LSO-air interface. Optical photons traveling on this path would hence exit the crystal at the side instead of being internally reflected towards the given APD pixel, i.e. there is no virtual light source at that side of the block for that pixel. Adding a similar cut-off factor to the terms f representing the virtual sources in (4.15) removes the contribution of the non-existing virtual sources to the APD pixel signal.

The constrained minimization algorithm used (see section 4.6) requires a continuous objective function [17]. The cut-off factor, to describe the crossover from the refractive regime into the reflective regime, is therefore added into the function $f(x, y, z)$ using a sigmoid function. In case of the approximate solid angle, f becomes

$$\begin{aligned} f &= A_0 \times \Omega \times \sigma\left(\theta_c - \theta\right) \\ &= A_0 \times \frac{z}{\left(\left(x - x_m\right)^2 + \left(y - y_m\right)^2 + z^2\right)^{3/2}} \times \frac{1}{1 + e^{-\beta \times (\theta_c - \theta)}} \end{aligned} \tag{4.21}$$

where

$$\theta = arctan\left[\frac{\sqrt{\left(x - x_m\right)^2 + \left(y - y_m\right)^2}}{z}\right] \tag{4.22}$$

is the angle between the line from the light source to the detector pixel and the bottom surface normal, θ_c is the critical angle for the bottom surface and the smooth factor β determines the transition gradient between the two regimes. We found that the value of β is not very critical and set $\beta = \frac{1}{100^o}$ in our model.

4.6 Minimization methods

For each event, the 3D interaction position (x, y, z), the term A_0 and the background parameter C_{est} are estimated by fitting model (4.15) using either a LS (4.13) or WLS (4.14) method. A boundary constrained minimization method should be used, since $\hat{x}, \hat{y}, \hat{z}$ should be within the physical boundaries of the crystal and both C_{est} and A_0 should be non-negative. To simplify the problem, we

will first discuss the linear least squares problem, and then the unconstrained nonlinear minimization problem [56, 26, 60, 66], followed by the bound-constrained problem.

4.6.1 Linear problem

A linear least squares problem can be stated as:

$$\underset{\theta}{argmin}||Y - X\theta||^2 \tag{4.23}$$

where Y are the dependent variables, and X are the independent variables,

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}, X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix}, \theta = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_n \end{bmatrix}$$

The solution of (4.23) is [60]:

$$\theta = (X^T X)^{-1} X^T Y \tag{4.24}$$

4.6.2 Unconstrained minimization

The general unconstrained minimization problem has a simple mathematical formulation :

$$\underset{\theta}{argmin} f(\theta) \tag{4.25}$$

Here we assume that $f(\theta)$ is a smooth function, i.e. whose second derivative exists and is continuous.

There is no such solution as (4.24) existing for a nonlinear problem. Therefore, iterative searching algorithms are needed to minimize a nonlinear problem. The basic idea of iterative searching is that, at the k^{th} iteration, a new solution $\theta^{(k+1)}$ is generated by

$$\theta^{(k+1)} = \theta^{(k)} + \delta\theta^{(k)} \tag{4.26}$$

where $\delta\theta^{(k)}$ is chosen in such a way that it lowers the value of the objective function $f(\theta)$. The iterative procedure continues until it has converged, i.e. $\theta^{(k)}$ is sufficiently close to the optical solution θ_{min} of f.

Two main techniques to generate $\delta\theta^{(k)}$ are line search and trust region search. Based on these two strategies, various methods exist such as Newton's method, which is a line search method, and Levenberg–Marquardt algorithm, which is a trust region method.

Line Search methods

In line search algorithms, (4.26) becomes

$$\theta^{(k+1)} = \theta^{(k)} + \alpha^{(k)} p^{(k)} \tag{4.27}$$

where $\alpha^{(k)}$ is the step length, determined to achieve the optimal solution along the direction $p^{(k)}$. The direction $p^{(k)}$ is required to be a descent direction of $f(\theta)$ at $\theta^{(k)}$, i.e. $p_k^T \nabla f(\theta^{(k)}) < 0$ where

$$\nabla f(\theta) = \begin{bmatrix} \partial f(\theta)/\partial\theta_1 \\ \partial f(\theta)/\partial\theta_2 \\ \vdots \\ \partial f(\theta)/\partial\theta_n \end{bmatrix} \tag{4.28}$$

This guarantees that the function $f(\theta)$ can be reduced along this direction.

Once the descent direction $p^{(k)}$ is chosen, the step length $\alpha^{(k)}$ should be determined along the descent direction. One possibility among several ways to find a step size $\alpha^{(k)}$ is the Armijo Rule [66]:

Let $s > 0$ be a constant, $\rho \in (0,1)$ and $\mu \in (0,1)$. Take $\alpha^{(k)}$ to be the largest α in $\{s, s\rho, s\rho^2, \ldots\}$ such that

$$f(\theta^{(k)}) - f(\theta^{(k)} + \alpha p^{(k)}) \geq -\alpha\mu p_k^T \nabla f(\theta^{(k)}) \tag{4.29}$$

Trust region methods

The idea behind a trust region method for unconstrained minimization is that, in a neighborhood of the current point $\theta^{(k)}$, a simpler function, usually a quadratic model $q_k(\theta)$, is used to approximate the objective function in that neighborhood. The step is then chosen to be the minimizer of this approximate model in this region. This neighborhood region is called trust region. In fact, contrary to the line search method, the step and direction are obtained simultaneously.

At each of the iterations in the trust-region method, the following subproblem need to be solved:

$$p^{(k)} = minq_k(\theta) = f(\theta^{(k)}) + \nabla f(\theta^{(k)}) \cdot p + \frac{1}{2} p \cdot B^{(k)} p \quad (4.30)$$

$$subject\,to\,||p|| \leq \triangle^{(k)} \quad (4.31)$$

where the quadratic approximation $q_k(\theta)$ is defined by the first two terms of the Taylor approximation to $f(\theta^{(k)} + p)$, $\triangle^{(k)}$ is a trust region radius and $B^{(k)}$ is a symmetric approximation to the local Hessian matrix $H^{(k)}$, where

$$H(\theta) = \nabla^2 f(\theta) = \begin{bmatrix} \frac{\partial^2 f(\theta)}{\partial \theta_1^2} & \frac{\partial^2 f(\theta)}{\partial \theta_1 \partial \theta_2} & \cdots & \frac{\partial^2 f(\theta)}{\partial \theta_1 \partial \theta_n} \\ \frac{\partial^2 f(\theta)}{\partial \theta_2 \partial \theta_1} & \frac{\partial^2 f(\theta)}{\partial \theta_2^2} & \cdots & \frac{\partial^2 f(\theta)}{\partial \theta_2 \partial \theta_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\theta)}{\partial \theta_n \partial \theta_1} & \frac{\partial^2 f(\theta)}{\partial \theta_n \partial \theta_2} & \cdots & \frac{\partial^2 f(\theta)}{\partial \theta_n^2} \end{bmatrix} \quad (4.32)$$

Figure 4.6 illustrates an example of a trust-region approach and a line search approach on a function f in which the current point $\theta^{(k)}$ and the optimal solution θ_{min} lie at opposite ends of a curved valley. A line search method searches along its descent direction and the step length α_k. For the trust region method, the step direction is obtained at the same time by minimizing (4.30), the step size is restricted by 4.31 within the dotted circle.

4.6.3 Bound-constrained optimization

The bound constrained optimization problems have the general form of

$$\underset{\theta}{argmin} \quad f(\theta)$$

subjected to

$$l \leq \theta \leq u$$

where $f(\theta)$ is a smooth function, The lower and upper bounds on the variable θ are given by l and u. The feasible region is often called a "box" because of its rectangular shape.

Several methods have been proposed to solve this problem. The two main methods are active set Sequential Quadratic Programming (SQP) and interior point methods [56].

This thesis will use the trust region reflection method, which is also a kind of interior point method, available in the MATLAB optimization toolbox. This

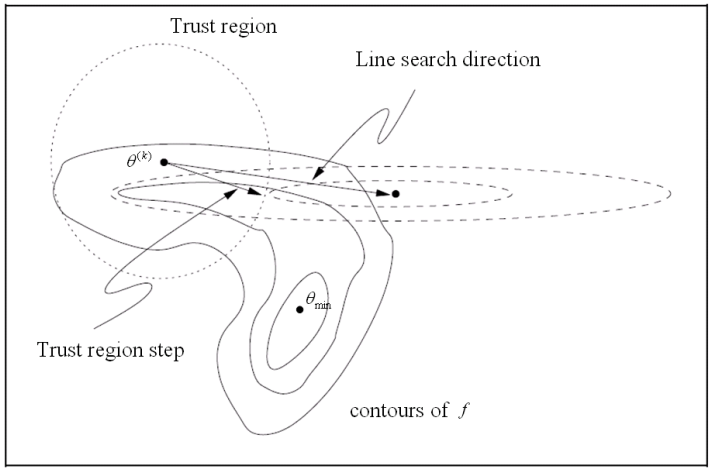


Figure 4.6 – An illustrative example of trust-region and line search step. The objective function f has the optimal solution θ_{min}, and the current point is at $\theta^{(k)}$. A line search method searches along its descent direction and the step length α_k. For the trust region method, the step direction is obtained at the same time by minimizing (4.30), the step size is restricted by 4.31 within the dotted circle. Figure from [56].

method was called as a kind of "interior-point" method because the estimates in each iteration is strictly feasible. The essential difference with the classic interior point methods is that the classical methods use a barrier function to ensure feasibility. The algorithm used follows a piecewise linear path, reflecting back as they come across the constraints.

A reflective transformation from the original θ space to another β space, where $\theta = R(\beta)$ and R is a continuous mapping, is used to guarantee the feasibility. The constrained problem in θ space becomes an unconstrained problem in β space, and therefore a standard descent direction searching method can be used in β space. The condition to guarantee that a local minimizer in β space is also a local minimizer in θ space is that the reflection R has to be a so called Open Mapping[2] (Proven in [18]).

[2]Open Map is a continuous linear operator, which is surjective between Banach spaces.
An operator $f : X \to Y$ is a surjective operator if and only if for every y in Y there is at least one matching x in the domain X, such that $f(x) = y$.

The reflective transformation $\theta_j = R(\beta_j)$ proposed by [17] is:

$$
\begin{aligned}
&Case\,1: (l_j > -\infty,\, u_j < \infty)\; upper\; and\; lower\; bound: \\
&\qquad w_j = |\beta_j - l_j| \bmod [2(u_j - l_j)],\; \theta_j = min(w_j, 2(u_j - l_j) - w_j) + l_j \\
&Case\,2: (l_j > -\infty,\, u_j = \infty)\; lower\; bound: \\
&\qquad if\; \beta_j \geq l_j,\; \theta_j = \beta_j,\; else\; \theta_j = 2l_j - \beta_j \\
&Case\,3: (l_j = -\infty,\, u_j < \infty)\; upper\; bound: \\
&\qquad if\; \beta_j \leq u_j,\; \theta_j = \beta_j,\; else\; \theta_j = 2u_j - \beta_j \\
&Case\,4: (l_i = -\infty,\, u_i = \infty)\; unbounded \\
&\qquad unconstraint
\end{aligned}
$$

where j is the index of the j^{th} dimension in θ.

An illustration of a 1-dimensional reflective transformation with finite upper and lower bound is given in figure 4.7.

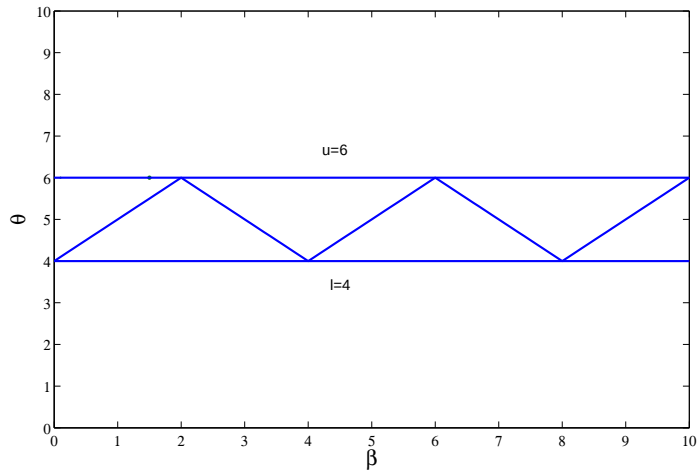


Figure 4.7 – A 1-Dimensional reflective transformation example with finite lower and upper bound of respectively of 4 and 6. [17]

4.6.4 Initial value and local minima

A numerical searching algorithm for nonlinear optimization problems usually seeks only a local solution, where the objective function is smaller than its nearby points.

This might not be the global solution over all feasible points. Good initial guesses of the starting values are therefore extremely important in order to reach the global minimum of the cost function.

In our search, the initial value of X and Y are set to the coordinate of the center of the pixel that produced the biggest signal. The initial value of the DOI coordinate z is set to the average depth of the first interaction of a 511 keV photon. Given a mean free path (MFP) of 11.3 mm in LSO, this yields an initialization value of 4.2 mm. The initial value of the background constant C_{est} is set to the average of the 10 to 15 smallest APD pixel values. The initial value of parameter A_0 is set to 0.8, assuming that the 64 APD pixel signals have been normalized for energy, i.e. each pixel value is divided by the sum of the 64 pixels.

4.7 Confidence intervals

Knowing the best-fit parameter is not the only goal of a regression. It is also important to know the precision of the estimated parameters. This is presented as confidence intervals (CI) (or standard errors). A 95% CI is an interval that has a 95% chance of containing the true value of the parameter. There are normally two ways to compute the CI: an asymptotic method, which is used in most nonlinear regression programs, and a Monte Carlo method.

4.7.1 Asymptotic method

Using the asymptotic method, a CI can be computed from the best-fit value $\hat{\theta}$, the standard errors of the estimated parameters σ_θ, and the Student's t-distribution with v degree of freedom using the equation:

$$\hat{\theta} + [-\sigma_\theta \times t(1 - \alpha/2,\, v),\, \sigma_\theta \times t(1 - \alpha/2,\, v)] \tag{4.33}$$

where v is equal to the number of data points minus the number of parameters fit by nonlinear regression and α equals the percentage of the CI. The term $t(1 - \alpha/2,\, v)$ is a constant, which is near 2.0 when there are plenty of degrees of freedom (more than a dozen).

The CI computed using these errors should also be considered approximate since the standard errors of the estimated parameters returned by most nonlinear regression programs are "approximate".

The problems with asymptotic confidence intervals computation is:

1. The intervals are always symmetrical, as we can see in (4.33).

2. The computation is based on a mathematical simplification and it is only an approximation for nonlinear equations. Only when there are plenty of data that clearly define the curve without a huge amount of scatter, the approximation for the confidence intervals gives a good indication on the precision of the estimates. With marginal data, the approximation doesn't always work so well [55].

Those issues are the reason that a standard CI checking is not performed in this thesis. Since in each fitting, we only have 64 data points. In addition, it has been found that CI estimations are unstable for data close to the boundary, i.e, when gamma ray interaction close to the boundary of the crystal.

4.7.2 Monte Carlo method

The idea of Monte Carlo confidence interval calculation is to self generate a bunch of (e.g.1000) data sets, using the same independent x value and the same model with the "best-fit" parameters. Then different fluctuations on those simulated ideal data are added. The fluctuation obeys a zero-mean Gaussian distribution with a standard deviation which is estimated from the residuals of a previous fitting. Then each simulated data is used to estimate the parameters. Finally, the distribution of parameter estimators create the confidence intervals.

This method is not realistic in our case in a real system. For example, in a BrainPET scanner, up to 400.000 coincidences happen per second [50]. If the CI bound has to be computed for the estimated interaction position of each of these events, 400.000.000 simulated events would need to be generated and fitted in a single second.

Therefore, no CI bound are generated to check the parameter fitting in this thesis, i.e. all results are accepted to generate a final spatial resolution.

Chapter 5

Model Evaluation

In this chapter, the models described in Chapter 4 will be evaluated using simulation and experimental data. The initial values of the model parameters have been discussed in section 4.6.4. For X and Y, the initials are set to the coordinate of the center of the pixel that produced the biggest signal and for the DOI coordinate, the inital value is set to the average depth of the first interaction of a 511 keV photon. The initial value of the background constant C_{est} is set to the average of the 10 to 15 smallest APD pixel values. The initial value of parameter A_0 is set to 0.8, assuming that the 64 APD pixel signals have been normalized for energy, i.e. each pixel value is divided by the sum of the 64 pixels.

When not specified, the resolutions stand for "global resolution", which is the FWHM and FWTM obtained from one single error histogram generated by all events over the entire crystal surface. In both simulation and experiment, the 511 keV gamma beam pitch equals 0.5 mm in both the X and the Y direction.

5.1 Detector setup description

The detector used in this study was introduced in section 4.3. Here we add some more details of the set-up used to generate the experimental data.

Each of the 64 APD pixels is equipped with its individual readout chain consisting of a CREMAT CR-110 front-end preamplifier, a CAEN N568B spectroscopic amplifier channel (16 channels/module), and a CAEN V785 peak sensing ADC channel (32 channels/module) (figure 5.1). A homemade circuit provides the analog sum of the 64 channels to yield the total energy of an event. The front-end part of the readout, i.e. the scintillator, the APD arrays and the CREMAT

pre-amplifiers are enclosed in an air and light-tight box, cooled to a temperature of $18 \pm 0.6\,^oC$ by a flow of cold liquid from a cooling unit.

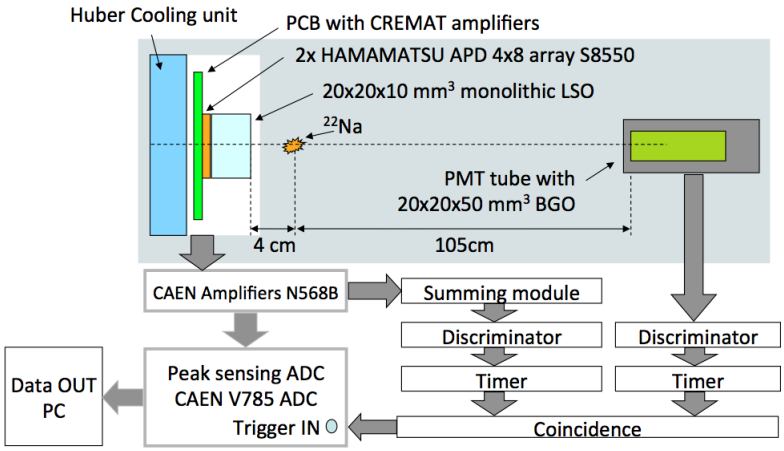


Figure 5.1 – Schematic diagram of the experimental set-up and data acquisition.

The detector characteristics are given in table 5.1. The average detection quantum efficiency (DQE) of the S8550 APD arrays was computed using the LSO emission spectrum [20], the Meltmount transmission curve [38] and the APD QE curve [54]. The excess noise factor (ENF) of the APD arrays was also reported in [54]. The operational APD gain of about 85x, was estimated by comparing the output signals relative to those obtained at a gain of 50x. The high voltage that needs to be applied to the APD arrays for a gain of 50x was specified by the manufacturer. A measurement on a large number of the APDs used in this study, showed a pixel gain variation of 4.6 % FWHM [1]. The Pre-amplifier ENC is measured using the same way as in [48], where the electronic noise was determined from the RMS on the analog signal going into the ADCs. To convert the RMS voltage fluctuation to an equivalent noise charge (ENC) at the input of the Cremat pre-amplifiers, the gain of the readout channel was calibrated by injecting a known charge signal into the input of the pre-amplifiers.

Three of the 64 channels (channels 17, 26 and 64 in figure 5.2) were dead or damaged . These were not used when minimizing the cost function (4.13) and (4.14) to determine the 3D interaction position.

Parameter	Value
APD DQE	70.4 %
APD gain	85 x
APD ENF	1.75
Pre-amplifier ENC	1800 e^- RMS

Table 5.1 – Detector parameters.

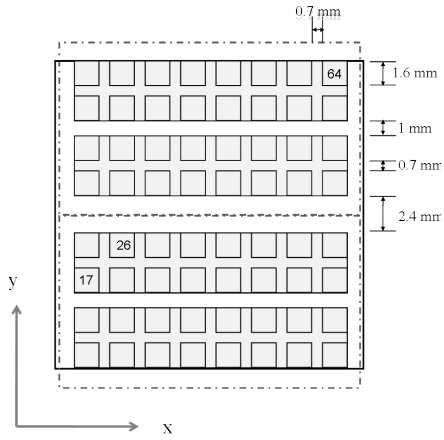


Figure 5.2 – Three of the damaged channels

5.2 Simulation analysis

5.2.1 Generation of Monte Carlo modeling data

Monte Carlo simulations of the detector were performed using GATE [35]. The light yield of the LSO scintillator was set at 26,000 optical photons/MeV [21]. The total attenuation length was set to 138mm[53]. The surface condition of the LSO block is modeled using micro-facets according to the UNIFIED model [25]. The normals to the micro-facets are randomly distributed with a mean of 0 and standard deviation of σ_α, i.e. σ_α =0 means a perfectly polished surface. We use σ_α= 0.1^o for the polished surfaces in our simulations [14]. The diffuse reflection of the Teflon surrounding the LSO block was characterized by a reflection coefficient of 95%.

Scintillation light distributions measured by the matrix of APD pixels were simulated for 511 keV photons impinging on top LSO surface on a rectangular grid of beam positions (0.5 mm pitch in both x and y direction) covering the

complete surface. The 511 keV photon beam was assumed to be perfect, i.e. no influence of a finite beam size was included in the simulation.

Adding detector and readout induced noise

The GATE Monte Carlo simulation did not take the influence of the APD DQE, the APD gain, the APD excess noise factor and the pre-amplifier noise into account. An additionally normally distributed random number with standard deviation σ_{de} is therefore added onto the simulation data. Hence, the total simulated pixel signal variance equals

$$\sigma^2_{total} = \sigma^2_{N_\gamma} + \sigma^2_{de} \tag{5.1}$$

where N_γ is the simulated number of the scintillation photons impinging on a single photo detector.

Suppose the fraction of optical photons emitted from the gamma ray interaction that will reach the m^{th} photo detector pixel, $\eta_m(\mathbf{R})$, is small. As has been discussed in 2.3.1 and 3.2.1, this small fraction makes that N_γ is Poisson distributed, because a low-efficiency binomial selection of a non-Poisson random process is a Poisson random process [5]. Hence

$$\sigma^2_{N_\gamma} = \overline{N}_\gamma \tag{5.2}$$

Substituting (5.2) in (2.17), we can write the inverse of the signal-to-noise ratio (SNR^{-1}) of the signal generated by the m^{th} photo detector pixel as

$$\begin{aligned} SNR^{-1}_{output} &= \frac{\sigma}{\overline{E}} \\ &= \sqrt{\frac{ENF}{QE \times CE \times \overline{N}_\gamma} + \left(\frac{ENC}{\overline{G} \times QE \times CE \times \overline{N}_\gamma}\right)^2} \end{aligned} \tag{5.3}$$

Since SNR is unit free, we can use it to compute an equivalent noise level at any stage of the readout chain, expressed in the unit of the signal at that level. If we apply this to the APD pixel level, we can express the total noise as an equivalent fluctuation on the number of optical photons hitting a pixel, i.e.

$$\sigma^{APD}_{total} = N_\gamma \cdot SNR^{-1}_{output} \tag{5.4}$$

Using expression (5.3), equation (5.4) becomes

$$\sigma_{total}^{APD} = N_\gamma \sqrt{\frac{ENF}{\overline{N}_\gamma \cdot DQE} + \left(\frac{ENC}{\overline{N}_\gamma \cdot DQE \cdot G}\right)^2} \tag{5.5}$$

with $DQE = QE \times CE$ the APD detection quantum efficiency. Assuming $\overline{N}_\gamma$ can be approximated by N_γ, the detector and electronic related noise components σ_{de} can now be expressed in an equivalent number of optical photons hitting the APD pixel using (5.1) and (5.5)

$$\begin{aligned} \sigma_{de} &= \sqrt{(\sigma_{total}^{APD})^2 - \sigma_{N_\gamma}^2} \\ &= \sqrt{\frac{ENF \cdot N_\gamma}{DQE} + \left(\frac{ENC}{DQE \cdot G}\right)^2 - N_\gamma} \end{aligned} \tag{5.6}$$

This expression for σ_{de} is finally used to simulate the detector and electronic related noise contribution which is added to the simulated number of optical photons hitting a pixel. The values for DQE, CE, ENF, G and ENC are the same as those shown in table 5.1.

After adding the detector and electronic noise, only events that cross an energy threshold of 350 keV were accepted. For these events, the predicted X, Y and DOI position was compared with the true interaction point in the LSO block. The FWHM and FWTM of the distribution obtained by the difference between the predicted and true coordinates was used to define the resolution.

5.2.2 Solid angle model and parameter fitting procedure

Table 5.2 shows the FWHM and FWTM resolutions in both X, Y and DOI direction. Using an exact or approximate solid-angle calculation yields similar results. Because most of the interactions occur in the upper part of the LSO block, where there is no numerical difference between the pixel signal values computed using (4.19) and (4.20), the difference between the results obtained with the two solid-angle models is small. However, the approximate solid-angle model is simpler to compute and hence results in shorter execution times (about 30% faster).

A comparison of the LS (4.13) and the WLS (4.14) parameter fitting methods shows that the achieved X-Y resolutions are similar. The LS fitting of the DOI coordinate is slightly better. It is not surprising that WLS does not perform better since the premise to use WLS (i.e. observables should be Poisson distributed) is not fulfilled due to the influence of ENF and the ENC.

Parameter fitting method	Ω model	X-direction		Y-direction		DOI-direction		
		FWHM (mm)	FWTM (mm)	FWHM (mm)	FWTM (mm)	FWHM (mm)	FWTM (mm)	Bias (mm)
LS	Exact	1.14	3.84	1.21	4.41	1.84	5.93	0.45
LS	Approximate	1.13	3.98	1.20	4.41	1.91	5.79	0.40
WLS	Exact	1.16	3.95	1.28	4.20	2.15	6.37	0.60
WLS	Approximate	1.16	3.98	1.27	4.20	2.23	6.36	0.54

Table 5.2 – Solid angle versus approximate solid angle.

Using (5.6) and the parameters in table 5.1, it follows that the electronic noise dominates when the number of optical photons hitting an APD pixel is less than 340. The distribution of the optical photons over the APD pixels was studied Using simulation data. For each event, the number of optical photons impinging on each pixel was sorted from high to low. This resulted in 64 data sets, i.e a set of the biggest signals, second biggest signals and so on.

Figure 5.3 shows the mean of each of these data sets. The number at ordinate value 1 corresponds to the pixels receiving the most photons whereas ordinate value 64 correspondence to the smallest number of photons receive on a pixel. The distribution shows that most of the useful signals in a scintillation flash are measured by 5 to 10 pixels. On average they receive 100 to 300 optical photons. This corresponds to a signal of 6000 to 19000 electrons at the input of the preamplifier. This is to be compared with a noise level of 1800 electrons ENC.

The photon distribution levels off to a slowly decreasing trend for the remaining pixels. On average 40 to 60 optical photons will hit these pixels, resulting in an electronic signal of 2400 to 3600 electrons.

All models and fitting procedures also introduced a bias (i.e. a systematic error) on the predicted DOI. On average the interaction position was estimated 0.4 to 0.6 mm closer to the bottom of the LSO block than the true interaction point due to multiple interactions. On average, the estimated X-Y position coincides with the true one, i.e. there is no global detectable bias on the coordinates in the X-Y plane. The local response (resolution and bias) is detailed in section 5.2.8.

Given these results, we will continue to use the approximate solid angle model and LS fitting for all the remaining results.

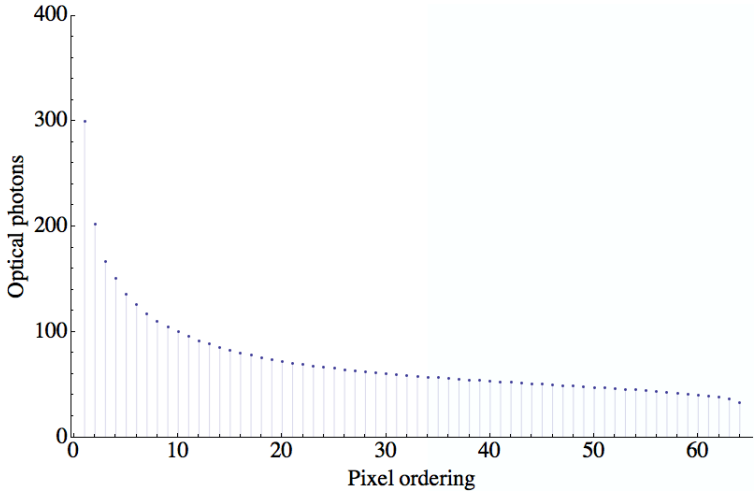


Figure 5.3 – Average number of optical photons impinging on the APD pixels as a function of the pixel order.

5.2.3 How many mirror sources are needed?

Optical photons emitted at the interaction site of a 511 keV gamma ray can reach an APD pixel via a direct path or after a reflection on one of the 5 sides of the LSO block. Internal reflections are modeled using mirror sources. Table 5.3 shows the impact of adding these internal reflections in the model.

When 4 mirrors sources are placed around the side surfaces, nearly all resolutions improve. Especially the FWTM resolutions are significantly enhanced.

Adding an additional fifth mirror source above the upper block surface, has little effect on the FWHM resolutions. This indicates that optical photons reflected on the top surface and reaching the APD pixels don't add significant information to enhance the extraction of the interaction positions.

It should also be noted that adding the virtual sources in the model increases the bias in the DOI estimate slightly. It should also be noted that adding the virtual sources in the model increases the bias in the DOI estimate slightly. This is due to the fact that the critical angle was not taken into account in the model. This makes that for some of the pixels, the model assumes that the photons are coming from the direct source and the mirror sources, but in fact they are only coming from the direct source. As a result, the model assumes the shape of light is actually narrower than in reality, which causes the DOI estimation to have a bias towards the bottom.

Ω model	X-direction		Y-direction		DOI-direction		
	FWHM (mm)	FWTM (mm)	FWHM (mm)	FWTM (mm)	FWHM (mm)	FWTM (mm)	Bias (mm)
no mirror source	1.21	5.02	1.39	5.36	1.84	6.86	0.21
4 mirror sources	1.10	3.98	1.20	4.41	1.91	5.79	0.40
5 mirror sources	1.13	4.31	1.19	4.27	2.09	6.45	0.50

Table 5.3 – Resolutions when using a model with no virtual sources, 4 virtual sources or 5 virtual sources.

5.2.4 Influence of detector noise level

To investigate the influence of the ENC on the achievable resolution, simulation studies were performed using different noise levels. Table 5.4 shows the X, Y and DOI resolution as well as the energy resolution as a function electronic noise (ENC expressed in electrons at the input of the preamplifier). The intrinsic X-Y resolution can be improved by 0.2-0.3 mm FWHM when the ENC is reduced from 1800 electrons to 600 electrons. The DOI resolution improves by 0.6 mm FWHM for a similar reduction of the electronic noise level. The energy resolution improves from 18.2% to 13.1%. Lowering noise levels below 600 electrons ENC does not seem to have a significant additional effect on the resolutions.

ENC (electrons)	X-direction		Y-direction		DOI-direction			ΔE (%)
	FWHM (mm)	FWTM (mm)	FWHM (mm)	FWTM (mm)	FWHM (mm)	FWTM (mm)	Bias (mm)	
300	0.77	2.62	0.83	2.70	1.34	3.94	0.35	12.9
600	0.78	2.82	0.99	2.70	1.39	4.42	0.37	13.1
1200	0.91	3.20	1.06	3.23	1.69	5.10	0.39	15.2
1800	1.10	3.98	1.20	4.41	1.91	5.79	0.40	18.2
2100	1.20	4.24	1.71	4.16	2.05	6.65	0.43	19.5

Table 5.4 – Spatial and energy resolution as a function of the electronic noise level.

5.2.5 Extended approximate solid angle model

Table 5.5 gives the FWHM and FWTM when critical angles are included in the approximate solid angle model. Although this extended model is more realistic, the achieved DOI resolution is worse compared to those in table 5.2. The X-Y

resolution is also slightly worse. This seems to indicate that the problem lies with the model fitting. To study this problem, APD pixel signals were generated using the extended approximate solid angle based model itself, i.e. using (4.15) and (4.21) for an interaction at some point. Then, the cost function (4.13), is evaluated using the generated pixel values for positions around the chosen interaction point. One typical example is shown in figure 5.4 for the approximate solid angle model and extended approximate solid angle model. The interaction point was chosen as (x, y, z) = (2, 5, 3). The values for C and A were chosen to be 2 and 200. The 3D plots in figure 5.4 show the value of the cost function when X and DOI are varied around the interaction point. The cost function generated by the approximate solid angle model (figure 5.4(a)) is very smooth, while the cost function which includes the critical angle into the model creates local minima in the cost function (figure 5.4(b)). From this example, we can see that the extended approximate solid angle model is very sensitive to good initial values due to local minima in the cost. This is especially the case for the DOI estimation.

To overcome this initial value problem, we first used the approximate solid-angle model to get a first estimate of the model parameters. These values are then used as initial values for the extended approximate solid model fitting. The results, shown in table 5.5, are now very comparable to those in table 5.2, but still not better.

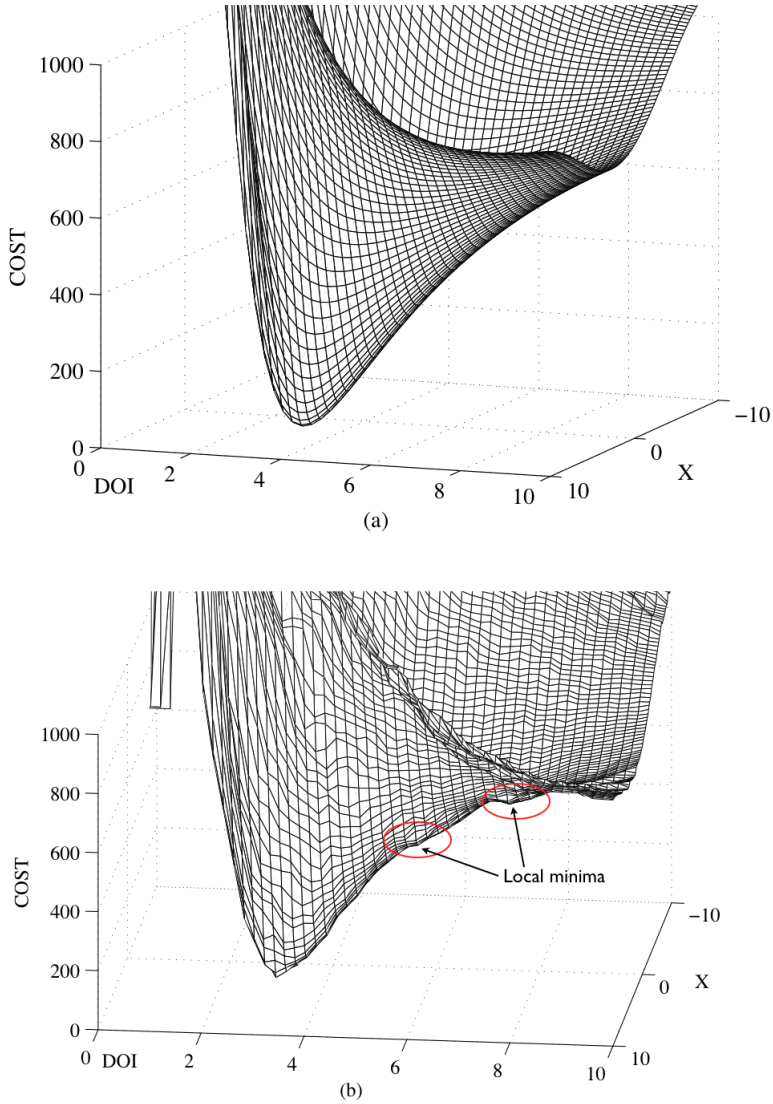


Figure 5.4 – (a) Cost function as a function of X and DOI for an interaction at (x,y,z)=(2,5,3) using the approximate solid angle model. (b) Similar plot for the extended approximate solid angle model

Initialization method	X-direction		Y-direction		DOI-direction		
	FWHM (mm)	FWTM (mm)	FWHM (mm)	FWTM (mm)	FWHM (mm)	FWTM (mm)	Bias (mm)
Standard parameters	1.26	4.70	1.42	4.98	5.37	7.40	1.49
Approximate solid angle model	1.13	4.24	1.21	4.18	2.02	6.04	0.34

Table 5.5 – Resolution obtained with extended approximate solid angle model and initialization using either the standard parameters or results from the normal approximate solid angle model.

5.2.6 Influence of Compton scattering inside the crystal

Using a monolithic scintillator block, it is not possible to reject those events that undergo multiple interactions in the block, thereby creating multiple light emission points. The influence on the spatial resolution when multiple interactions are either included or excluded from the simulated data set is shown in table 5.6. Multiple interactions have more influence on FWTM than FWHM. On the FWHM, they add an extra component to the X-Y resolution of about 0.6 mm FWHM. For the DOI resolution, the extra resolution component is about 1.0 mm FWHM. Multiple interactions have more influence on the FWTM in all three directions. The larger influence on the DOI resolution can be understood from the fact that Compton scattering at 511 keV is mostly forward along the DOI direction.

	X-direction		Y-direction		DOI-direction	
	FWHM (mm)	FWTM (mm)	FWHM (mm)	FWTM (mm)	FWHM (mm)	FWTM (mm)
No multiple interaction	0.96	3.22	1.05	2.99	1.61	4.93
With multiple interaction	1.13	4.24	1.21	4.41	1.91	5.79

Table 5.6 – Influence of multiple interactions on the resolution.

5.2.7 Influence of surface polishing

To study the influence of unpolishing the four side surfaces and top surface, we set σ_α =5.6^o [14] in the UNIFIED model to simulate a rough surface. With an unpolished surface, the scintillation light reaches an APD pixel after being internally reflected on the many micro-facets. Each micro-facet creates its own virtual source. The result is a cluster of virtual sources. Since a cluster of mirror sources is difficult to model, we will still use a single mirror source to represent the cluster of sources. The position of this mirror source is identical to the one used for polished surfaces. Table 5.7 compares the spatial resolution obtained for a polished versus an unpolished LSO block, using the approximate solid angle model with 4 mirror sources. It can be concluded that, under the assumption of using point-like virtual sources, polishing all surfaces has nearly no influence on the global spatial resolution.

	X-direction		Y-direction		DOI-direction	
	FWHM (mm)	FWTM (mm)	FWHM (mm)	FWTM (mm)	FWHM (mm)	FWTM (mm)
polished	1.10	3.98	1.20	4.41	1.91	5.49
unpolished	1.14	4.20	1.33	4.39	1.82	6.01

Table 5.7 – Influence of surface polishing on the resolution.

5.2.8 Local resolution and bias

To study how the resolution and the bias vary across the monolithic block, the crystal volume was divided in 1 mm intervals in X direction and 2 mm in DOI direction. No division was made in the Y direction, i.e. the Y coordinate was not taken into account. For each of the $1(X) \times 20(Y) \times 2(Z)$ mm^3 volumes, X and DOI resolution profiles were extracted using simulation data, followed by the computation of the FWHM and the bias. The simulation model used polished block surfaces, 1800 electrons ENC and included multiple interaction events. Figure 5.5 shows the results as color-coded images. For symmetry reasons only half of the X coordinate range was studied.

In general, the X resolution degrades as the interaction point moves further away from the APD and is the worst in the top corner. In the lower third of

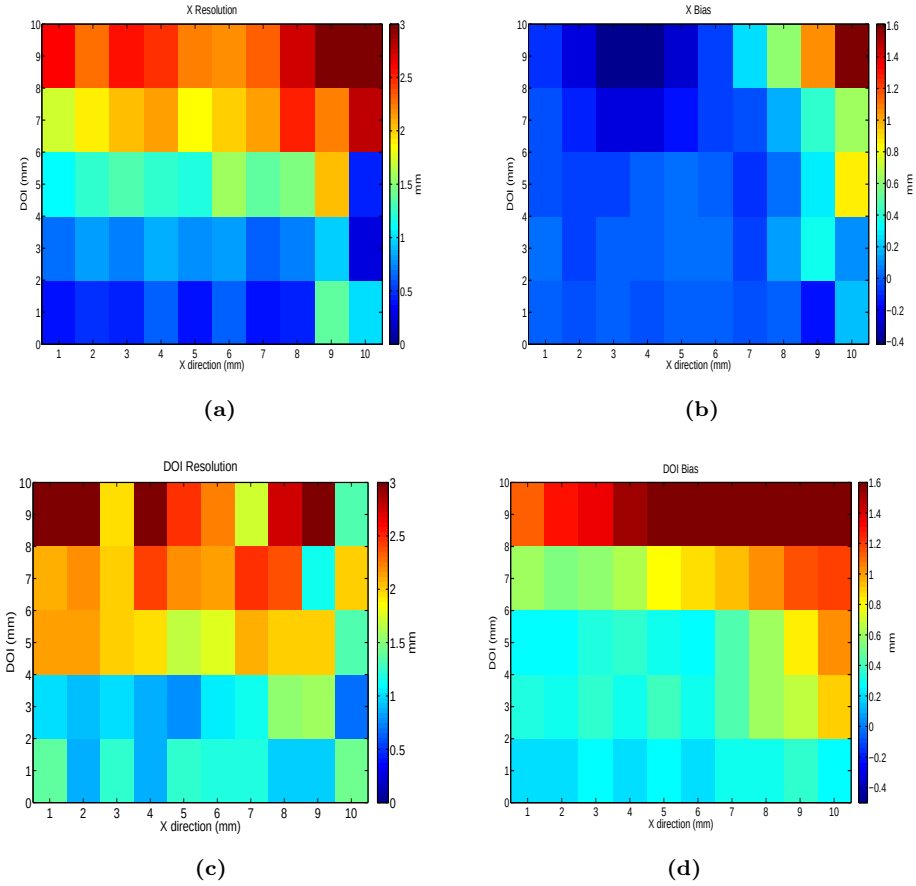


Figure 5.5 – X-resolution (a), X bias (b), DOI resolution (c) and DOI bias (d) as a function of the X-position and DOI of the true interaction point. The DOI coordinates go from bottom (0 mm) to top (10 mm) and the X coordinates go from the crystal center (0 mm) to the crystal edge (10 mm).

the block, the resolution is sub-millimeter. However, there is rapid resolution degradation in the upper two millimeter of the LSO block.

The X bias is usually small except for interactions within 1 mm from the edge, especially in the top corner. Here, the events are localized closer to the center compared to the true interaction point. Towards the top-middle of the crystal, there is a region where the X bias is slightly negative. This needs to be looked into further.

The accuracy of the predicted interaction depth also degrades for interactions further away from the APD. In the lower part of the block, the DOI resolution is better than 1.5 mm FWHM. The bias on the DOI coordinate is unidirectional towards the bottom, i.e. always positive. In about 75% of the block volume, it is around 0.5 mm and it increases when either approaching the side or top surface. Hence, the predicted interaction position should be moved 0.5 mm towards the top. This compensates for the DOI bias of events that interact more than 2 mm away form the side or top. The DOI bias of the remaining events is only partially compensated. An exact compensation of the DOI bias for the edge events is not so easy because the DOI resolution of these events is larger than their systematic miss-positioning along the DOI direction.

5.2.9 Reduction the electronic readout channels

Up to now we used all 64 APD pixel signals in the NLS modeling process. In order to simplify the read out electronics, one could think of adding a summing preamplifier which adds the signals from the rows and columns. Hence the number of read out channels reduces from 64 to 16. In this case the model used in the NLS fitting is changed appropriately, i.e. the solid angles are now the sum of the solid angles of all pixels in one row or column.

The results of using this new readout geometry are shown in table 5.8. The resolution gets about 30% worse in the X dimension and DOI dimension while the Y direction suffers by a 50% resolution increase. This shows that in order to use a simpler electronic readout, a compromise on the spatial resolution has to be made.

5.2.10 Impact of the size of the APD's

When the APD pixel size would be reduced, we have more data points to fit the model to. On the other hand, smaller pixels collect less scintillation light and

	X-direction		Y-direction		DOI-direction	
	FWHM (mm)	FWTM (mm)	FWHM (mm)	FWTM (mm)	FWHM (mm)	FWTM (mm)
Individual channels	1.10	3.98	1.20	4.41	1.91	5.49
Sum row and column	1.45	5.13	1.83	4.87	2.61	7.94

Table 5.8 – Influence of reducing the number of electronic readout channels.

hence the SNR per pixel will be worse.

To test whether smaller pixels could improve the performance of our NLS method, a virtual APD with 20×20 pixels measuring 1×1 mm each was simulated. The size of the LSO block was kept at $20 \times 20 \times 10$ mm. It was assumed that no gaps exist in-between the pixels, i.e. 100% packing fraction.

When no electronic noise is added, the simulated resolution in the X direction improves from 0.74 mm FWHM in case of the real APD to 0.62 mm for the virtual APD. However, the DOI error distribution using the virtual APD has a strange shape (figure 5.6(a)) compare to the DOI error distribution obtained with the real APD (figure 5.6(b)): when an event interacts in the top part of the LSO block, the scintillation light spreads most and the measured light distribution is very noisy for the virtual APD with smaller pixels. An example is shown in figure 5.7(a) for an event interaction at a depth of 3.863mm (from the top surface). A similar event detected by the real APD with the 1.6x1.6 mm pixels generates a less noisier light distribution (5.7(b)) .

Therefore making the pixels smaller does not necessarily improve the performance of our NLS algorithm due to the degradation of the photon statistics per pixel which makes the modeling less accurate. Optimizing the pixel size as a function of the light yield and crystal thickness could be part of future optimizations.

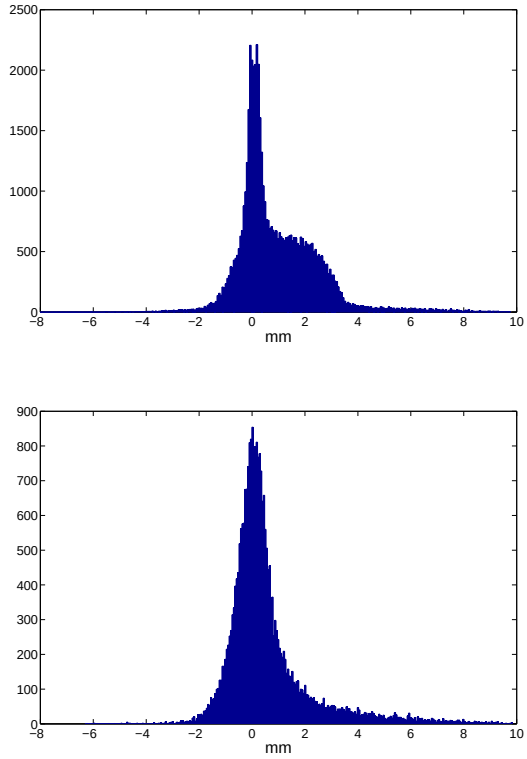


Figure 5.6 – (a) the DOI error distribution using the virtual APD (b) the DOI error distribution using the real APD

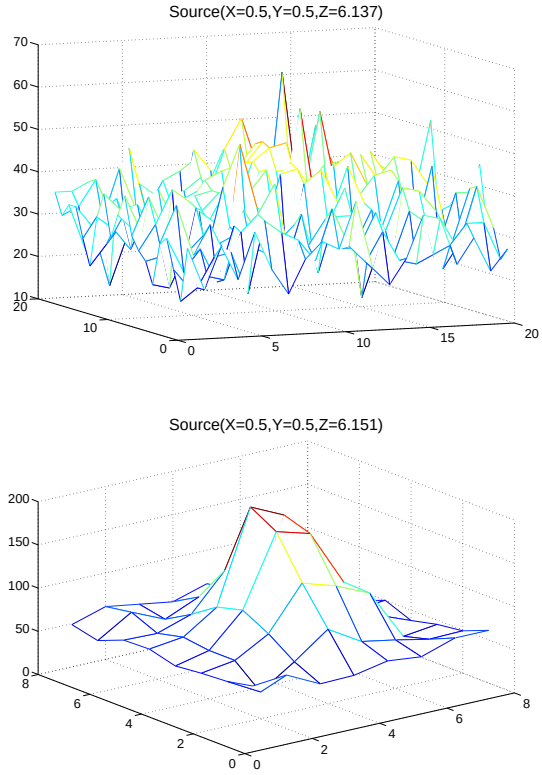


Figure 5.7 – (a) an example of light spread for virtual APD when the interaction is at the top part of the crystal (3.863mm from the top surface) (b) similar example of light spread for real APD.

5.3 Experimental results

5.3.1 Acquiring experimental modeling data

To shoot 511keV photons at a known position onto a polished LSO block, a narrow beam is defined by geometrical collimation using a ^{22}Na point source (ϕ=0.5 mm) and a 2x2x5 cm Bismuth Germinate (Bi4Ge3O12 or BGO) scintillator mounted on a photomultiplier tube (PMT) at a distance of $\approx$ 105 cm from the source (figure 5.1). The peak sensing ADCs are triggered when both the PMT signal and the analog sum of the 64 APD signals cross a preset threshold level within a coincidence time. After digitization, the values of the 64 channels are read out by a Labview program.

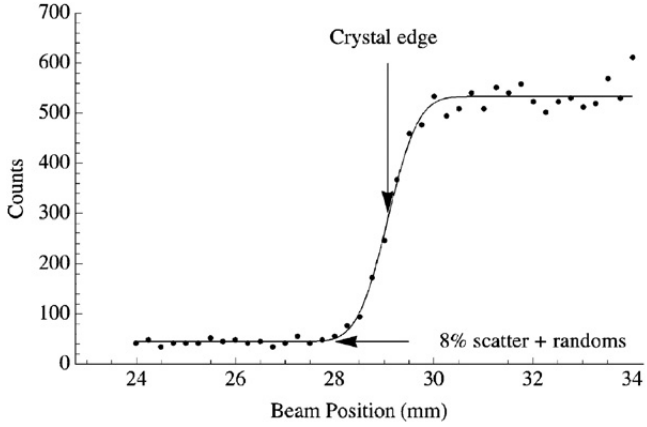


Figure 5.8 – Count rate profile corresponding to a 1.2 mm ± 0.2 mm FWHM 511 keV gamma beam being scanned over the edge of the LSO block.

The width of the beam was estimated by measuring a count rate profile when the beam is scanned over the edge of the block in 0.5 mm steps (figure 5.8). Assuming that a Gaussian distribution can describe the photon beam, the FWHM is obtained by fitting an erf(x) function through the count rate profile. This yielded a value of 1.2 ± 0.2 mm FWHM. The position halfway the rising edge of the erf(x) function was defined as the edge of the LSO block. The background events due to scattered and random coincidences correspond to 8 % of the total number of events measured when the beam is positioned on the block.

The detector box (white box in figure 5.1) is placed on a precision computer controlled X-Y stage, allowing automated scanning over the 20x20 mm LSO block

surface in 0.5 mm steps in both direction. At each of the 1600 beam positions, about 150 events were measured. Only events in a 380 keV – 630 keV energy window are accepted. The detector can also be rotated relative to the direction of the incoming gamma beam to vary the 511 keV gamma incidence angles on the LSO block.

5.3.2 X-Y resolution

The X and Y positions are estimated using the approximate solid angle model and then compared with the known beam position. The FWHM and FWTM of the resolution profile in the X direction are 1.87 mm and 5.21 mm respectively. For the Y direction we obtained 2.51 mm FWHM and 6.48 mm FWTM. The resolutions obtained with experimental data are slightly higher than what would be expected from adding the effect of a 1.2 mm FWHM photon beam to the simulated resolutions, i.e. 1.62 mm FWHM in the X direction and 1.70 mm FWHM in the Y direction. This could be due to the 0.2 mm uncertainty on the measured beam width and the fact that the tails in the true beam profile are bigger compared to those expected from the Gaussian assumption.

The resolution along the Y direction is worse than along the X direction. One of the reasons is that the APD pixels are not evenly distributed in Y direction. This trend is also visible in the simulation results (Table 5.2). Another reason is that the manual mounting of the two APDs on the LSO block resulted in a slightly different geometry than the one used in the model (figure 5.9), i.e. the coordinates (x_i, y_i) of the i^{th} pixel center are not at the same position relative to the LSO block as those used in the model. This was specifically the case in the Y direction.

Figure 5.10 shows the local FWHM and bias for beam positions in 1 mm intervals along the X-axis for all beam positions at least 2 mm from the edge. Hence we used 150×80 events to generate each of the local resolution profiles. The average local FWHM resolution is 1.7 mm and is rather constant along the length of the LSO block. In this region the bias also remains small. For beam positions within 2 mm from the edge, the resolution profiles become very non-Gaussian and sometimes exhibit multiple peaks (figure 5.11) because of the boundary constraints imposed on the cost minimization during the fitting procedure. For these distributions, the FWHM representation is not very indicative.

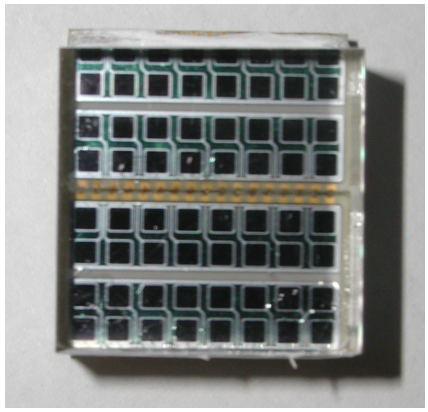


Figure 5.9 – The manual mounting of the two APDs on the LSO block.

5.3.3 DOI resolution

Contrary to the simulations, the true DOI coordinate is unknown in the experimental data. The DOI position of an interaction was obtained indirectly by rotating the detector over 45^o. When a 511 keV photon impinges at a 45^o angle, the DOI is equal to the difference ΔX between the coordinate X_0 where it enters the LSO block and the coordinate X where it interacts (figure 3.8 on page 75). Hence, the distance ΔX is obtained from the known beam incidence point and the predicted X coordinate using the approximate solid angle model. Comparing the obtained ΔX with the predicted DOI coordinate yields a DOI resolution of 3.4 mm FWHM and 7.61 mm FWTM. There was also a systematic error on the estimated DOI of -0.3 mm. These results contain a number of degrading effects:

1. Uncertainty on the estimation of interaction X coordinate, i.e. the resolution in the X direction.

2. A larger beam width of 1.2mm $/cos(45^o)$ = 1.70 mm FWHM due to the inclined incidence of the photon beam.

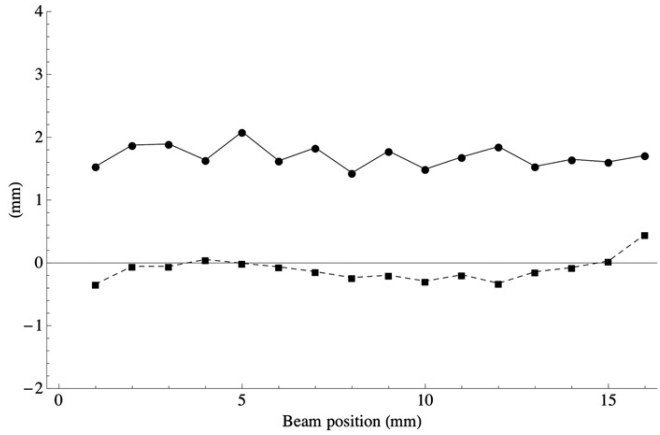


Figure 5.10 – Local FWHM (full line) and bias (dashed line) in the X direction obtained from experimental data using the approximate solid angle model.

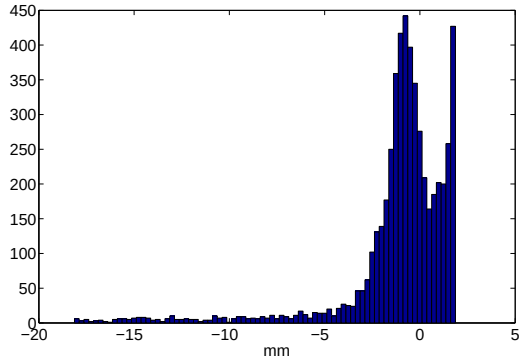


Figure 5.11 – Multiple peaks in the x-estimation error histogram, where the beam is 1.5mm from the edge in X direction.

3. Uncertainty on the incidence coordinate X_0 relative to the edge.

4. Beam angle might not be perfectly at 45^o.

The first two effects add to the FWHM/FWTM of the DOI resolution. If the X resolution is adjusted to take the larger beam width into account by subtracting it from the obtained DOI resolution, we estimate the intrinsic DOI resolution to be around 2.6 mm FWHM. The last two degrading effects add an additional bias to the DOI estimation. This could explain the difference in bias observed between the DOI estimates from simulated data (systematic error towards the bottom)

and the DOI bias using experimental data (systematic error towards the top).

5.4 Conclusion

Using simulation data, two models have been studied: exact solid angle based and approximate solid angle based. The approximate solid angle based model yielded similar results to the exact solid angle based model but requires less computing time. A weighted least squares method to fit the model to the data was slightly inferior to a normal least square fitting.

Adding a critical angle effect to the model made the fitting of the model very sensitive to the initial values and it became easily trapped into local minima that appeared in the cost function.

From the local study of the resolution and bias in X-direction and DOI, it follows that the performance improves as the interaction occurs closer to the APD. Hence a detector design with the APD on the front-side of the LSO block seems more appropriate since more interactions occur in the upper part of the crystal [64]. Edge effects only seem to manifest themselves significantly within 1 mm from the edge.

In our current experimental setup, the electronic noise is the limiting factor. Reducing the electronic noise from the 1800 electrons ENC to around 600 electrons ENC would clearly improves the resolution. Using a model with point-like virtual sources, it was shown that unpolished crystal surfaces yield similar performances than fully polished crystal surfaces.

Applying the approximate solid model with 4 virtual sources to experimental data resulted in an average resolution over the complete block of 1.85 mm FWHM in the X direction and 2.5 mm FWHM in the Y directions and 3.4 mm FWHM DOI. If the influence of the 1.2 mm FWHM beam size is subtracted from these results, the intrinsic resolution in X and Y direction becomes 1.4 mm and 2.2 mm FWHM respectively. An evenly distributed APD geometry and more accurate mounting of the LSO block on the APD matrices will probably improve the resolution in Y direction. The estimated intrinsic DOI resolution is estimated at 2.6 mm FWHM.

The model used on the experimental results did not take inter-pixel gain variations into account. Measuring the individual gain for each APD pixel and normalizing the signals accordingly, could potentially further improve the achieved experimental resolutions.

Chapter 6

Comparison of Algorithms

This chapter compares our nonlinear least square fitting algorithm (NLS) with positioning approaches investigated by other groups. The first part of this chapter gives some details about the implementation. In the second part, these algorithms will be applied on our simulation and measurement data to compare their performance. In addition, an extra comparison is done with an ML clustering method (section 3.2.3) performed by another group using a crystal of a different size and a different photo detector.

6.1 Alternative positioning algorithms

The alternative algorithms which we will compare with include Maximum likelihood (ML) maximization based methods, a Neural Network (NN) approach and the theoretical Cramér-Rao lower bound (CRLB). The implementation of the NN approach was already detailed in chapter 3. The following paragraphs in this section explain how the other methods were practically implemented for the comparison analysis.

6.1.1 3D-ML position estimation

In simulation data, the type of interaction (single or multiple) and the exact interaction position are known. We will take advantage of this information and use an ideal 3D-ML algorithm on the simulation data in order to compare the best results from 3D-ML to NLS. First, no electronic noise was considered. Next, a comparison will be done including the effect of the electronics.

6.1.1.1 Implementation in the absence of electronic noise

The research groups that investigate the ML positioning methods assumed that the number of photons g_m detected by the m^{th} photo detector pixel are Poisson distribution [34]. The likelihood of observing a set of signals $\mathbf{g} = \{g_m, m = 1 \cdots M\}$ as a function of the gamma interaction position R and gamma energy E is then given by the following multivariate Poisson equation (See chapter 3):

$$pr(\mathbf{g}|\mathbf{R}, E) = \prod_{m=1}^{M} \frac{\left[\overline{N}_m(\mathbf{R}, E)\right]^{N_m}}{N^m!} exp\left[-\overline{N}_m(\mathbf{R}, E)\right] \tag{6.1}$$

with

$$N_m \equiv u\left({}^{g_m}/_{G_m}\right), \ \overline{N}_m \equiv {}^{\langle g_m(\mathbf{R},E)\rangle}/_{G_m} \tag{6.2}$$

and where $\langle \cdots \rangle$ is an ensemble average, $u(\cdots)$ is a rounding operator and G_m is the gain of the m^{th} photo detector pixel. Since $\overline{N}_m(R, E)^{N_m}$ can become extremely large, the logarithm of the likelihood is usually used. Therefore, the estimation of the maximum likelihood becomes:

$$\begin{aligned}
(\hat{x}, \hat{y}, \hat{z})_{ML} &= \underset{(x,y,z)}{argmax}\, log\{pr(\mathbf{g}|\mathbf{R}, E)\} && (6.3)\\
&= \underset{(x,y,z)}{argmax}\, log\{\prod_{m=1}^{M} \frac{\left[\overline{N}_m(\mathbf{R}, E)\right]^{N_m}}{N_m!} exp\left[-\overline{N}_m(\mathbf{R}, E)\right]\} && (6.4)\\
&= \underset{(x,y,z)}{argmax} \sum_{m=1}^{M} \left(N_m log\overline{N}_m - \overline{N}_m - logN_m!\right) && (6.5)\\
&= \underset{(x,y,z)}{argmax} \sum_{m=1}^{M} \left(N_m log\overline{N}_m - \overline{N}_m\right) && (6.6)
\end{aligned}$$

$\overline{N}_m(R, E)$ is the mean number of detected photoelectrons and they should be pre-calibrated for each scintillator in order to compute the ML. If the pre-calibrated look-up table of $\overline{N}_m(R, E)$ is only in two dimensions (x, y), the ML will give an estimate of the 2D spatial position. If entries in the look-up table are obtained for R= (x, y, z), the interaction position can be estimated in 3D.

Parametric calibration of $\overline{N}_m(R, E)$ It is impossible to calibrate $\overline{N}_m(R, E)$ in the z (DOI) dimension directly in an experiment since one cannot control the depth of interaction of a gamma ray. An approach to overcome this problem is

proposed in [34]. Here $\overline{N}_m(\{x,y,z\},E)$ is presented as a polynomial function of the DOI coordinate z (3.6), and the coefficients of the polynomial function are calculated using Maximum Likelihood estimation from all the collected data using prior knowledge of the distribution of the gamma rays path lengths in the scintillator (3.11). This distribution is defined relying on simulation data.

It is not the goal of this thesis to repeat the work of [34]. Therefore, we will compare the 3D-ML with our NLS method using simulation data, where the ML estimation of coefficients in the polynomial function can be done using the known DOI coordinate. For each beam position (x,y) and for each photon detector pixel m, the coefficients along the DOI direction z are estimated using

$$\hat{c}_{mn} = \underset{c_{mn}}{argmax} \sum_{j=1}^{J} \left(N_m log \overline{N}_m - \overline{N}_m \right) \tag{6.7}$$

$$= \underset{c_{mn}}{argmin} \sum_{j=1}^{J} \left(\sum_{n=0}^{3} c_{mn} z^n - N_m log \sum_{n=0}^{3} c_{mn} z^n \right) \tag{6.8}$$

where J is the total number of single interaction events collected at a given beam position. Here, we used a third order polynomial to describe the depth dependence of $\overline{N}_m$, i.e.

$$\overline{N}_m = \sum_{n=0}^{3} c_{mn} z^n \tag{6.9}$$

The polynomial parametrization as a function of DOI is only valid in case the impinging gamma only interacted once in the scintillator. In a simulation, these events can be easily selected. In an experimental setting, it is difficult to distinguish single and multiple interactions. Compton filtering is one strategy to obtain only single interaction events in measurement data [34]. This is based on the fact that a secondary interaction often occurs away from the beam axis. Therefore, a maximum likelihood threshold can be used to filter out all the events which are estimated off-center from the beam axis. This won't of course be perfect. Here, the 3D-ML will only be compared with the ideal case of single interactions.

To find the values $\hat{C}_{mn}$ which minimize the unconstrained multivariate function (6.8), we use the function "fminunc" build in the Matlab Optimization toolbox. The algorithm used in this function is based on the interior-reflective Newton method described in [17, 18].

The initial value $\hat{c}^0_{mn}$ used as a starting point in the search for the optimal value of $\hat{c}_{mn}$ is calculated from a simple linear least squares function, where for

each beam position (x, y) and for each photon detector m, the value of N_{mj} from the j^{th} event is equal to

$$N_{mj} \quad = \quad \overline{N}_{mj} + e = \sum_{n=0}^{3} c^0_{mn} z^n_j + e$$

where e is the error between the value of $\sum_{n=0}^{3} c^0_{mn} z^n_j$ and the measured N_{mj}. Therefore, the estimated initial value of $\hat{c}^0_{mn}$ is:

$$\hat{c}^0{}_{mn} \quad = \quad \mathbf{Z} \setminus \mathbf{N_m}$$

where

$$\mathbf{Z} = \begin{bmatrix} 1 & z_1 & z_1^2 & z_1^3 \\ 1 & z_2 & z_2^2 & z_2^3 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & z_J & z_J^2 & z_J^3 \end{bmatrix} \; and \; \mathbf{N_m} = \begin{bmatrix} N_{m1} \\ N_{m2} \\ \vdots \\ N_{mJ} \end{bmatrix}$$

The searching method requires the user to supply the gradient of the objective function. Our objective function (6.8) is a sum composed by J sub-parts, hence the gradient is also the sum of the J sub-gradients.

Figure 6.1 shows an example of how well the polynomial function (6.9) can present $\overline{N}_m$ once the optimal values of $\hat{C}_{mn}$ have been found. The beam position is at the point $(x, y) = (5, 5)$ in figure 6.1 (e), and four examples of the fittings from channel 19, 39, 47, 59 are shown in figure 6.1 (a), (b), (c), (d) separately. The corresponding channel positions are illustrated in 6.1 (e).

We notice that there is one outlier in figure 6.1 (a) and (b), whose DOI is 7.582. The reason is that this is data from an event whose first interaction is a Rayleigh Scattering (also called Coherent Scattering) at position $(5, 5, 7.582)$ in the crystal, followed by a photo electric absorption at (5.304, 5.721, 1.448), which is 1.4 mm from the bottom. A Rayleigh Scattering is an elastic interaction without energy loss but with momentum change (see Appendix A). Hence, all the energy is deposited in the second interaction place close to the APD array. That's why pixel 47 closest to the interaction side receive more light, and the neighboring pixel 39 has less light than expected.

The chance that this kind of events can occurs in an LSO crystal is 5% [67] and according to the simulation data based on our setup, 2% of the total detected

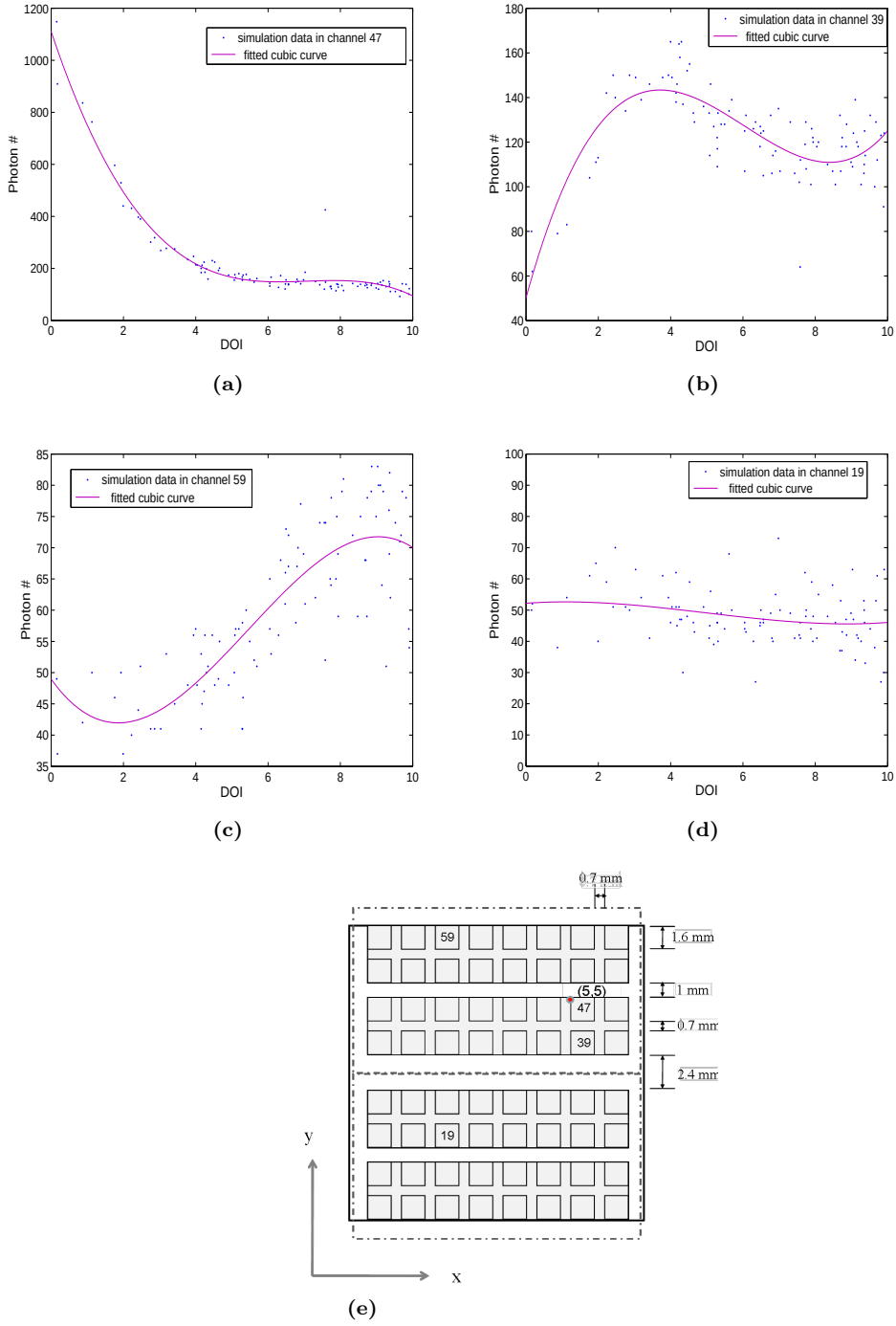


Figure 6.1 – (a-d) Number of photons detected in pixels 47,39,59 and 19 as a function of the DOI of each of the 150 events impinging at position (5,5). The red line is the corresponding third order polynomial fit. (e) Schematic showing the beam impinging position (red dot) relative to the four pixels examined.

events first had a Rayleigh Scattering and then underwent a Compton Scatter or Photo electric absorption. The remaining 3% escapes without a second interaction.

Interpolation of c_{mn} The coefficients c_{mn} are calculated for each photo detector m and for each beam position (x, y). A beam step size of 0.5 mm between -9.5 and 9.5 mm, 64 pixels and 4 coefficients for each polynomial to represent the DOI dependent $\overline{N}_m$, results is a $19 \times 19 \times 64 \times 4$ look-up table (LUT) of C_{mn}. The 3D-$\overline{N}_m$ at those specific positions (x, y) present in the LUT can hence be computed.

To obtain a 3D-$\overline{N}_m$ at any other position (x, y), the LUT is interpolated. To make the Grid searching method efficient (3.2.1), a discontinues interpolation of C_{mn}, using a cubic spline method, in steps of 0.125mm in x and y dimensions has been done. The dimensions of C_{mn} therefore enlarge to $73 \times 73 \times 64 \times 4$. The precision of the estimation based on the interpolated LUT is in the order of 0.125mm. This is acceptable since the spatial resolution are not better than 0.5 mm.

6.1.1.2 Implementation in the presence of electronic noise

Adding electronic noise to the simulation data makes the results more realistic, especially for APD photo detectors since the internal gain of APD is normally only around 100. Compared to PMT gains of up to a few 10^6, APDs are more sensitive to the electronic noise. If electronic noise is dominant, a Gaussian distribution is a better choice over a Poisson distribution to present the detected signals. The likelihood to detect a signal g_m then becomes:

$$pr = \prod_{m=1}^{M} \frac{1}{\sqrt{2\pi}\sigma_{g_m}} exp\left(-\frac{(g_m - \overline{g}_m)^2}{2\sigma_{g_m}^2}\right) \tag{6.10}$$

with

$$\overline{g}_m \equiv \langle g_m(\mathbf{R}, E)\rangle \tag{6.11}$$

and where $\langle \cdots \rangle$ is an ensemble average, σ_{g_m} is the standard deviation of g_m, M is the number of pixels in the photo detector. The estimation of the position then becomes

$$(\hat{x},\hat{y},\hat{z})_{ML} = \underset{(x,y,z)}{argmax}\, log\{pr(\mathbf{g}|\mathbf{R},E)\} \quad (6.12)$$

$$= \underset{(x,y,z)}{argmax}\, log\{\prod_{m=1}^{M}\frac{1}{\sqrt{2\pi}\sigma_m}exp\left(-\frac{(g_m-\overline{g}_m)^2}{2\sigma_m^2}\right)\} \quad (6.13)$$

$$= \underset{(x,y,z)}{argmax}\sum_{m=1}^{M}\left(log\frac{1}{\sqrt{2\pi}}-log\sigma_m-\frac{(g_m-\overline{g}_m)^2}{2\sigma_m^2}\right) \quad (6.14)$$

$$= \underset{(x,y,z)}{argmin}\sum_{m=1}^{M}\left(log\sigma_m+\frac{(g_m-\overline{g}_m)^2}{2\sigma_m^2}\right) \quad (6.15)$$

Parametric calibration of $\overline{g}_m$ and σ_m Two ways to present $\overline{g}_m$ and σ_m as a function of DOI have been investigated. One is similar to the method used in the previous section, i.e.

$$\overline{g}_m = \sum_{n=0}^{3} c_{g_{mn}} z^n$$

and

$$\sigma_{g_m} = \sum_{n=0}^{3} c_{\sigma_{g_{mn}}} z^n$$

The alternative approach is to represent σ_{g_m} using:

$$\sigma_{g_m} = \sqrt{c_{\sigma m_0} + \frac{c_{\sigma m_1}}{z^2+1}} \quad (6.16)$$

where z is the DOI coordinate. The underlying motivation is that the variance changes according to the amount of detected photons. This amount is proportional to the solid angle, which is approximately proportional to $1/z^2$ (4.20). The $+1$ term in the denominator is added to ensure that the function does not diverge when z comes close to the bottom of the crystal, i.e z$\rightarrow$ 0.

However, for both ways, the searching process always converged to a point very close to any given starting values. Therefore, no results were obtained for a Gaussian model based 3D-ML method using simulation data with electronic noise. A better model to represent $\overline{g}_m$ and σ_m and more calibration data are required to perform a Gaussian based 3D-ML method. Therefore, we will use a 2D-ML (only in (x,y) dimensions) to compare with NLS and NN methods.

6.1.2 2D-ML position estimation

In the presence of electronic noise, the generation of a look-up table of the mean and variance at each beam position is much simpler for 2D-ML then for the 3D case. Since the two dimensions (x, y) of the beam position are known, the mean and variance can be generated directly from the data.

An example of the signals of the 64 channels for a central beam position are plotted in figure 6.2. Channel 17, 26 and 64 are the three bad channels, which are not used in ML method. We can see that four channels (28, 29, 36, 37 within the blue square), which are near the source, are not really Gaussian shaped due to the varying DOI. The signal histograms of the other pixels are of a Gaussian shape. The mean and variance of the 61 channels are computed and saved in look-up table. This process is repeated for each of the 40×40 calibration positions.

After the look-up table has been built, it can be used to estimate the 2D position of a new perpendicular impinging event using equation (6.15):

$$(\hat{x}, \hat{y})_{2DML} = \underset{(x,y)}{argmin} \sum_{m=1}^{61} \left(log\sigma_m + \frac{(g_m - \overline{g}_m)^2}{2\sigma_m^2} \right) \tag{6.17}$$

where g_m is the signal of m^{th} channel, and $\overline{g}_m$ and σ_{g_m} are taken from pre-calibrated look-up table.

6.1.3 Cramér–Rao lower bound

One can derive a CRLB to express the smallest value of the standard deviation any unbiased estimator can achieve, i.e. the lower bound is independent of the estimation algorithm [71]. In case of 2D position estimation, the lower bound shows us the theoretical best performance for estimating the coordinates (x, y) from a known set of APD outputs. This lower bound can be used to verify whether a specific estimation algorithm is efficient. If the standard deviation of our estimator is close to the CRLB, it means the estimation algorithm is doing well. Here we check the CRLB for an additional purpose, i.e. to confirm the difference of the spatial resolution in the x and y direction obtained from experimental data using either 2D-ML method, NLS method or NN methods (as we shall see in the results in table 6.3).

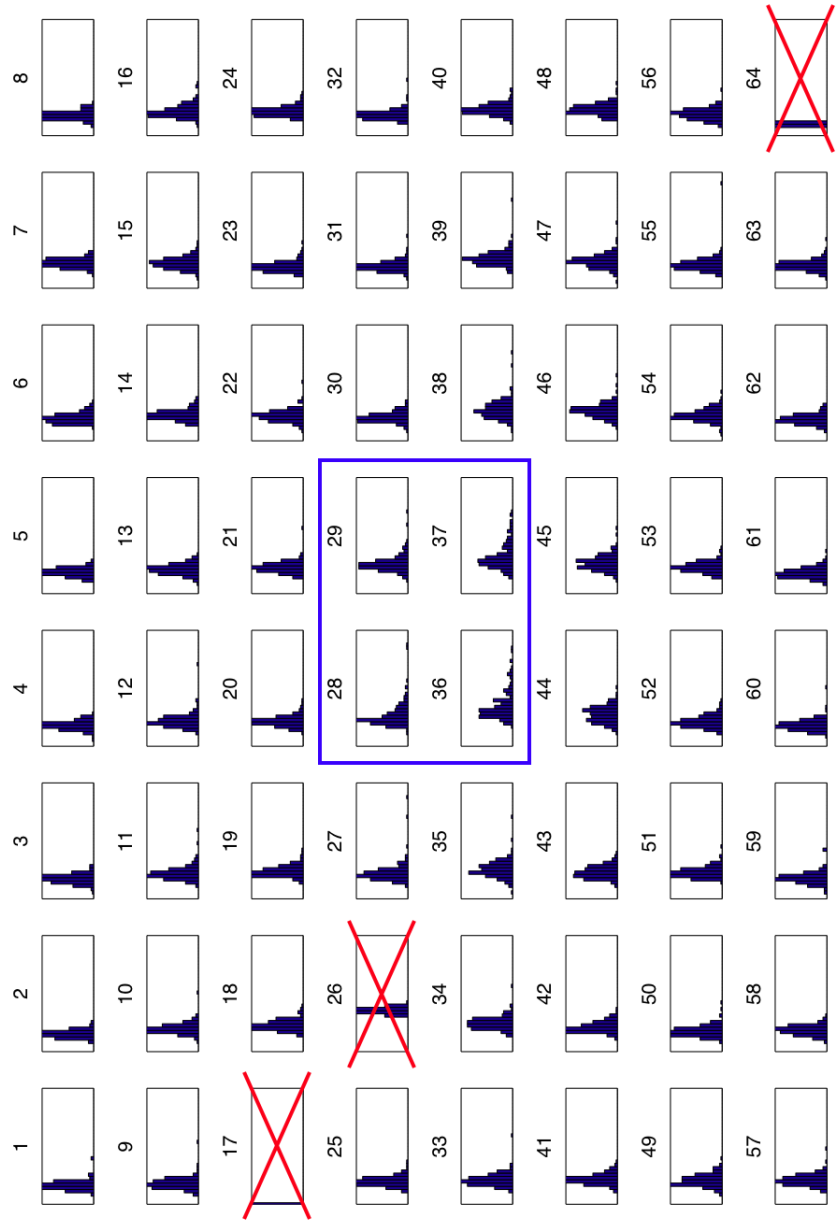


Figure 6.2 – Signals of 64 channels for beam position at the central. Channel 17, 26 and 64 are bad channels. The mean and variance of other channels are computed to save in look-up-table. This is repeated for each of the 40×40 calibration positions.

The CRLB of any unbiased estimator is calculated from the Fisher information matrix I [70].

$$I_{v,w}(x,y) = -\sum_{m=1}^{M} E\left[\frac{\partial^2 ln\left(pr\left(g_m\right)\right)}{\partial v \partial w}\right] \tag{6.18}$$

where E is the expectation value, v, w are equal to x or y, g_m is the signal detected in the m^{th} APD, and pr the likelihood to detect a signal g_m given by equation (6.10), where the mean $\overline{g}_m$ and variance σ_{g_m} of the beam position (x, y) are calculated from measurement data. The CRLB for e.g. the x coordinate is given by:

$$\sigma_x \geq \sqrt{I_{xx}^{-1}(x,y)} = \sqrt{\frac{I_{yy}(x,y)}{I_{xx}(x,y)I_{yy}(x,y) - I_{xy}^2(x,y)}} \tag{6.19}$$

Substituting (6.10) in (6.18), the information matrix becomes [70]

$$I_{v,w}(x,y) = \sum_{m=1}^{M} \frac{1}{2\sigma_{g_m}^2}\left(-\frac{1}{\sigma_{g_m}^2}\frac{\partial \sigma_{g_m}^2}{\partial v}\frac{\partial \sigma_{g_m}^2}{\partial w} + 2\frac{\partial \overline{g}_m}{\partial v}\frac{\partial \overline{g}_m}{\partial w}\right) \tag{6.20}$$

In our experiment, the training data contains about 150 events per beam position. For each beam position (x, y) the mean $\overline{g}_m$ and variance σ_{g_m} for each APD are therefore calculated from these 150 events. Since the data contain noise, they should be smoothed first before calculating the derivatives numerically using:

$$\frac{\partial \overline{g}_m(x,y)}{\partial x} = \frac{\overline{g}_m(x+h,y) - \overline{g}_m(x-h,y)}{2h} + O(h^2) \tag{6.21}$$

where $h = 0.5mm$.

6.2 Results

6.2.1 Comparing all algorithms using simulation data

We first compare the performance of the different algorithms using Monte Carlo data. Because the NN method (see Chapter 3) yields the 2D impact position on the surface, and the 3D-ML method did not work in the presence of additive noise (i.e. when we have to assume a Gaussian signal distribution), the comparison is limited to the 2D achievable resolution in the x and y direction. The data used in the different algorithms includes 1800 e^- ENC electronic noise. At each beam position, 300 events are simulated which deposit more than 350keV energy.

Table 6.1 shows that the theoretical lower limit given by the 2D CRLB is just below 1 mm for the x direction and a little above 1 mm for the y direction. Hence, it looks that the non-regular distribution of the APD pixels in the y direction causes some loss of information, resulting in a slightly inferior resolution.

The NLS algorithm is closest to the CRLB. The 2D-ML and NN algorithms are significantly worse. This seems to indicate they are more sensitive to noise and the impact of multiple interactions than the NLS method. Also note that the NN method is not sensitive to the different light sampling pattern in x and y direction.

With electronic noise	X-direction	Y-direction
	FWHM (mm)	FWHM (mm)
2D CRLB	$0.87 \pm 0.32mm$	$1.16 \pm 0.65mm$
2D-ML	1.65	1.75
NN	1.76	1.75
NLS	1.10	1.20

Table 6.1 – FWHM resolution in x and y direction using CRLB, 2D-ML, NN and NLS applied on simulated data with electronic noise.

6.2.2 Influence of multiple interactions on ML and NLS performance

About half of the simulated events used in the comparison study undergo multiple interactions. To study their influence, the resolution achieved by the ML and NLS algorithms was determined in case of single interaction events and in case of multiple events.

For the 3D-ML method, two thirds of the single interaction events (i.e. 100 events from each beam position) is used to calibrate the coefficients c_{mn}. The resolution of the ML algorithm is then evaluated using either the last one third of the single interaction events, or all the multiple interaction events . As for the NLS method, all the single interaction or multiple interaction data are used to evaluate the algorithm.

Table 6.2 shows the 3D resolution components for the 3D-ML method based on Poisson model in section 6.1.1.1 and the NLS method for simulation data without additive electronic noise. We see that in the absence of multiple interactions,

the 3D-ML method achieves the best resolution and has no detectable bias on the DOI. The disadvantage of the 3D-ML algorithm is that it degrades more rapidly than the NLS approach when testing on multiple interaction events. This is not surprising since the 3D-ML approach used here is based on a model that is incorrect when more than one interaction occurred. Hence more parameters are required in the ML method to model the multiple interactions, but this will make the model complicated and more research is required [34].

One solution would be to filter out the multiple events [34] and only use single interaction events. This will obviously result is a significant loss of sensitivity of about 50% in our detector.

No electronic noise	X-direction		Y-direction		DOI-direction		
	FWHM (mm)	FWTM (mm)	FWHM (mm)	FWTM (mm)	FWHM (mm)	FWTM (mm)	Bias (mm)
ML (single int.)	0.46	1.31	0.48	1.32	0.89	2.04	-0.06
ML (multiple int.)	1.09	3.05	1.20	3.05	1.41	4.95	-0.23
NLS (single int.)	0.52	1.65	0.66	1.51	0.98	2.85	0.32
NLS (multiple int.)	1.00	3.87	0.99	3.56	1.53	5.41	0.1

Table 6.2 – Global spatial resolution from ML and NLS comparison without electronic noise.

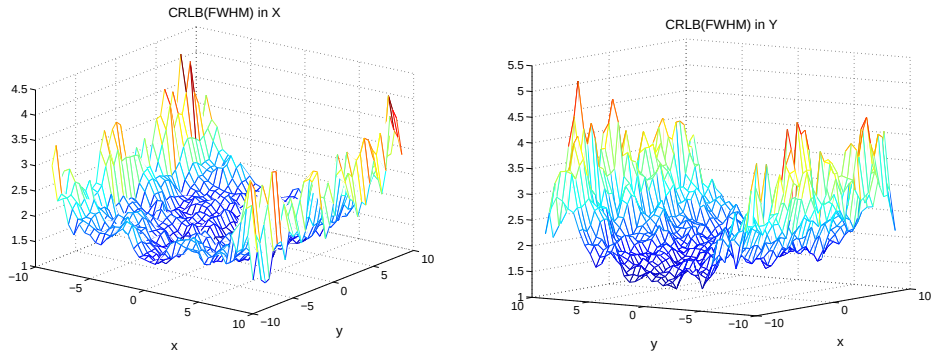
6.2.3 Comparing all algorithms using experimental data

Using the experimental data, only the signals from the 61 good channels are used to calculate the 2D CRLB, and evaluate the 2D-ML and NLS methods. As for the NNs, the 3 bad channels were set to be 0 before taking the sum of the rows and columns.

Figure 6.3 plots the FWHM CRLB on the resolution in the x and y direction, starting at 1 mm from the edge. The FWHM CRLB in x direction, averaged over the complete surface equals $1.84 \pm 0.47mm$, while in y direction it is $2.25 \pm 0.62mm$. Again we noticed a difference in the theoretical lower limit.

For the 2D-ML method, two third of the experimental data are used to build the look-up table. The rest is used to test the resolution. For the NN approach, about one tenth of the data are used to train the NN, and all others are used to test. The results are shown in table 6.3, together with the result from NLS. The

results from NLS and 2D-ML are quite similar, while NN has the best resolution in y direction but worse in the x direction.



(a) **(b)**

Figure 6.3 – CRLB(FWHM)

Experimental data	X-direction	Y-direction
	FWHM (mm)	FWHM (mm)
CRLB	$1.84 \pm 0.47mm$	$2.25 \pm 0.62mm$
2D-ML	1.82	2.55
NN	2.20	2.30
NLS	1.87	2.51

Table 6.3 – FWHM resolution in x and y direction using CRLB, 2D-ML, NN and NLS applied on experimental data. The contribution of the 1.2mm FWHM beam size is still included.

6.2.4 Comparison to 2D-ML+DOI clustering using experiment from another setup

It is difficult to compare with results obtained from with a crystal of a different size and read out by a different kind of photo detector. Here we shown one result that was recently reported in [45]. The crystal they are using is rather big, i.e. $50 \times 50 \times 15$ mm LSO. The photo detector is a Hamamatsu H8500 PMT coupled to the bottom of the crystal. The entrance surface of the crystal was painted

white but the sides of the crystal were painted black. Gamma rays impinged on the entrance surface of the crystal using a 0.6 mm FWHM beam.

The method they are using is a 2D-ML+DOI clustering algorithm, which has been introduced in section 3.2.3, where a look-up table of the means and variances of the Gaussian distributions needs to be built for 7 DOI regions, and a third-order polynomial fit is then applied to those means and variances. The fitting results are used to interpolate from 7 DOI layers to 15 DOI layers. The DOI estimation is therefore discontinues in steps of 1 mm.

Their results are shown in table 6.4 which is not a single number of the global resolution as we have been shown until now, but it is shown by the mean and standard deviation from all beam positions. The global resolution includes the local bias indirectly into the result.

	X-direction (mm)	Y-direction (mm)	DOI-direction (mm)
ML cluster (raw)	1.61±0.29	1.61±0.28	4.80±0.36
ML cluster (intrinsic)	1.49±0.30	1.49±0.32	4.48±0.37
NLS (raw global)	1.87	2.51	3.4
NLS (intrinsic global)	1.43	2.20	2.6

Table 6.4 – Spatial resolution with/without beam correction using ML cluster method, data of first two rows from [45].

If we subtract the beam size effect, (they have 0.6 mm FWHM beam, we have about 1.2 mm), the intrinsic resolution in the X direction become quite comparable. We have a worse Y resolution due to the incorrect mounting of the LSO block on the two APDs, and the irregularly scintillation light sampling geometry. Finally, due to our smaller crystal size, the impact of the biased evens near the edge have a larger relative contribution to the broadening of the global error histogram.

However, the NLS shows significant better resolution in the DOI direction. This could due to the fact that the size of the crystals are different (they have $50\times50\times15$mm, we have $20\times20\times10$mm). But also, the worse resolution in DOI in the 2D ML+DOI cluster method is due to the Compton scattering effects which are mainly forward [45]. The events contaminate some of the DOI estimating capabilities of the look-up table. As for the NLS method, since no pre-calibration is required, the position estimation is not effected by previous events.

6.3 Discussion and conclusion

In an ideal, noise-free world with single interaction events only, the 3D-ML Poisson model based method has better resolutions than the NLS method and shows no DOI bias. Its resolution degrades faster than with NLS when used on multiple interaction events. In order not to lose sensitivity by only using single interaction events, one hence has to compromise on the spatial resolution.

A comparision of our NLS method with a 2D-ML and a NN algorithm, using simulation data including electronic noise contributions showed that the NLS performance is closest to the theoretical lower limit given by the CRLB. The 2D-ML and NN methods are more sensitive to the presence of model errors (due to the pre-calibration or training) and of significant levels of electronic noise.

For experimental data, the CRLB in the X direction is 1.84 mm FWHM, and 2.25 mm FWHM in the Y direction. This means that the data contain less spatial information to estimate the y-coordinate than to estimate the x-coordinate of the interaction position. This trend is observed in all three methods: NLS, NN and 2D-ML.

The NN has the best global resolution of 2.3mm in Y direction , which hints that the NN is influenced least by the sampling geometry differences and by the imperfect mounting of the detector. The Y resolution of the other two methods is about 2.5mm. On the other hand, the NN has the worst resolution in X direction (2.2mm FWHM), while the other two algorithms achieve 1.82 mm FWHM (2D-ML) and 1.87 mm FWHM (NLS) which are close to the CRLB in X direction.

There are some other reasons that could explain why pre-calibrated 2D-ML and pre-trained NN are not better than NLS. Both NN and 2D-ML methods require the acquisition of calibration (or training) data sets. These data sets can get corrupted by bad events, i.e. events whose interaction location is not along the gamma beam used to generate the calibration data:

1. In our experimental data setup, about 2% of the events has first undergone an elastic Rayleigh scattering.

2. About another 8.6 % of the registered events are actually due to 1275 keV gamma which are also emitted during the decay of Na^{22} [73]. These high energy photons can undergo a Compton scattering in the LSO block and deposited an amount of energy that lies within the energy window imposed to select good events.

3. More than 8% events are from scatter or randoms events (5.8).

4. Finally, the presence of multiple interaction events in the calibration data set could degrade the performance of the trained algorithm.

Other factors that can influence the usage of pre-calibrated algorithms are varying detector conditions such as temperature or bias voltages. If these parameters differ between the moment the calibration data are taken and the moment the pre-calibrated algorithms are used, it can degrade the performance of the algorithms.

An additional comparison with results from another group using a 2D-ML+DOI clustering method has also been shown. The resolution along the X direction is very similar with the result obtained using the NLS method on our experimental data. In the DOI direction, they have a worse resolution of about 4.4 mm (intrinsic) compared to the NLS intrinsic DOI resolution of about 2.6 mm. The main reason of this difference, is that they need to build a look-up table for DOI positioning. The inclusion of multiple interaction events makes the DOI information less correct since most 511 keV gamma will be forward scattered after a first Compton interaction. The inaccurate look-up table will hence deteriorate the DOI position estimation.

A drawback of methods that need calibration data is that it takes a long time to obtain the training data. For our current setup, we need one day to get about 150 events in each of the 40×40 beam positions. This is even insufficient for 3D-ML and 2D-ML+DOI cluster methods. To speed up their data collection, they use multiple sources [51] to reduce the time it requires to record sufficient data for the 2D-ML+DOI cluster method to 6 hours. Since in theory, each crystal needs to be calibrated separately, the calibration of a whole PET system will be extremely time consuming. In addition, because PMT gains might change after some period and APD are sensitive to voltage and temperature, re-calibration of the detector from time to time might be required.

Chapter 7

Conclusion

Based on the target application area, each kind of PET scanner requires a different kind of performance. A pre-clinical (e.g. small animal) PET requires a very high spatial resolution ($< 1.5mm$), a whole-body human PET benefits more from TOF and a Brain-PET (or other specialized imaging systems, e.g. Positron Emission Mammography (PEM)) need both high spatial resolution and a very good sensitivity. The research presented in this work is targeted to dedicated Brain-PET systems, and maybe also pre-clinical PET systems, that can be used together with an MRI.

To improve the sensitivity, the ring of the PET system is normally made as small as possible. However, this also introduces more parallax errors which require DOI-measurement capabilities in the system. Therefore, a 3D positioning facility is also important. However, commercial PET systems rarely provide this DOI information.

The 2D spatial resolution of current PET systems based on discrete scintillator detectors is determined by the size of the individual crystals. Making the crystals smaller improves the spatial resolution but also reduces the SNR (which makes the image noisier), decreases the sensitivity, worsens energy and time resolution, and increases the cost and complexity. In addition, the existing methods to obtain DOI information in discrete detectors usually need extra hardware and are therefore not cost-effective.

In recent years, continuous detectors have shown to be an interesting alternative to simultaneously improve energy resolution and sensitivity without degrading the spatial resolution. Because the scintillation photons can travel freely over some distance within the crystal, the light distribution can be used to determine

the 3D gamma ray position. The challenge of this type of scintillator is the complexity of the extraction of the 511keV gamma ray interaction position. Numerous methods have been investigated to achieve this goal (see Chapter 3). However, most of them need a time-consuming calibration procedure. Considering there are in order of hundreds of crystals in one PET system, a method which does not need pre-calibration or training is very interesting.

The goal of this thesis is to investigate a method that does not need prior calibration data or extra hardware to obtain the 3D coordinates in one step and this with an accuracy equal to or better than other existing methods.

To achieve this goal, a nonlinear least-squares (NLS) estimation method is used. The difference between the NLS method and the classic maximum likelihood (ML) method, is that we define a general model to describe the light distribution. Hence, the 3D interaction coordinates are parameters in the model, which can be estimated directly from the measured pixels signals using a numerical searching method. The classic ML method assumes that the amount of photons detected in each channel obeys a Poisson (or Gaussian) distribution and the mean of each distribution should be pre-calibrated and saved in a look-up table. Afterward, the gamma ray interaction position can be obtained using this look up table to check which position is most likely to have generated the detected signals.

The general model used in our NLS method is based on the solid angles subtended by the detector pixels as seen from the gamma ray interaction location and from the additional virtual light sources mirrored around the surfaces of the crystal. Those mirror sources are used to take internal reflections into account. A constant parameter is also included in the model to represent the background light, mainly due to the reflection of the optical photons on a diffuse reflector (e.g. Teflon) surrounding the scintillator block. The parameters (i.e. the gamma ray position (x, y, z), the amplitude and the constant parameter) are estimated using a interior-reflective Newton minimization method with boundary constraints. This algorithm was implemented in Matlab.

The NLS method has been evaluated using both simulation and experimental data. The detector used in this study consists of a $20\times20\times10mm^3$ LSO scintillator block coupled to two Hamamatsu S8550 APD arrays. The APD arrays consist of 8×4 pixels, each measuring $1.6 \times 1.6mm^2$. The LSO crystal is polished on all sides and wrapped in Teflon. Simulation gave insight in how the model should be optimized, such as how many mirror source should be add, if the exact solid angle model can be simplified, the influence of the surface, etc. This is not easy

to check experimentally since the DOI dimension cannot be obtained directly.

After the verification using simulations, the optimized model is evaluated using experimental data. For perpendicular incident gamma rays, whose beam size is about 1.2mm FWHM, the measured global resolution in X direction is 1.87mm FWHM and 2.51mm FWHM in Y direction (beam effect still included). The resolutions obtained with experimental data are slightly higher than what would be expected from adding the effect of a 1.2 mm FWHM photon beam to the simulated resolutions, i.e. 1.62 mm FWHM in the X direction and 1.70 mm FWHM in the Y direction. This could be due to the 0.2 mm uncertainty on the measured beam width and the fact that the tails in the true beam profile are bigger compared to those expected from the Gaussian assumption. Another reason is that the manual mounting of the two APDs on the LSO block resulted in a slightly different geometry than the one used in the model. The third reason could be that the gain variance in experimental setup has some influence on the resolution.

The worse resolution in the Y direction is mainly due to the un-evenly distributed APD geometry (this trend is also visible in the simulation results) and the inaccurate mounting of the LSO block on the APD matrices (this was specifically the case in the Y direction).

By shooting the gamma ray beam at 45 degrees on the surface, DOI information can be obtained indirectly to evaluate the NLS estimation of the DOI coordinate. The measured resolution of the DOI is about 3.4mm FWHM, and the intrinsic resolution is estimated at 2.6 mm FWHM.

The same experimental data are used to compare the NLS method with the performance of a 2D ML algorithm, a NN approach and the CRLB. The latter is the theoretical lower bound on the achievable resolution, independent of the algorithm. The CRLB in the X direction is $1.84 \pm 0.47mm$ FWHM, and $2.25 \pm 0.62mm$ mm FWHM in the Y direction (including beam size effect), which indicates that the data contain less spatial information to estimate the y-coordinate than to estimate the x-coordinate of the interaction position. This trend is observed in all three methods: NLS, NN and 2D-ML. The 2D-ML achieves a resolution in X direction of 1.82 mm FWHM. Both the 2D-ML and the NLS methods provide a resolution very close to the CRLB in X direction, but their resolution in Y direction is about 2.5 mm FWHM.

The NN approach results in a worse resolution in X direction (2.2 mm FWHM), but a better resolution in Y direction (2.3 mm FWHM). This can be explained

by the observation that the NN is influenced least by the sampling geometry differences and by the imperfect mounting of the detector. Its worse performance in the X direction might be caused by the three damaged channels, which were not used in NLS and 2D ML method. To compute the sum of rows and columns for the input of NN, their value was set to 0.

There are some reasons that could explain why the pre-calibrated 2D-ML and the pre-trained NN are not better than the NLS. The multiple interactions, for example, add model errors; outlier calibration data will also degrade the model performance (i.e. the 2% Rayleigh scattering or less than 8.6% events in our setup due to 1275 keV gamma ray decayed from the ^{22}Na source or more than 8% events are from scatter or randoms events). Also, the APDs are sensitive to the temperature and bias voltages fluctuations. This could cause a change in the behavior of the detector after it has been calibrated.

The NLS approach has shown its attractive features such as good overall performance and no need to pre-calibrate. It also has some potential downsides: one needs to ascertain that the geometrical detector configuration applied in the model corresponds with the true physical layout. This can be achieved by proper mechanical mounting. Another limitation is the fitting algorithm. The Matlab implementation is rather slow. For real-time usage, one might have to use a GPU or hardware (i.e. FPGA) based solutions. The continuing improvement in the speed of these devices will probably enable the future real time implementation of the proposed least squares fitting algorithm.

Future Work

Three main research orientations can be considered to continue on the work in this thesis. The first one is to optimize the detector design, i.e. the shape and size of the crystal, the layout of the photo detector, etc. This will rely on Monte Carlo simulation. Second is to test the method using new types of photo detectors, such as GM-APD. A third aspect is to speed up the implementation of the algorithm with the final goal to have a real time implementation.

Optimize the detector design As we have discussed in section 5.2.10, making the size of the pixels smaller does not necessary result in a better spatial resolution performance, and this due to the lack of statistics from the amount of detected photons. More studies are needed for the NLS method in order to have the best spacial sampling.

In addition to the layout, the size and the shape of the crystal should also be invested. For example, to use the NLS method with trapezoidal shaped blocks to reduce the inter-block dead space in a ring configuration. Then, the position of the mirror sources should be adjusted, and the corresponding performance needs to be further evaluated.

As pointed out in section 5.2.8, the NLS method has a better performance when the interaction occurs close to the APD. On the other hand, more than 60 % of the events interact in the top part of the scintillator. Hence, placing the photo detector on top of the crystal might be a better choice.

Testing on new photo detectors GM-APD's draw a lot of attention in recent years. It has the combined advantage of a PMT (i.e. high Gain, low ENF and fast timing) and of an APD (compact and insensitive to magnetic fields). Testing NLS performance on this new kind of photo detectors will be very interesting.

Speeding up the implementation For a pixelated crystal, the gamma ray position is only determined by which small crystal generates the light. Therefore, the positioning is usually very fast. As for monolithic crystals, the positions are calculated by more complicated algorithms, requiring more processing time. Thus far, only a run time implementation was reported for the neural network that can be used on systems like Brain-PET. (i.e. processing up to 25 million events/second). The 2D-ML are being implement in CUDA and reach processing speeds up to 500 events/second [34]. As for the NLS method, no work has been done for optimizing the speed. Future research on a CUDA or an FPGA implementation to speed up the implementation are necessary.

Appendix A: Gamma Ray Interactions with Matter

Interaction processes

Because gamma ray photons have zero mass and zero charge, they interact differently in matter compared to charged particles in the sense that a charged particle penetrating in matter will interact with all electrons and nuclei on its path. Gamma rays, due to their electrical neutrality, are not influenced by Coulomb forces. They will travel some considerable distance without losing any energy before undergoing an interaction with one single atom, thereby partially or totally transferring their energy to an electron. These electrons will ultimately deposit their energy in the medium.

There are four gamma ray interaction mechanisms in matter: pair production, Compton scattering (or inelastic scattering), photo electric absorption and Rayleigh scattering (or elastic scattering). The following introductions are partially based on [12].

Pair production In pair production, a gamma ray interacts with the electric field of the nucleus of an atom. The photon energy is transformed into an electron-positron pair. Since the rest mass energy equivalent of an electron and positron is 0.511 MeV, pair production can only happen when the energy of the gamma ray is more than 1.022 MeV. Photon energy above this threshold is divided between the electron and positron as kinetic energy. The electron and positron loose their energy in multiple interactions. The positron will eventually annihilate with an electron, generating two 511 keV annihilation photons (figure A.1).

Since the highest energy occurring during the PET image formation is 511 keV, pair production is irrelevant for PET.

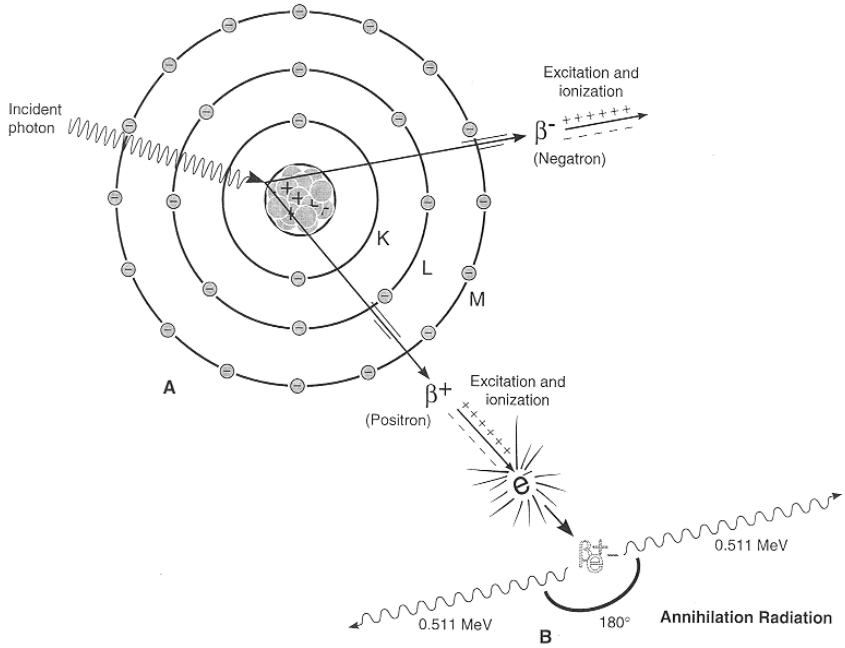


Figure A.1 – Pair production results in the production of an electron-positron pair. The positron will eventually annihilate in two 511 keV gamma rays, figure from [12].

Compton scattering The process of Compton scattering is an inelastic collision between an incident gamma ray photon and an outer shell electron. The gamma ray transfers part of its energy to a recoil electron and scatters over an angle θ. The recoil electron obtaining the energy is ejected from the atom (figure A.2). The probability distribution of the scattering angle θ (also called the differential cross section) depends on the initial energy of the gamma rays and is given by the Klein-Nishina formula :

$$\frac{d\sigma}{d\Omega} = 0.5 \times r_e^2 \times f(\epsilon,\theta)^2 \times [f(\epsilon,\theta) + f(\epsilon,\theta)^{-1} - sin^2(\theta)] \qquad (1)$$

where r_e is the classical electron radius (2.818×10^{-15} m), $\epsilon = \frac{E_0}{511\,kev}$ is the incident photon energy in units of the electron rest energy and $f(\epsilon,\theta) = 1/[1 + \epsilon(1 + cos\,\theta)]$. Figure A.3 shows the probability distribution for 10 keV, 100 keV and 511 keV gamma's. The length from the central point in the plot to one of the curves presents the probability of scattering along that angle. We can see from

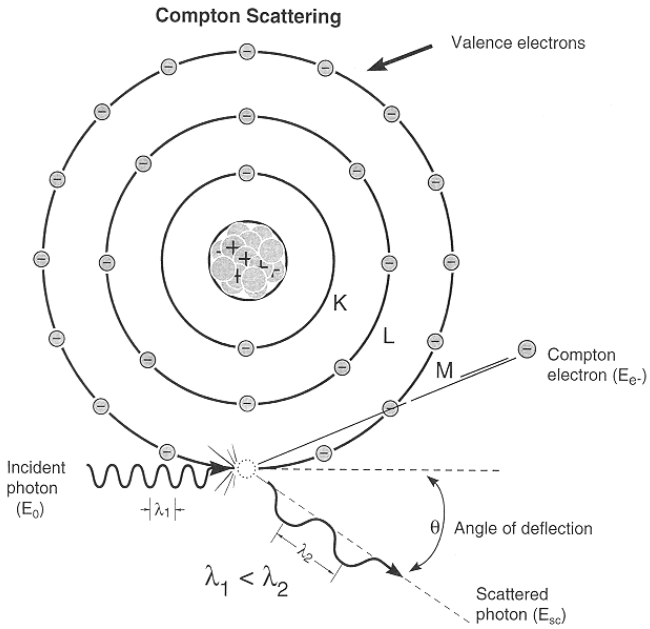


Figure A.2 – A Compton interaction results in the incoming gamma being scattered over an angle θ and transferring some of its energy to a recoil electron emitted from the atom. Figure from [12].

the figure that for the 511keV gamma rays, forward scattering is most probable. Large scattering angles are least likely to occur. As the photon energy drops, the scattering angle distributions become more symmetric along the forward and backward direction.

The energy of the scattered photon can be calculated from the energy of the incident photon and the angle θ of the scattered photon :

$$E_{sc} = \frac{E_0}{1 + \frac{E_0}{511\,kev}(1 - cos\,\theta)} \tag{2}$$

where E_0 is the initial gamma energy expressed in keV and E_{sc} is the remaining energy of the gamma after being scattered. The energy transferred to the ejected electron increases with the scattering angle. The maximum energy lost by the incoming gamma occurs during a 180 degree backscatter. In this case a 511 keV photon looses 340.6 keV, i.e. its energy after back scattering is still 170.3 keV. Hence it is not possible for a gamma to completely disappear during a Compton scattering process.

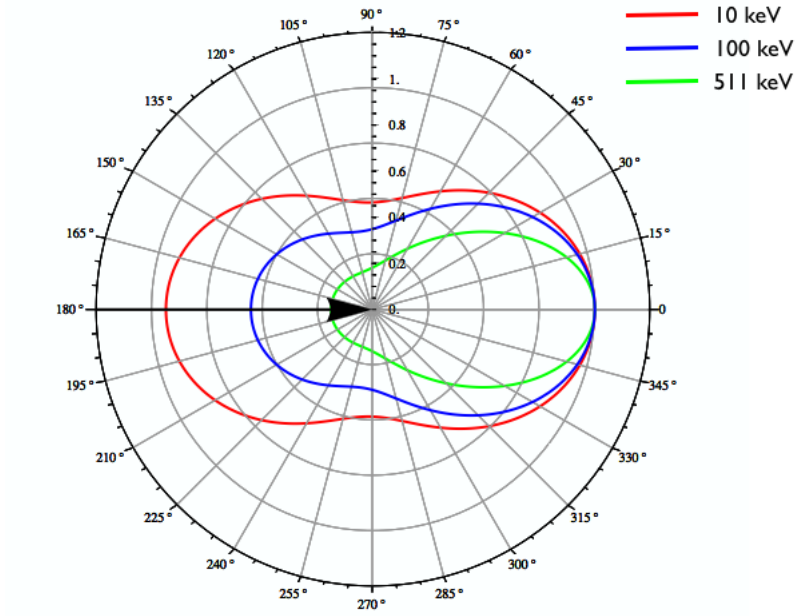


Figure A.3 – Differential cross section for the Compton interaction at 10 keV, 100 keV and 511 keV

Photo electric absorption In a photo electric absorption process, a gamma ray photon transfers all its energy to an electron which is ejected from the atom. The gamma completely disappears (figure A.4). The kinetic energy of the ejected electron, also called a photo electron, is equal to the incident gamma energy minus the binding energy of the electron. The electron quickly looses this kinetic energy to the surrounding material in multiple scattering interactions.

This ionization process creates a vacancy in the atomic electron shell, which is quickly filled by a free electron or an electron from an outer shell. This filling of the vacancy is accompanied by the emission of a characteristic X-ray (figure A.4). Because the energy of these X-rays is usually rather low, they are reabsorbed in the neighborhood. Sometimes no characteristic X-ray is emitted, but instead a second electron called Auger electron is ejected from the atom, carrying away the energy.

Rayleigh scattering In Rayleigh scattering, the incident photon interacts with and excites the total atom, as opposed to individual electrons in Compton scat-

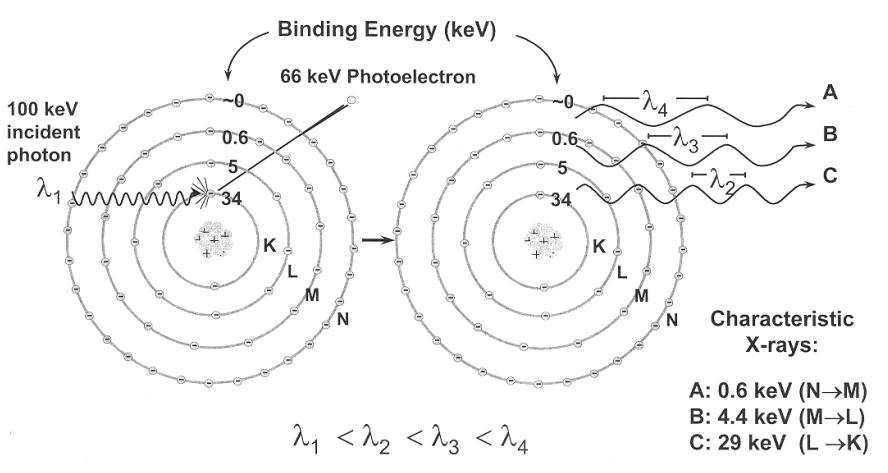


Figure A.4 – During a photo electric interaction the incoming gamma is completely absorbed, resulting in the the ejection of a photo electron and emission of characteristic X-rays. The wave length λ (or energy) of the characteristic X-rays depend on the energy difference between the electron orbitals involved in the transition of the electron filling up the vacancy created by the ejection of the photo electron. This wave length (energy) is material dependent. Figure from [12].

tering or photoelectric effect. During a Rayleigh scattering event, the complete energy of the incident photon is given off, causing all of the electrons in the scattering atom to oscillate in phase. The atom's electron cloud immediately radiates this energy again, emitting a photon of the same energy but in a slightly different direction (figure A.5). In this interaction, electrons are not ejected and thus ionization does not occur. In general, the scattering angle increases as the photon energy decreases.

Interaction cross sections

The probability that a given process occurs upon interaction of a gamma ray with material is determined by the cross sections of each of the possible interaction mechanisms. These cross sections depend on the energy of the gamma ray and the material it is traveling. Figure A.6 shows the cross sections of each of the three possible interaction processes in water and LSO as a function of energy.

At 511 keV the Compton scattering process is most dominant. In LSO, a 511 keV photon has 66% chance to be Compton scattered upon its first interaction. The cross section for Compton scattering depends on the electron density of the

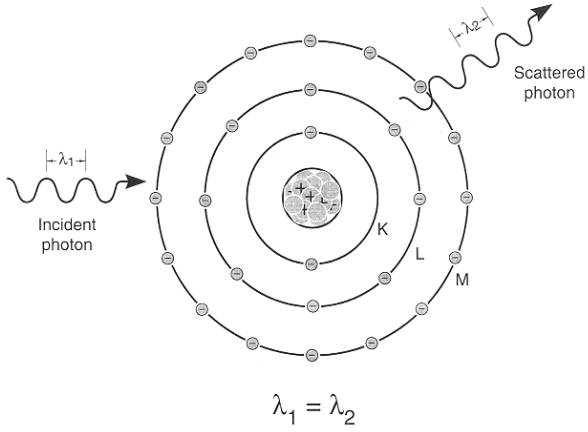


Figure A.5 – Rayleigh scattering is an elastic interaction. The wave length λ_2 of the scattered photon is the same as the wave length λ_1 of the incoming photon, i.e. both photons have the same energy.

material. With the exception of hydrogen, the total number of electrons per gram is fairly constant for most materials. Hence the probability of Compton scatter per unit of mass is rather independent of the atomic number Z.

At gamma energies below 400 keV the photo electric absorption is the most likely interaction to occurs in LSO. In water, Compton scattering remains the most important process until the gamma energy drops below 30 keV. In general, the cross section for photo electric absorption varies as $\frac{Z^4}{E^3}$ where Z is the atomic number and E is the energy of the gamma ray. High Z materials will hence increase the probability of photo electric absorption relative to Compton interaction. This is important for detectors since it results in more gamma's depositing all their energy.

Although Rayleigh scattering is only significant at very low energies (i.e. X-rays), it still has a probability of 5% to occur for 511 keV gammas interacting in LSO [67].

The total cross section, i.e the sum of the cross sections of the three interaction mechanisms, is also shown in figure A.6. This number relates to the mean free path (MFP) of the gamma ray when it enters the material. The MFP is the average distance it will travel before undergoing an interaction. Since the cross sections are expressed in $\frac{cm^2}{g}$, they need to be multiplied by the density of the material to obtain MFP^{-1}.

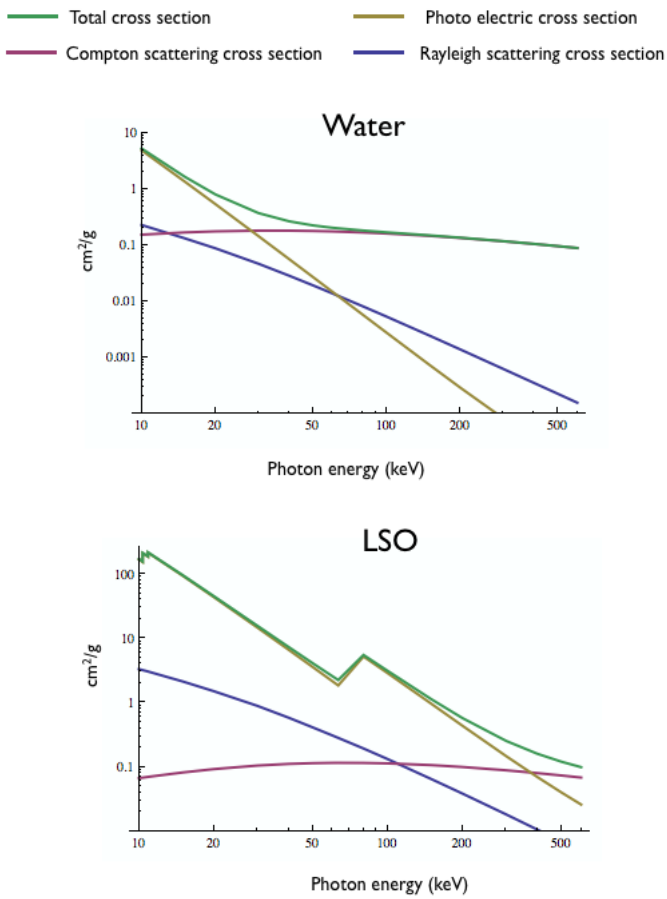


Figure A.6 – Cross sections in water and LSO for photo electric absorption, Compton interaction and Rayleigh scattering as a function of the gamma energy. The total cross section (proportional to the mean free path of the gamma) is also shown.

Bibliography

[1] C. Abreu et al. Characterization and quality control of avalanche photodiode arrays for the Clear–PEM detector moduless. *Nucl. Instrum. Methods*, 567:19–22, 2007.

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

[3] K. Arisaka. Principle of PMT and its calibration. *PMT Calibration school*, 2003.

[4] H. H. Barrett, W. C. J. Hunter, et al. Maximum–likelihood methods for processing signals from gamma-ray detectors. *IEEE Trans. Nucl. Sci.*, 56:725–735, 2009.

[5] H. H. Barrett and K. Myers. *Foundations of image science.* Wiley–Interscience, 2003.

[6] H. H. Barrett, C. J. H William, et al. Maximum-likelihood methods for processing signals from gamma-ray detectors. *IEEE Transactions on Nuclear Science*, 56:725–735, 2009.

[7] T Beyer, DW Townsend, T Brun, et al. A combined PET/CT scanner for clinical oncology. *J Nucl Med.*, 41:1369–1379, 2000.

[8] A. Blanco et al. Spatial resolution on a small animal RPC–PET prototype operating under magnetic field. *Nuclear Physics B (Proc. Suppl.)*, 158:157–160, 2006.

[9] A. Bronstein, M. Bronstein, M. Zibulevsky, and Y. Y. Zeevi. High-energy photon detection in positron emission tomography using adaptive non-linear parametric estimation algorithms. 2002.

[10] P. Bruyndonckx et al. Neural network-based position estimators for PET detectors using monolithic LSO blocks. *IEEE Trans. Nucl. Sci.*, 51:2520–25, 2004.

[11] P. Bruyndonckx et al. Evaluation of machine learning algorithms for localization of photons in undivided scintillator blocks for PET detectors. *IEEE Trans. Nucl. Sci.*, 55:918–24, 2008.

[12] J. T. Bushberg, J. A. Seibert, E. M. Leidholdt Jr, and J. M. Boone. *The Essential Physics of Medical Imaging.* Lippincott Williams & Wilkins, 2002.

[13] P. Buzhan et al. An advanced study of silicon photomultiplier. *ICFA Instrumentation Bulletin*, 2001.

[14] C. Moisan C, A. Levin, and H. Laman. Testing scintillation transport models with photoelectron yields measured under different surface finishes. *IEEE Nucl. Sci. Symp. Conf. Rec. (Vancouver)*, 1997.

[15] S. R. Cherry. Of mice and men (and positrons)–advances in PET imaging technology. *J Nucl Med*, 47:1735–45, 2006.

[16] S. R. Cherry, J. Sorenson, and M. Phelps. *Physics in Nuclear Medicine.* Saunders; 3 edition, 2003.

[17] T. F. Coleman and Y. Y. Li. On the convergence of interior-reflective newton methods for nonlinear minimization subject to bounds mathematical programming. *Mathematical Programming*, 67:189–224, 1994.

[18] T. F. Coleman and Y. Y. Li. A reflective newton method for minimizing a quadratic function subject to bounds on some of the variables. *SIAM Journal on Optimization*, 6-4, 1996.

[19] M. Defrise, D. Townsend, and A. Geissbuhler. Implementation of three-dimensional image reconstruction for multiring positron tomographs. *Phys. Med. Biol.*, 35:1361–1372, 1990.

[20] H. Du, Y. Yang, et al. Continuous depth-of-interaction encoding using phosphor-coated scintillators. *Phys. Med. Biol.*, 54:1757–71, 2009.

[21] C. W. E. Van Eijk. Inorganic scintillators in medical imaging. *Phys. Med. Biol.*, 47:R85–106, 2002.

[22] R. D. Evans. *The Atomic Nucleus.* Krieger Publishing Company, N.Y., 1982.

[23] J. A. Fessler. Penalized weighted least-squares image-reconstruction for positron emission tomography. *IEEE Trans. Med. Imag.*, 13:290–300, 1994.

[24] L. R. Furenlid, J. Y. Hesterman, and H. H. Barrett. Scintillation camera. *Rev. Sci. Instr.*, 29:27–33, 1958.

[25] Geant4. Geant4 users guide for application developers version: Geant4.9.3. *Geant4 Collaboration*, 2009.

[26] M. S. Gockenbach. *Optimality conditions for unconstrainted minimization.* Michiganchnological University, Lectures MA5630, 2005.

[27] H. Gotoh and H. Yagi. Solid angle subtended by a rectangular slit. *Nucl. Instrum. Methods*, 96:485–486, 1971.

[28] R. M. Gray and A. Macovski. Maximum a posteriori estimation of position in scintillation cameras. *IEEE Trans. Nucl. Sci.*, 23:849–852, 1976.

[29] Hamamastu. Si APD array S8550.

[30] J. S. Huber, W. W. Moses, W. F. Jones, and C. C. Watson. Effect of ^{176}lu background on singles transmission for LSO-based PET cameras. *Phys. Med. Biol.*, 47:3535–41, 2002.

[31] H. M. Hudson and R. S. Larkin. Accelerated image-reconstruction using ordered subsets of projection data. *IEEE Trans. Med. Imag.*, 13:601–609, 1994.

[32] S. Van Huffel and J. Vandewalle. *The total least squares problem: Computational aspects and analysis.* SIAM, Philadelphia, 1991.

[33] W. C. J. Hunter. Modeling stochastic processes in gamma-ray imaging detectors and evaluation of a multi-anode PMT scintillation camera for use with maximum-likelihood estimation methods. *Ph. D Dissertation, Dept. Phys. Univ. Arizona*, 2007.

[34] W. C.J. Hunter, H. H. Barrett, and L. R. Furenlid. Calibration method for ML estimation of 3D interaction position in a thick gamma-ray detector. *IEEE Transactions on Nuclear Science*, 56:189–196, 2009.

[35] S Jan et al. Gate: a simulation toolkit for PET and SPECT. *Phys. Med. Biol.*, 49:4543–61, 2004.

[36] M. Kaul et al. Thick continuous crystal design for PET. *Nuclear Science Symposium Conference Record, IEEE, Dresden, German*, 2008.

[37] Th Kirn et al. Absorption length, radiation hardness and ageing of different optical glues. *CMS NOTE 1999/003*.

[38] Th. Kirn et al. Absorption length, radiation hardness and ageing of different optical glues. *CMS NOTE 1999/003*, 1999.

[39] G.F. Knoll. *Radiation Detection and Measurement*. John Wiley & Sons, Inc., 1999.

[40] C. Lemaitre. Use of machine learning algorithms for gamma detection in positron emission tomography. *dissertation, Department of physics*, 2009.

[41] C. W. Lerche et al. Depth of gamma-ray interaction within continuous crystals from the width of its scintillation lightdistribution. *IEEE Trans. Nucl. Sci.*, 52:560–572, 2005.

[42] C. S. Levin and E. J. Hoffman. Calculation of positron range and its effect on the fundamental limit of positron emission tomography system spatial resolution. *Phys. Med. Biol.*, pages 781–799, 1999.

[43] C. S. Levin and H. Zaidi. Current trends in preclinical pet system design. *PET Clin 2*, pages 125–160, 2007.

[44] T. K. Lewellen. Recent developments in PET detector technology. *Phys. Med. Biol.*, 53:R287–R317, 2008.

[45] Xiaoli Li, W. C. J. Hunter, T. K. Lewellen, and R.S. Miyaoka. Spatial resolution performance evaluation of a monolithic crystal pet detector with Cramer-Rao lower bound (CRLB). *Medical Imaging Conference Record, IEEE, Knoxville, Tennessee*, 2010.

[46] T. Ling, T. H. Burnett, T. K. Lewellen, and R. S. Miyaoka. Parametric positioning of a continuous crystal PET detector with depth of interaction decoding. *Phys. Med. Biol.*, 53:1843–1863, 2008.

[47] T. Ling, T. K. Lewellen, and R. S. Miyaoka. Depth of interaction decoding of a continuous crystal detector module. *Phys. Med. Biol.*, 52:2213–2228, 2007.

[48] M. C. Maas, D. J. van der Laan, et al. Signal to noise ratio of APD–based monolithic scintillator detectors for high resolution PET. *IEEE Trans. Nucl. Sci.*, 55:842–852, 2008.

[49] B. McIntosh et al. Validation of gate simulations of the ^{176}lu intrinsic activity in LSO detectors. *Nuclear Science Symposium Conference Record, IEEE, Orlando, FL*, pages 3426–28, 2009.

[50] P. R. Mendes, P. Bruyndonckx, M. C. Castro, Zhi Li, J. M. Perez, and I. S. Martin. Optimization of a monolithic detector block design for a prototype human brain pet scanner. *IEEE Nucl. Sci. Symp. Conf. Rec. Dresden, Germany*, 2008.

[51] R. S. Miyaoka, T. Ling, et al. Calibration procedure for a continuous miniature crystal element (cMiCE) detector. *IEEE Trans. Nucl. Sci.*, 57(3):1023–1028, 2010.

[52] S. Moehrs et al. A detector head design for small animal PET with silicon photomultipliers (SiPM). *Phys. Med. Biol.*, 51:1113–27, 2006.

[53] C. Moisan, D. Vozza, and M. Loope. Simulating the performances of an LSO based position encoding detector for PET. *IEEE Trans. Nucl. Sci.*, 44:2450–8, 1997.

[54] J. B. Mosset. Developpement dun module de detection phoswich LSO/LuYAP pour le prototype de camera a positrons ClearPET. *These EPFL, no 3596*, 2006.

[55] H. Motulsky and A. Christopoulos. *Fitting Models to biological data using linear and nonlinear regression–A practical guide to curve fitting.* PRISM, GraphPad Software, Inc., 2003.

[56] J. Nocedal and S. J. Wright. *Numerical Optimizaiton.* Springer, 2006.

[57] S. A. Payne et al. Nonproportionality of scintillator detectors: Theory and experiment. *IEEE Trans. Nucl. Sci.*, 56, 2009.

[58] C. Pedrini. Scintillation mechanisms and limiting factors on each step of relaxation of electronic excitations. *Physics or the solic state*, 47:1359–11363, 2005.

[59] M. E. Phelps. *PET physics, instrumentation, and scanners*. Springer, 2006.

[60] R. Pintelon and J. Schoukens. *System Identification: A Frequency Domain Approach*. Wiley-IEEE Press, 2001.

[61] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in C, The Art of Scientific Computing, Second Edition*. CAMBRIDGE UNIVERSITY PRESS, 1992.

[62] M. Rudin. *Molecular Imaging*. Imperial College Press, 2005.

[63] A. Sanchez-Crespo, P. Andreo, and S. Larsson. Positron flight in human tissues and its influence on PET image spatial resolution. *European Journal of Nuclear Medicine and Molecular Imaging*, 31:44–51, 2004.

[64] D. R. Schaart et al. A novel, SiPM-array-based, monolithic scintillator detector for PET. *Phys. Med. Biol.*, 54:3501–3512, 2009.

[65] D. J. Schlyer. PET tracers and radiochemistry. *Ann. Acad. Med. Singapore*, 33:146–54, 2004.

[66] ZhenJun Shi and XiangSun Zhang. From line search method to trust region method. *International Symposium on OR and Its Applications*, pages 156–170, 2005.

[67] J. R. Stickel and S. R. Cherry. High-resolution PET detector design: modelling components of intrinsic spatial resolution. *Phys. Med. Biol.*, 50:179–195, 2005.

[68] S. Tavernier. *Experimental Techniques in Nuclear and Particle Physics*. Springer, 2010.

[69] 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 A*, 537:321–325, 2005.

[70] D. J. van der Laan. Modelling monolithic scintillator detectors for PET. *Ph.D Thesis, department of radiation, Delft University of Technology*, 2009.

[71] D. J. van der Laan et al. Using Cramer-Rao theory combined with Monte Carlo simulations for the optimization of monolithic scintillator PET detectors. *IEEE Transactions on Nuclear Science*, 53:1063–1070, 2006.

[72] R. Vinke et al. Optimization of digital time pickoff methods for LaBr3–SiPM TOF–PET detectors. *Nuclear Science Symposium Conference Record, IEEE, Orlando, FL*, 2009.

[73] M. Wedrowski. Artificial neural network based position estimation in PET. *Ph.D Thesis, Interuniversity Institue for High Energies, Vrije Universiteit Brussels*, 2010.

[74] K Ziemons. A monolithic block detector design for a hybrid MR-PET scanner. In *AstroMed09 Conference Talk*, 2009.

Abbreviations

ANN	artificial neural networks
APD	avalanche photo diodes
BGO	Bismuth Germanate
CE	collection efficiency
COG	center-of-gravity
CT	computed tomography
CZT	Cadmium Zinc Telluride
DOI	depth-of-interaction
DQE	detection quantum efficiency
ENC	equivalent noise charge
ENF	excess noise factor
FBP	filtered-back projection algorithm
FDG	Fluoro-deoxyglucose
FOV	the field of view
FWHM	full-width at half-maximum
FWTM	full-width at tenth-maximum
LOR	line-of-response

MFP mean free path

ML maximum likelihood

MPPC multi-pixel photon counters

MRI magnetic resonance imaging

PDE photon detection efficiency

PET positron emission tomography

PMT photomultiplier tubes

PSF point spread function

QE quantum efficiency

SiPMT silicon photo multipliers

SNR signal-to-noise ratio

SPECT single photon emission computed tomography

US ultrasound

List of publications

Peer Reviewed Paper

1. P. Bruyndonckx, C. Lemaître, D. J. van der Laan, M. Maas, D. Schaart, Wang Yonggang, Zhi Li, M. Krieguer, and S. Tavernier, "Evaluation of Machine Learning Algorithms for Localization of Photons in Undivided Scintillator Blocks for PET Detectors." *IEEE Trans. Nucl. Sci.*, 55, pp 918-924, 2008

2. Zhi Li, M. Wedrowski, P. Bruyndonckx and G. Vandersteen, "Nonlinear least-squares modeling of 3D interaction position in a monolithic scintillator block." *Phys. Med. Biol.* 55, pp 6515-6532, 2010

Conference proceeding

1. P. R. Mendes, P. Bruyndonckx, J. Navarrete, J. M. Pérez and Zhi Li, "Evaluation of Monolithic Detector Blocks for High- Sensitivity PET Imaging of the Human Brain." *IEEE Nucl. Sci. Symp. Conf. Rec.* Hawaii, USA, 2007

2. P. Bruyndonckx, Zhi Li, C. Lemaitre, J. M. Perez, P. Rato, D. Schaart, M. Maas, D. J. Van der Laan, S. Tavernier and Wang Yonggang, "Impact of instrumentation parameters on the performance of neural network based positioning algorithms for monolithic scintillator blocks.", *IEEE Nucl. Sci. Symp. Conf. Rec.* Hawaii, USA, 2007

3. Zhi Li, P. Bruyndonckx, Jun Dang, M. Wedrowski, J. M. Perez, P. R. Mendes, K. Ziemons and S.Tavernier, "Monte Carlo Evaluation of Monolithic Scintillator Block Detectors Using Silicon PMTs." *IEEE Nucl. Sci. Symp. Conf. Rec.* Dresden, Germany, 2008

4. P. R. Mendes, P. Bruyndonckx, M. C. Castro, Zhi Li, J. M. Pérez and I. S. Martín, "Optimization of a Monolithic Detector Block Design for a Prototype Human Brain PET Scanner," *IEEE Nucl. Sci. Symp. Conf. Rec.* Dresden, Germany, 2008

5. Zhi Li, P. Bruyndonckx, M. Wedrowski and Gerd Vandersteen, "3D Nonlinear Least Squares Position Estimation in a Monolithic Scintillator Block," *IEEE Nucl. Sci. Symp. Conf. Rec.* Orlando, USA, 2009

6. M. Wdrowski, P. Bruyndonckx, S. Tavernier, Zhi Li, Jun Dang, P. R. Mendes, J. M. Perez and K. Ziemons, "Robustness of Neural Networks Algorithm for Gamma Detection in Monolithic Block Detector, Positron Emission Tomography," *IEEE Nucl. Sci. Symp. Conf. Rec.* Orlando, USA, 2009