Estimation of the Image Quality in Emission Tomography: Application to Optimisation of SPECT System Design.

Niccoló Fuin
UCL

A thesis submitted to University College London
for the Degree of Doctor of Philosophy
in Medical Physics applied to Nuclear Medicine

September 2013

I, Niccoló Fuin, confirm that the work presented in this thesis is my own.

Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

In Emission Tomography the design of the Imaging System has a great influence on the quality of the output image. Optimization of the system design is a difficult problem due to the computational complexity and to the challenges in its mathematical formulation. In order to compare different system designs, an efficient and effective method to calculate the Image Quality is needed.

In this thesis the statistical and deterministic methods for the calculation of the uncertainty in the reconstruction are presented. In the deterministic case, the Fisher Information Matrix (FIM) formalism can be employed to characterize such uncertainty. Unfortunately, computing, storing and inverting the FIM is not feasible with 3D imaging systems. In order to tackle the problem of the computational load in calculating the inverse of the FIM a novel approximation, that relies on a sub-sampling of the FIM, is proposed. The FIM is calculated over a subset of voxels arranged in a grid that covers the whole volume. This formulation reduces the computational complexity in inverting the FIM but nevertheless accounts for the global interdependence between the variables, for the acquisition geometry and for the object dependency.

Using this approach, the noise properties as a function of the system geometry parameterization were investigated for three different cases. In the first study, the design of a parallel-hole collimator for SPECT is optimized. The new method can be applied to evaluating problems like trading-off collimator resolution and sensitivity. In the second study, the reconstructed image quality was evaluated in the case of truncated projection data; showing how the subsampling approach is very accurate for evaluating the effects of missing data.

Finally, the noise properties of a D-SPECT system were studied for varying acquisition protocols; showing how the new method is well-suited to problems like optimizing adaptive data sampling schemes.

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my primary supervisor, Brian Hutton, whose ideas and encouragements have represented a crucial contribution to this project.

I would like to thank my secondary supervisors, Sebastien Ourselin and Simon Arridge, for their helpful input into the project.

I would like to thank all members of the physics research group at the Institute of Nuclear Medicine UCL, for their support throughout my time in the department. Thank you also to everyone at CMIC who have supported me during my time at UCL. In particular, I would like to thank Stefano Pedemonte for the help, the discussions and because his passion for this job is contagious.

Alla mia famiglia, per il supporto, la fiducia e l'affetto che mai mi sono mancati.

Un ringraziamento speciale a Juliette per l'incoraggiamento e per le attenzioni durante la stesura di questa tesi.

Contents

1 Introduction **18**

1.1 Preamble . . . 18

1.2 Aim and Contributions . . . 19

1.3 Outline . . . 21

2 Background **23**

2.1 Emission Computed Tomography . . . 23

2.2 Discrete System Model . . . 32

2.3 Noise Model . . . 33

2.4 Space-Invariant vs Space-Variant Systems . . . 34

2.5 Tomographic Image Reconstruction . . . 36

2.6 Data Collection . . . 44

3 Statistical Evaluation of Image Quality **45**

3.1 Statistical Method for the Calculation of the Uncertainty in the Estimation . . . 47

3.2 Quantifying Resolution . . . 51

3.3 The Linearised Local Impulse Response . . . 56

3.4 Figures Of Merit for Image Quality . . . 58

3.5 Proposed Experimental Set-Up . . . 60

4 Deterministic Evaluation of Image Quality **62**

4.1 Introduction . . . 62

4.2 Mean and Covariance of Implicit Estimators . . . 64

4.3 Linear Local Impulse Response . . . 69

4.4 The Fisher Information Matrix 70
4.5 Efficient Calculation of the FIM 74
4.6 Image Quality Quantification for System Design 83
4.7 Discussion 87
4.8 Proposed Experimental Set-Up 89

5 Collimator Design: Resolution-Sensitivity trade-off 91

5.1 Collimator Design Parameters 93
5.2 Experiments 97
5.3 Results 100
5.4 Validation 104
5.5 Discussion 108

6 Region Of Interest Reconstruction 111

6.1 Introduction 111
6.2 Singular Value Decomposition 113
6.3 Experiments 114
6.4 A Data Sufficiency Condition for The Interior Problem 116
6.5 Results 120
6.6 Discussion 130

7 Changing Acquisition Trajectory: The D-SPECT System 137

7.1 Introduction 137
7.2 System Description 139
7.3 SVD for Different Acquisition Protocols 142
7.4 Statistical Calculation of the Uncertainty 142
7.5 Experiments 144
7.6 Results 145
7.7 Validation 154
7.8 Discussion 158

8 Conclusions and Future Work 165

8.1 Main Contributions and General Discussion 165

8.2 Suggestions for Future Work 170

8.3 Publications Arising from Thesis Work 174

A Methodology for the Optimisation of Novel Collimator Design 176

A.1 Design Parameters and Cost Function 177

A.2 Ray-Tracing Algorithm 179

A.3 Optimisation Using Genetic Algorithm 180

A.4 Conclusion 182

List of Figures

2.1 Simplified diagram for a SPECT camera. 25

2.2 A - Scintillation event on a SPECT detector. B - Schematic representation of the collimator. 27

2.3 Modelling the detector response in a discrete tomographic system model. Ideal and depth-dependent detector-collimator response. 35

2.4 Comparison of space-invariant and space-variant geometric responses. 36

2.5 Statistical reconstruction with and without regularisation. A (left) - Maximum Likelihood Reconstruction. B (right) - Regularised Likelihood Reconstruction. 40

3.1 Variance as function of Number of Iterations and of Number of Noise Instances. 52

3.2 Non uniformities for PL estimator in ideal emission tomography reconstruction using a test phantom. 54

3.3 Non uniformities in SPECT reconstruction using a Jaszczak test phantom. . . . 55

3.4 LLIRs contours and profiles, at different locations of the phantom, for a SPECT system and PL estimator. 59

4.1 Example of grids for the estimation of the uncertainty. 78

4.2 Rotation-based algorithm for fast computation of the Fisher Information Matrix: 3-D schematic representation. 79

4.3 Comparison between full FIM, subsampled FIM and Circulant FIM. 84

4.4 Comparison between the Covariance matrix obtained from the full FIM, the subsampled FIM and the matrix obtained by row-by-row inversion of the circulant FIM. 85

4.5 Schematic representation of a grid and respective subsampled covariance matrix. 90

5.1 Objective function Ψ curves as a function of number of iterations. The different curves represent the objective function for different collimator apertures. 98

5.2 Variance images for the NCAT phantom obtained with the reference method and the subsampled FIM method, for a standard SPECT system. 102

5.3 Variance images of a uniform sphere obtained with a standard SPECT system. . 103

5.4 CNRs for different collimator apertures and different target resolutions. 105

5.5 Validation of the Fisher information-based standard deviation calculation method for different grids for a conventional SPECT camera and the NCAT phantom. 107

6.1 Phantoms for ROI reconstruction and Interior Problem 118

6.2 Objective function Ψ curves as a function of number of iterations. The different curves represent the objective functions for different levels of truncation. 119

6.3 SVD spectra for the interior problem. 122

6.4 Sufficient condition for the interior problem: Non-truncated case 124

6.5 Sufficient condition for the interior problem: truncated case 125

6.6 Mean image from truncated projection data, for a uniform sphere phantom and uniform non-zero background 127

6.7 nterior tomography: Variance images for a uniform sphere phantom obtained with truncated projection data with ROI diameter $u = 36$. 128

6.8 Interior tomography: CNR for different levels of truncation for a voxel in the center of the sphere. ROI diameter (from left to right) $u = 96, \cdots, 16$. 129

6.9 Validation of the Fisher information-based standard deviation calculation method for different grids, in case of truncated data. 131

6.10 Comparison between full FIM, subsampled FIM and Circulant FIM, for the interior problem. 134

6.11 Comparison between the Covariance matrix obtained from the full FIM, the subsampled FIM and the matrix obtained by row-by-row inversion of the circulant FIM; for the interior problem. 135

7.1 A - Position of pD-SPECT detectors. B - Angular movement of a single pD-SPECT detector. 142

7.2 Objective function Ψ curves as a function of number of iterations. The different curves represent the objective function for different scanning patterns of a pD-SPECT system. 143

7.3 SVD spectra for the D-SPECT system for different acquisition protocols. . . . 147

7.4 Mean images obtained from the pD-SPECT system with varying scanning pattern time ratio. The slice of the digital NCAT phantom includes the left and right ventricle myocardium. 150

7.5 Variance images obtained from the pD-SPECT system with varying scanning pattern time ratio. The slice of the digital NCAT phantom includes the left and right ventricle myocardium. 151

7.6 Mean images obtained from the pD-SPECT system with varying scanning pattern time ratio. The slice of the digital NCAT phantom includes the apex of the heart and part of the liver. 152

7.7 Variance images obtained from the pD-SPECT system with varying scanning pattern time ratio. The slice of the digital NCAT phantom includes the apex of the heart and part of the liver. 153

7.8 Variance images of a uniform sphere ($\lambda = 8kBq/cm^3$) and uniform background ($\lambda = 2.2kBq/cm^3$) obtained from the pD-SPECT system with scanning pattern time ratio $S = 0.45$. 155

7.9 Variance images of a uniform sphere ($\lambda = 8kBq/cm^3$) and uniform background ($\lambda = 2.2kBq/cm^3$) obtained from the pD-SPECT system with scanning pattern time ratio $S = 0.9$. 156

7.10 CNRs for different scanning patterns of the pD-SPECT system and for different levels of background. 161

7.11 Images of covariance for a point in the centre of the uniform sphere for different acquisition protocols of a pD-SPECT system. 162

7.12 Validation of the Fisher information-based standard deviation calculation method for a pD-SPECT acquisition protocol with time ratio $S = 0.45$. 163

7.13 Validation of the Fisher information-based standard deviation calculation method for a pD-SPECT acquisition protocol with time ratio $S = 0.9$. 164

A.1 Flow chart of the method for optimisation of novel collimator designs. 178
A.2 Raytracer: rays are traced from the source to the detector. 180
A.3 Flow chart of the Genetic Algorithm. 182

List of Tables

4.1 Computation times for the calculation of the FIM and its inverse. 80

5.1 Validation of the Subsampled FIM for the NCAT phantom. 106
5.2 Validation of the Subsampled FIM for the uniform phantom. 108

6.1 Condition numbers for the interior problem. 121
6.2 Validation of the Subsampled FIM for the interior problem. 130

7.1 pD-SPECT System Parameters 140
7.2 Condition numbers for the pD-SPECT system with different acquisition protocols. 148
7.3 Validation of the Subsampled FIM for the pD-SPECT system. 157

List of Acronyms

CDR	collimator detector response	28
CNR	contrast to noise ratio	60
CRC	Cramér-Rao bound	72
CRC	contrast recovery coefficient	58
CT	computed tomography	18
$CUDA$	aka compute unified device architecture	80
CZT	cadmium zinc telluride	26
DFT	discrete Fourier Transform	75
$D-SPECT$	Dynamic-SPECT. Spectrum Dynamics, Caesarea, Israel	137
EM	expectation maximisation	42
FFT	fast Fourier transform	80
FIM	Fisher information matrix	70
FIM	figure of merit	44
FOV	field of view	35
$FWHM$	full width at half maximum	26
$FWHM$	graphics processing unit	43
GA	genetic algorithm	176
$IFFT$	inverse fast fourier transform	80
$LEHR$	low energy high resolution collimator	55
LIR	local impulse response	56
$LLIR$	linearised local impulse response	57
LOR	line of response	25
MLE	maximum likelihood estimation	38

MRI	magnetic resonance imaging	18
$NaI(Tl)$	thallium-activated sodium iodide	25
OS	ordered subset	42
$pD - SPECT$	pseudo Dynamic-SPECT	140
PET	single photon emission computed tomography	19
PH	parallel hole collimator	91
PL	penalised likelihood	41
PLE	penalised likelihood estimator	40
PMT	photo-multiplier tube	25
PSF	point spread function	29
ROI	region of interest	111
SNR	signal-to-noise ratio	60
$SOPS$	sums of products	77
$SPECT$	positron emission tomography	19
SVD	singular value decomposition	113
$UCRB$	uniform Cramér-Rao bound	83

List of Symbols

a	attenuation factors	33
$\boldsymbol{A}$	matrix of attenuation factors	33
b	geometric system model term	33
$\boldsymbol{B}$	geometric system model matrix	33
$\mathcal{B}$	Fourier transform of the LLIR	86
c	detector efficiency	33
CD	distance between a point source and the detector plane	95
d	detector unit index	33
$D[\cdot]$	produces a diagonal matrix of the vector in argument	75
D	diameter of uniform circle	81
e_i	unit basis vector	35
E	geometric efficiency of the collimator	96
f	one column of the Fisher information matrix	75
$\boldsymbol{F}$	Fisher information matrix	70
$\mathcal{F}$	Fourier Transform	75
$\mathcal{F}^{-1}$	Inverse Fourier Transform	75
g	grid for the calculation of the subsampled FIM	89
$\boldsymbol{G}$	grid for the calculation of the subsampled FIM (subset of voxel indexes)	76
h	system matrix element	33
$\boldsymbol{H}$	system matrix	33
i	voxel index	33
k	set of voxels whose intensity are to be combined in a penalty function	41
l	length of the holes through the collimator	94

L	log-likelihood function	37
M	total number of recorded counts	33
n	noise realisations index	47
n_l	number of collimator layers	177
N	total number of voxels	32
N_G	number of voxels for a grid G	76
N_r	number of noise realisations	47
p	conditional likelihood function	34
$\boldsymbol{P}$	anisotropic post-smoothing filter	83
$\boldsymbol{P_t}$	target bias gradient	83
$\mathcal{P}$	Fourier transform of the anisotropic post-smoothing filter	86
$\mathcal{P}_t$	Fourier transform of the target bias gradient	86
$\boldsymbol{o}$	distance between two collimator layers	177
$\boldsymbol{q}$	thickness of the collimator septa	94
$\boldsymbol{Q}$	bias gradient of the estimator	72
R	penalty function	37
s	spatial variation in sensitivity term	33
S	time ratio for the acquisition protocol of a D-SPECT system	142
t	generic term for threshold	114
t_a	constant scanning time for every angular step for a D-SPECT system	142
T	total scanning time	142
u	limited number of detector bins used to measure data in case of truncation	115
$\boldsymbol{U}$	mean gradient of the estimator	72
v	diameter of the collimator holes	94
V	singular vector of the system matrix	114
w	penalty weight	41
x, y, z	cartesian coordinates in the image space	31
β	regularization parameter (or hyper-parameter) for PL estimation	37
$\boldsymbol{\Gamma}$	Hessian of the objective function	65
$\boldsymbol{\lambda}$	activity distribution (vector of emission rates)	32

$\hat{\boldsymbol{\lambda}}$	estimate of the activity distribution	38
$\check{\boldsymbol{\lambda}}$	mean of the estimate of the activity distribution	47
μ	attenuation coefficient	31
$\boldsymbol{\nu}$	vector of measurement data	33
$\overline{\boldsymbol{\nu}}$	vector of noiseless measurement data	56
σ	singular values of the system matrix	114
$\boldsymbol{\Sigma}$	partial derivatives of the objective function	65
φ	angles tangent to the ROI contour for a D-SPECT system	141
Ψ	objective function for statistical reconstruction	37

Chapter 1

Introduction

1.1 Preamble

Emission computed tomography is a nuclear medicine tomographic imaging technique involving radio-active emission. This type of tomography differs from medical image modalities such as computed tomography (CT) or magnetic resonance imaging (MRI), in which morphological differences between tissues are exploited to gather information about the anatomy of the patient. While these modalities are referred to as anatomical imaging modalities, nuclear medicine modalities are applied for the purpose of functional imaging, in which specific metabolic or physiological processes are visualised.

In order to achieve this aim, first a radionuclide is combined with a molecule or with an existing pharmaceutical compound, to form a radio-pharmaceutical. Second, a very low dose of this radio-pharmaceutical is administered to the patient, usually by intravenous injection. This radio-pharmaceutical, once administered to the patient, can localise to specific organs or cellular receptors. This property of radio-pharmaceuticals allows nuclear medicine the ability to image a specific physiological process within a certain tissue or organ and to determine the uptake of certain substances within the body to evaluate the risk of developing a certain disease.

The radionuclides can be categorised into two classes. The first class are the single photon emitters which emit γ-photons that can be directly measured by the γ-camera. The second one are positron emitters which emit positrons that annihilate with electrons up to a few millimeters away, causing two γ photons to be emitted in opposite directions. Single photon emitters do not produce collinear photons, and the acquisition must rely on the detection of individual photons using collimated detectors with single photon emission computed tomography (SPECT) camera; whereas a positron emission tomography (PET) camera measures the number of emitted positrons by coincidence detection without the need of collimation.

A reconstruction algorithm is then used to provide a three-dimensional image of the radioactivity distribution that is more likely, according the characteristics of the imaging system and to prior belief, to have generated the photon emissions measured by the SPECT or PET camera. Such measurement is uncertain due to the limited amount of information that the scan may acquire. The dose of radio-pharmaceutical administered to the patient is limited by safety constraints. Consequently, the number of photons acquired during SPECT and PET medical acquisition is of the order of only a few millions to a few hundreds of millions. Discretising the imaging volume in a few millions voxels, it is immediately understood how the uncertainty associated with the photon counting process constitutes a major limitation in emission computed tomography. The characterisation of the uncertainty associated with the measurement of activity is essential in order to inform decision processes that arise in medical diagnosis and to merge the information provided by the measurement of activity with information provided by other imaging modalities and bio-markers. Furthermore it is essential to estimate the uncertainty of the measurement produced by a given system set-up, in order to compare it with other set-ups, thereby enabling the optimisation of the design of an imaging system.

1.2 Aim and Contributions

The main goal of this PhD work is to evaluate and optimise the design of SPECT imaging systems by computing figures of merit that characterise the uncertainty in the estimation. Such optimisation problems include the choice of a particular type of detector and collimator and tuning of their design parameters, the choice of the number of cameras and their position. While such class of design optimisation problems may be referred to as hard optimisation, the development of adaptive SPECT systems has introduced a second class of soft optimisation problems, where the parameters of the imaging system may be modified during the acquisition in order to image certain desired properties of the underlying object and to adapt to the imaging conditions.

The introduction of adaptive optimisation problems emphasises the need for an efficient method for the comparison of a set of different design parameters in a reasonably short time.

In such cases, numerical simulations can be prohibitively expensive and therefore it is useful to have approximate deterministic expressions of the figures of merit that characterise the uncertainty in the estimation. From the Bayesian perspective, the Fisher Information Matrix (FIM) can be employed to characterise the uncertainty of the reconstruction. Unfortunately, computing, storing and inverting the FIM is not feasible for the typical matrix size of 3-D imaging systems, due to the high dimensionality of the imaging volume.

In order to tackle the problem of the computational load in inverting the FIM an approximation has been previously proposed by Qi and Leahy [2000], who argued that if we are only interested in calculating the properties of an estimator in a single voxel i, it is acceptable to ignore the non-stationarity of the FIM. The computations are done for voxel i, and therefore only the i-th row of the FIM needs to be calculated. This local approximation of the FIM, is obtained by replacing all rows of the FIM with the shifted version of its $i - th$ row and then by inverting this shift-invariant matrix in order to estimate the variance in each voxel i. Consequently the FIM simply reduces to a circulant matrix and this approximation is referred to as the circulant approximation.

However, since the tomographic imaging system measures the integral along lines that traverse the entire imaging volume, the estimate of the activity in a given voxel and its uncertainty are related to the estimate and to the uncertainty in every other location. The full FIM accounts for such complex interdependence between all the voxels in the imaging volume; whereas the aforementioned circulant approximation makes use of a single row of the FIM and does not capture such interaction.

In this thesis, a novel algorithm for efficient estimation of the uncertainty in the reconstruction, based on the FIM, is introduced. Our new formulation relies on sub-sampling the FIM. The FIM is calculated over a subset of voxels arranged in a grid that covers the whole volume. Every element of the FIM at the grid points is calculated exactly, accounting for the acquisition geometry and for the object, without further approximation. This new formulation reduces the computational complexity in inverting the FIM but nevertheless accounts for the global interdependence between the variables.

The aim of this thesis is to describe the novel algorithm and to explore its use for the optimisation of SPECT systems; emphasising how it enables us to explore the design of highly shift variant systems, as a result of distance dependent resolution, data truncation or adaptive data sampling. Such systems include the standard rotating camera with parallel hole collimator and an adaptive system for cardiac imaging, similar to the commercially available D-SPECT.

1.3 Outline

Chapter 2 provides background information about SPECT imaging which is essential for the understanding of the subsequent chapters. The basic principle of SPECT imaging is briefly explained, followed by a description of typical collimator and detector designs used for SPECT systems. Then, the acquisition models are described, and the concepts of analytical and statistical image reconstruction are introduced.

Chapter 3 summarises the statistical approach for the calculation of the uncertainty in the reconstruction. First, the most common figures of merit used for image quality evaluation are listed. Second, a method to quantify resolution for Penalized Likelihood estimators will be discussed and the concept of Linearised Local Impulse Response will be introduced.

Chapter 4 summarises the deterministic method for the calculation of the uncertainty in the reconstruction based on the calculation and inversion of the Fisher Information matrix. First, the key concepts and the basic mathematical derivations of this method are described. Second, the new methodology for approximate calculation of the FIM is introduced, and the efficient implementation of the algorithm is described in detail. Third, a figure of merit for the estimation of the image quality, based on the fundamental trade-off between bias and variance that can be achieved in the reconstruction of emission tomograms, is presented.

In chapter 5 the approximated calculation of the FIM is employed for the optimization of of a parallel hole collimator for SPECT. A key challenge in SPECT system design is the achievement of a reasonable trade-off between resolution and detection efficiency. In order to prove the reliability of the proposed approximation, we show that different subsamples of the FIM yield the same optimal collimator aperture.

In chapter 6, in order to emphasize the benefits of the proposed approximation of the FIM with respect to existing methods for the deterministic evaluation of the uncertainty; we investigate how it can be employed to calculate the reconstructed image quality in the case of region-of-interest reconstruction from truncated projection data.

In chapter 7, the proposed novel algorithm is employed for the optimization of the camera trajectory in an adaptive SPECT system. This experiment is also meant to highlight the performance of the new method when used for optimization of systems with a highly shift-variant response, in comparison with other methods for the calculation of the uncertainty, such as the circulant approximation.

In conclusion, chapter 8 highlights the main contributions of this PhD work, discuss the usefulness and limitations of the novel algorithm, draws general conclusions and gives some suggestions for future work.

Chapter 2

Background

The topic of estimation of image quality is of interest in any image formation problem where a system is designed in order to produce images that will be used for a specific application. While many contributions brought in this thesis can be applied to generic imaging problems, the motivation for this work arose specifically from the problem of optimising a Single Photon Emission Computed Tomography (SPECT) system design. Therefore, most of the investigations and discussions presented in this thesis focus on image quality assessment in emission tomography with application to SPECT imaging. This background section introduces a summary of some important aspects of emission tomography. First the design of the imaging systems and the nature of the data produced by such systems are discussed. Second, we describe some of the reconstruction methods that can used to form images from these data.

2.1 Emission Computed Tomography

Emission computed tomography is a medical imaging modality which provides functional information about physiological processes in the body. Its development has enabled safe and non-invasive diagnostic capabilities for a variety of medical conditions where early and accurate diagnosis is critical. For instance, it can be applied for the detection of breast, lung or brain tumours, diagnosing myocardial perfusion defects, and imaging brain function in epilepsy, Alzheimer's disease or stroke patients [Brill and Beck, 2004]. Emission tomographic images are also used to monitor the treatment of specific patients, as well as to evaluate the effectiveness of new procedures and medications.

Typically, as a first step, a radio-tracer is introduced into a subject via injection. Radio-tracers are biological products whose carrier molecules, such as antibodies or ligands are coupled to radioactive nuclides. When a radio-tracer is administered into the body, the radioactive agent localises in the anatomical region or tissue of interest, thanks to the biochemical nature of the carrier molecule.

Gamma rays (γ-rays) are emitted isotropically as a product of decaying atoms of the radionuclide. These emissions radiate from within the patient's body at a rate proportional to the concentration of the radiopharmaceutical. Although the quality of the image significantly improves with the increase in the amount of radioactivity injected, the radiation exposure to internal organs imposes a limit on the dose. Therefore the dosage must be small enough to avoid having unwanted biochemical effects.

As the γ-rays are emitted from the patient's body, they are detected and recorded by the imaging hardware. The measurements thus form a sort of histogram of the number of photons absorbed at different locations on the surface of the detector. Generally speaking, the higher the radioactivity distribution in some anatomical region, the higher the number of photons that will be collected by the detector facing that region. Thus the photon counts collected from the detector provide information about the radio-tracer uptake, representing functional information about physiological processes in the body.

Emission tomography encompasses two families of imaging modalities, PET and SPECT. This thesis is concerned with SPECT imaging whose features will be introduced in the following.

2.1.1 Single-Photon Emission Computed Tomography

The radioactive tracers used in SPECT are γ - emitting radionuclides. When an atom of such a radionuclide decays, γ - ray photons are emitted isotropically. The photons that pass through the patient's body are collected by the detector. The typical SPECT detector consists of a single large NaI scintillation crystal that rotates around the patient and is known as an Anger gamma camera [Anger, 1964]. The orbit of the rotating detector can be arbitrary, but in practice, it is circular or elliptical. As the detector rotates at different angles along this orbit, it collects counts of γ -ray photons emitted from the patient's body. Figure 2.1 shows a representation of a SPECT camera.

When a detector collects a photon, ideally, it is assumed that the photon emission occurred

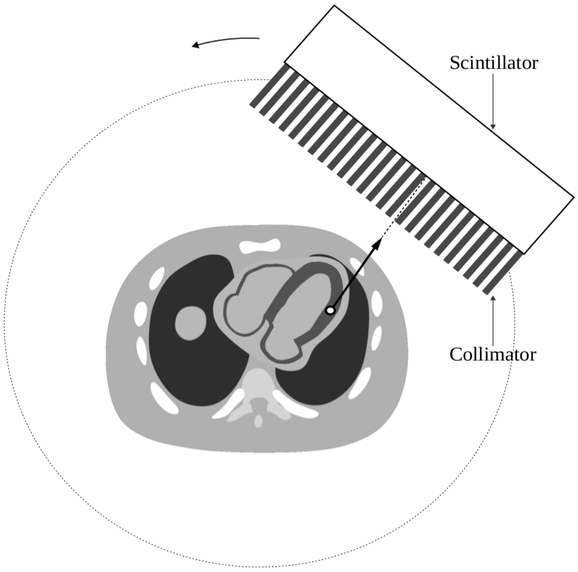

Figure 2.1: Simplified diagram for a Single Photon Emission Computed Tomography camera.

somewhere along a line of response (LOR), perpendicular to the detector surface. A high radioactivity distribution along a certain LOR leads to a high number of collected counts for this LOR at the detector. The count of detected photons for each LOR can be viewed as a projection (line integral) of the values of the radioactivity distribution along that line. All projection data, for a single transverse plane, are collectively called a sinogram and can subsequently be used to reconstruct an image of the radionuclide for that plane.

In SPECT, in order to obtain LOR information, a collimator is mounted in front of the scintillation crystal so that the gamma photons can only enter at known angles. The collimator holes are usually parallel to each other, although geometries with convergent or divergent openings also exist. As the detector rotates around the object (usually in a step-and-shoot mode), a full range of angles is obtained.

2.1.2 Detectors

A γ - ray detector is typically composed of a scintillation crystal that is coupled to a photomultiplier tube (PMT). The most commonly used scintillating material is thallium-activated sodium iodide (NaI(Tl)) . When a γ - photon interacts inside a scintillation crystal, the crystal absorbs the γ - ray photon and releases a burst of light photons onto the photocathodes of the PMTs. This burst of light is therefore converted in a short electrical signal by the PMTs and these signals are then counted so that the number of detected γ-photons is recorded.

In order to produce an image it is fundamental to know the LOR along which the gamma photon originated and the energy of the γ-ray interaction. It is therefore important to estimate the location of the detected γ-photons and their energy.

In an Anger camera [Anger, 1964], a large crystal is optically coupled with many PMTs (see Figure 2.2 - A). A PMT that is close to the position of interaction between the photon and the crystal will capture more light than a PMT that is further away from the scintillation event. The location of the interaction between the gamma ray and the crystal can be determined by processing the voltage signals from the photomultipliers; in simple terms, the location can be found by weighting the position of each photomultiplier tube by the strength of its signal, and then calculating a mean position from the weighted positions. The total sum of the voltages from each photomultiplier is proportional to the energy of the gamma ray interaction, thus allowing discrimination between different radio-tracers or between scattered and direct photons.
The emission of light photons is subject to uncertainty and it has a spatial distribution. Hence, the spatial location of a single scintillation event can only be measured with limited precision, which can be approximated by a 2-D Gaussian distribution [Helmer et al., 1967]. Typically, the full width at half maximum (FWHM) of the 2-D Gaussian distribution is used as a scalar measure to specify the spatial resolution and is referred to as intrinsic resolution. The intrinsic resolution of a gamma camera typically has a FWHM of 3-4 mm for a scintillation detector. The energy resolution depends on the material of the scintillation crystal. For the most commonly used crystal in SPECT (i.e. sodium iodide doped with thallium NaI(Tl)) the energy resolution is around 10% FWHM at 140 keV.

As an interesting alternative solid-state detectors based on Cadmium Zinc Telluride (CZT) crystals can be used. The most important feature of SPECT cameras based on CZT crystals is that they can convert and digitalize gamma radiation in a single step, eliminating the need for PMT technology. This allows the manufacturers to considerably reduce the bulkiness of the system and increase the mobility of the camera as both weight and size are drastically reduced. CZT are available as pixelated detector arrays with a typical intrinsic spatial resolution of 2.46 mm. Moreover CZT crystals offer higher energy resolution (in the range of $2-5\%$ for 140 keV) and intrinsic efficiency comparable to NaI(TI) [Butler et al., 2001] [Takahashi and Watanabe,

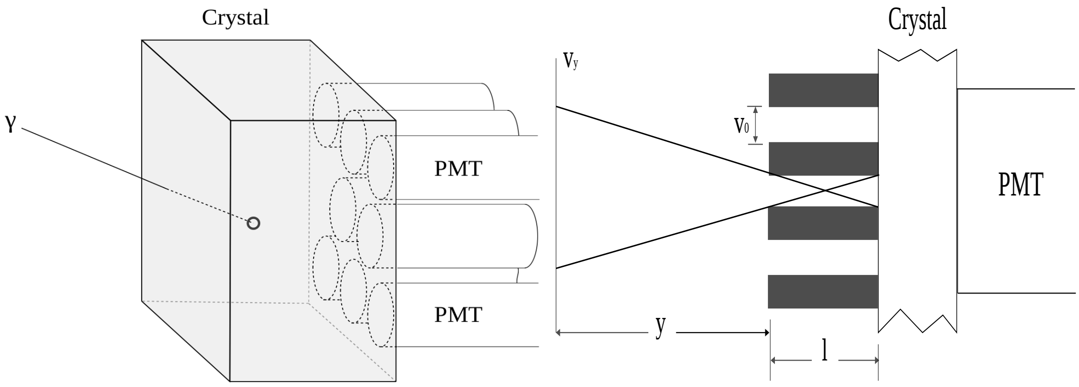

Figure 2.2: A (left) - Scintillation event on a SPECT detector. The scintillation crystal is optically coupled with many PMTs. B (right) - In SPECT, due to finite collimator length and hole size, photons are detected within some acceptance angle. This has the effect of increasing the uncertainty in location of the photons with increasing distance from the collimator.

2001].

2.1.3 Collimator

In SPECT, a collimator is mounted in front of the gamma camera. The purpose of the collimator is to achieve spatial localisation of the photon emissions in the imaging space. Since single photons are emitted isotropically, it would be otherwise impossible to obtain LOR information from the scintillation events in the crystal. The collimator is typically made of highly attenuating materials such as lead or tungsten and is meant to prevent oblique γ-photons from passing through to the scintillation crystal. Photons hitting on the collimator septa are ideally absorbed. In practice, although the collimator material greatly attenuates γ-rays, some photons pass through the lead with a probability given by Beer's law. This is called septal penetration. This penetration leads to an additional blur in the detector response.

Each collimator element allows the detection only of photons originating from a specific range of angles around the direction parallel to the collimator hole. Generally speaking, a narrow hole leads to a better LOR resolution but it also leads to a decrease in the number of photons that reach the scintillation crystal, thereby lowering the detector sensitivity.

The design of the collimator highly influences the detection sensitivity and the system resolution of a SPECT system, thus it is one of the key factors influencing the system response.

Even if various collimator designs have been presented in the literature, the most conventional collimator used for clinical SPECT is still the parallel hole collimator. It consists of a two-dimensional array of parallel holes whose shape can be circular, square or hexagonal. An enlarged view of a parallel hole collimator and of the detector for a SPECT system is shown in Figure 2.2 - B.

The need for collimation in SPECT systems is one of the main factors that contribute to shift-variance of the system response. Since, as illustrated in Figure 2.2 - B, photons originating from a specific range of angles enter the collimator hole, the system response varies as a function of distance from the detector. As a result, not all areas of the imaging volume contribute to the measurements in the same way. A photon entering at a particular hole of the collimator could have originated in a greater range of locations within the body if it originated at distance D_2 rather than at distance $D_1 < D_2$. Therefore, there is increasing spatial ambiguity with increasing distance from the detector. The response of a parallel hole collimator can be approximated by a depth-dependent Gaussian function whose FWHM increases linearly with distance.

A discussion on how the design of the collimator influences the system response and an analysis on the effects of the collimator design on the image quality will be presented in Chapter 5.

2.1.4 Collimator-Detector Response: Sensitivity and Resolution

As discussed in the previous sections, the projection of a point source at a single detector angular position is influenced by a number of factors related to collimator and detectors in gamma cameras, thus referred to as the collimator-detector response (CDR) . The CDR is characterised by two factors: the detection sensitivity and the system resolution.

Detection Sensitivity

The detection sensitivity can be defined as the percentage of emitted photons that are detected and recorded. It is determined by the detection efficiency of the γ - camera and by the geometric efficiency of the collimator.

The detection efficiency of the γ camera is given by the ratio between the number of photons that are actually counted by the detector and the number of photon that hit the surface of the scintillation crystal.

The geometric efficiency of a collimator is given by the ratio between the number of photons that hit the surface of the crystal and the number of photon emitted, in the ideal case in which the photons are not attenuated by the object.

In SPECT imaging, since physical collimation is needed in order to obtain LOR information, a large fraction of the incident photons are rejected. Collimators therefore exhibit low geometric efficiencies , of the order of $\sim 0.01\%$.

For the parallel hole collimator the geometric efficiency is nearly not distance-dependent. However, this is not always the case for other collimator designs.

System Resolution

The resolution of a SPECT system is determined by the intrinsic response of the gamma camera and the collimator geometric response.

Intrinsic response: Aside from the effect of collimators, the detector system itself demonstrates an intrinsic uncertainty in position estimation of incident gamma rays, as described in section 2.1.2. This uncertainty is referred to as intrinsic response and can be approximated by a 2-D Gaussian distribution.

Geometric response: As described in the previous section, the collimator dimensions define the acceptance angle within which incident photons are accepted. Subsequently, the geometric response function becomes wider with increasing distance from the collimator surface, and strongly depends on the particular design of each collimator. The uncertainty about the origin of the detected photons is modelled by a Point Spread Function (PSF) . The PSF of a gamma camera describes the photon count density distribution at the detector surface when a point source is imaged. For parallel hole collimators, the PSF is approximated by a Gaussian function whose width is determined by the collimator aperture. A scalar measure for the collimator resolution is then defined as the FWHM of that Gaussian, which is determined by the collimator geometry as well as the distance between the location of the point source and the gamma detector. Combining the intrinsic resolution and the collimator resolution yields the system resolution:

$$FWHM_{system} = \sqrt{FWHM_{detector}^2 + FWHM_{intrinsic}^2} \tag{2.1}$$

Typically, for parallel hole collimators, the system resolution is assumed to be shift-invariant on planes parallel to the detector surface but it is dependent on the distance from the detector.

2.1.5 Acquisition Model and Camera Trajectory

In SPECT imaging, the acquisition mode can be either planar or tomographic.

In planar imaging, a two dimensional (2D) projection is acquired at each detector head without performing the rotation of the gamma camera. Since anatomical structures along the projection direction are overlaid, the 3D information of the imaged subject is lost.

The aim of the tomographic imaging instead is to obtain a three dimensional (3-D) representation of the radiotracer distribution. The detector heads therefore need to rotate around the patient to obtain projection data from different angles. These projection data however cannot be interpreted directly but a reconstruction method needs to be applied in order to obtain a 3-D image.

The tomographic acquisition mode usually involves the rotation of the camera along a circular trajectory at constant speed around the centre of the imaging volume so that the detector collects photons during the same time interval for each angular position. Recently, new camera designs based on compact solid state detectors units (such as the D-SPECT system) allow movements that would not be achievable with conventional gamma cameras, allowing for a region centric acquisition (see chapter 7). Changing the trajectory, just like for any parameter of the acquisition system, the interdependence of the information changes. So, changing camera trajectory, one might collect more information about certain regions and less about other regions, leading to a shift-variance in the response.

The trajectory of the gamma camera has a profound effect on the uncertainty in estimation of the radio-tracer distribution. Therefore an accurate model of the camera trajectory is important for reconstruction.

2.1.6 Physical Effects: Attenuation and Scatter

The measurement data are influenced from certain physical effects that are not directly related to the detector response. Gamma photons are dependent on two physical effects known as attenuation and scatter.

Attenuation

A significant fraction of emitted photons are either absorbed or deviated as they interact with their surrounding material. As a result, the number of photons reaching the γ - camera is only a fraction of the number of photons emitted from the activity distribution within the object. This effect is generally called attenuation. Photons are attenuated as they travel through the surrounding material from their point of origin to the detector surface. The attenuation is an exponential function of the distance travelled and the attenuation coefficient of the surrounding materials μ (that expresses the probability of interaction per unit length). The survival probability of a photon [Macovski, 1983] as it travels between two points along the y-axis is given by:

$$p_{att} = \exp\left\{ - \int_{y_1}^{y_2} \mu(y) \mathrm{d}y \right\} \tag{2.2}$$

where $y_1 = (x, y_1, z)$ is the point of emission and $y_2 = (x, y_2, z)$ is a position on the collimator surface.

The effect of attenuation can be corrected, during reconstruction, by including the distribution of attenuation coefficients in the system model. In order to include attenuation effects, a transmission scan must be performed to estimate the attenuation map $\mu(x, y, z)$. For example, the attenuation map can be obtained from a Computed Tomography (CT) image which is registered with the SPECT image. However, since the attenuation coefficients are energy-dependent, the values from the CT image are first converted into the values corresponding to the energy of the emitted photons before they can be used for attenuation correction.

Scatter

A fraction of the emitted photons interact with free or bound electrons as they travel through light-materials such as water or soft tissues. When photons undergo this elastic collision with electrons, they lose part of their energy and they are deflected by an angle dependent on the amount of energy lost. This effect is called Compton scatter.

If a photon experiences Compton scatter, it has not originated along the LOR that is detected. This leads to a degradation in quantification and spatial resolution accuracy. Since scattered photons typically have less energy than unscattered photons, scattered events can be eliminated by energy discrimination. For example, in technetium-99m ($^{99\mathrm{m}}$Tc) SPECT imaging, the photopeak is at 140 keV and, if photons are detected at lower energies, they are considered to be scattered photons and they are rejected. However, since scintillating detectors have a limited energy resolution, some scattered photons will still be counted. For the most commonly used scintillation crystals in SPECT (i.e., NaI(Tl)), the energy resolution is around 10% FWHM at 140 kev. Solid-state CZT detectors present an energy resolution of 3-6$\%$ at 140 keV. They can therefore be considered an interesting alternative for the development of new SPECT cameras.

2.2 Discrete System Model

The acquisition model characterizes a linear operator $\boldsymbol{H}$, which we will refer to as the system operator, mapping the object space to the measurement space.

Generally speaking, when we model a real system, we must make certain approximations and therefore different models can have different properties. In tomography, several choices for the system model can be made; which include the idealised continuous model, used in analytical reconstruction, the continuous-to-discrete model and the discrete-to-discrete model.

Since, in practice, tomographic systems can acquire only a finite number of measurements and statistical reconstruction algorithms can estimate only a finite number of image intensity parameters, a discrete model is adopted in the rest of this thesis.

Let $\boldsymbol{\lambda}$ denote the radioactive tracer distribution to be imaged and underlying the measurement data $\boldsymbol{\nu}$ obtained by the imaging system. For this three dimensional (3D) model, both the measurement data and the object space are discretized. In a discrete representation of the object $\boldsymbol{\lambda}$ a set of basis function has to be chosen. We discretise the continuous function, expressing the rate of emission of γ-radiation, using a voxel basis, where $\boldsymbol{\lambda} = [\lambda_1, ..., \lambda_N]^T$ denotes the vector of emission rates. Our choice of the volume element for the basis function is the usual cubic voxel, but other shapes have been studied as well [Matej and Lewitt, 1996] [Yendiki and Fessler, 2004]. The basis functions are spatially localized volume elements, arranged on a grid over the object space. In the rest of the thesis, x and y denote the in-plane coordinates, z represents the axial coordinates of the discretized volume, N_x, N_y, N_z denote the number of voxels along each direction and $N = N_x \times N_y \times N_z$ denotes the total number of voxels.

In a discrete representation of the measurement data, the M photon counts collected by the detector are lexicographically reordered in a vector $\boldsymbol{\nu} = [\nu_1, ..., \nu_M]^T$. The detector unit responsible for collecting each of these M counts is referred to as a detector bin and is a conceptual unit since, in SPECT, it does not correspond to a physical detector element.

In emission tomography, the mean projection measurements $\overline{\boldsymbol{\nu}}$ are assumed to be related to the discretized object $\boldsymbol{\lambda}$, by the following discrete linear model:

$$\overline{\nu}_d(\boldsymbol{\lambda}) = \sum_{i=1}^{N} h_{di}\lambda_i \tag{2.3}$$

$$\overline{\boldsymbol{\nu}}(\boldsymbol{\lambda}) = \boldsymbol{H}\boldsymbol{\lambda} \tag{2.4}$$

the matrix $\boldsymbol{H}$ is the $M \times N$ system matrix whose elements $\{h_{di}\}$ represent the probability that photons emitted from voxel i are detected in detector unit d. It is possible to decompose the system matrix into multiple components that characterise the different physical aspects of the system response:

$$\begin{aligned} \overline{\nu}_d(\boldsymbol{\lambda}) &= \sum_{i=1}^{N} c_d a_{di} b_{di} s_i \lambda_i \\ &= [D[c_d](\boldsymbol{A} \cdot \boldsymbol{B}) D[s_i] \boldsymbol{\lambda}]_d \end{aligned} \tag{2.5}$$

where $D[\cdot]$ produces a diagonal matrix from the vector-valued argument.

For a typical SPECT model the c_d terms represent the detector efficiency, the a_{di} (or, equivalently, the matrix $\boldsymbol{A}$) terms represent the attenuation factors, b_{di} terms (or, equivalently, the matrix $\boldsymbol{B}$) denotes the geometric system model and the s_i terms denotes the spatial variation in sensitivity. The system matrix therefore models the propagation and detection of un-scattered photons, encompassing the depth-dependent response of the collimator, the position-dependent geometric efficiency, the scanning pattern of the detectors and attenuation through the propagating medium. In this model we do not take into account the contribution of scattered photons, though the system matrix $\boldsymbol{H}$, in principle, may encompass scatter events.

2.3 Noise Model

In Emission Tomography the inherent randomness in the activity emission results in randomness in the measurements, which can be treated as Poisson-distributed conditionally independent random variables [Rockmore and Macovski, 1976] [Yu and Fessler., 2000].

In the case of Emission Tomography, the conditional probability distribution associated with photon counting is therefore Poisson distributed with expectation $\boldsymbol{H\lambda}$:

$$p(\nu_d|\boldsymbol{\lambda}) = \frac{e^{-[\boldsymbol{H}\lambda]_d}([\boldsymbol{H\lambda}]_d)^{\nu_d}}{\nu_d!} \tag{2.6}$$

and the measurement data acquired by a real tomographic imaging system are related to the true radioactivity distribution, as in the following equation:

$$\boldsymbol{\nu} \sim Poisson\{\boldsymbol{H\lambda}\}, \tag{2.7}$$

If we consider $\boldsymbol{H}$ as known, following the discussion in the previous section, the image reconstruction problem in emission tomography consists of estimating $\boldsymbol{\lambda}$ given the data $\boldsymbol{\nu}$, where $\boldsymbol{\nu}$ is a realization of the random vector in (2.7).

2.4 Space-Invariant vs Space-Variant Systems

The definition of collimator-detector response has been presented in section 2.1.4. The collimator-detector response can be modelled in different ways.

A simple model of the collimator-detector response is the strip-integral model. In this model, illustrated in Figure 2.3 A , the CDR is simply modelled as a cuboid stripe. The b_{di} term is therefore proportional to the volume of intersection between the cuboid stripe corresponding to the d-th bin and the i-th voxel. Considering this model, the measurements are not depth-dependent and radially uniformly spaced. Therefore the system response is approximately space-invariant except for discretisation effects.

However in SPECT, as discussed in section 2.1.4, the projection of a point source at a single detector angular position is influenced by a number of factors related to collimator and detectors in gamma cameras. The CDR is typically modelled by a depth-dependent Point Spread Function (PSF) which takes the form of a Gaussian function. Since the b_{di} terms are chosen to specify a depth-dependent response with a Gaussian profile then the system response becomes space-variant.

Sample patterns for a single detector response are shown for the strip integral model and a depth-dependent Gaussian model in Figures 2.3 B and C, respectively.

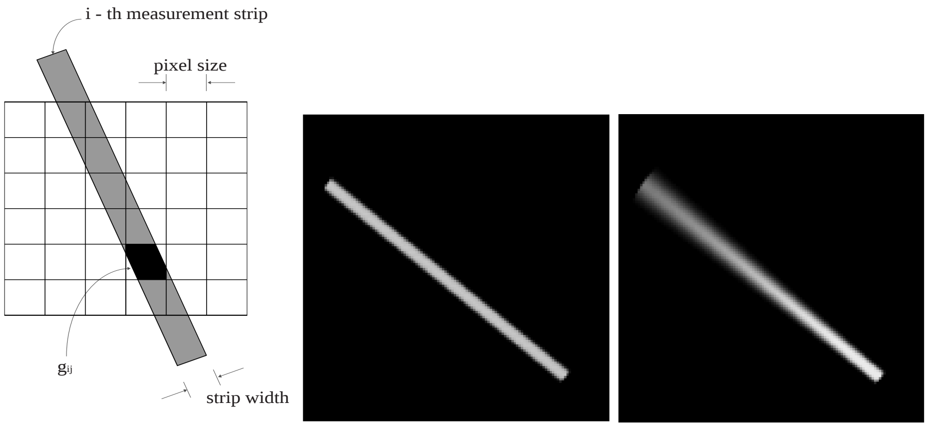

Figure 2.3: Modelling the detector response in a discrete tomographic system model. A (left) - Strip integral model. B (centre) - Sample pattern for an ideal detector-collimator response. C (right) - Sample pattern for a SPECT, depth-dependent, detector-collimator response.

The geometric response of the system is important when studying the properties of an imaging system. Thus, we discuss how the geometric response varies spatially.

For the discrete model, presented in section 2.2, the geometric response of the system is defined by:

$$\boldsymbol{B}'\boldsymbol{B}\mathrm{e}_i \tag{2.8}$$

where e_i is the unit basis vector for the voxel i. Hence, the geometric response is defined by the choice of the b_{di} terms in equation (2.5).

For an acquisition geometry that involves the rotation of the camera along a circular trajectory covering $360°$, we will have, on average, worst resolution for voxels near the centre of the Field Of View (FOV) rather than for voxels at the edges of the FOV; and since the resolution changes with detector angle, we will have an anisotropic response at the edges of the FOV. Moreover it should be noted that incorporating the effects of attenuation and detector efficiencies will lead to an increase in the space-variance of SPECT system response.

In Figure 2.4 A and B we illustrate the geometric response for the strip-integral model and for the depth-dependent Gaussian model respectively. We calculated the geometric response of the system, as in equation 2.8, for 37 voxels locations uniformly distributed on the central plane

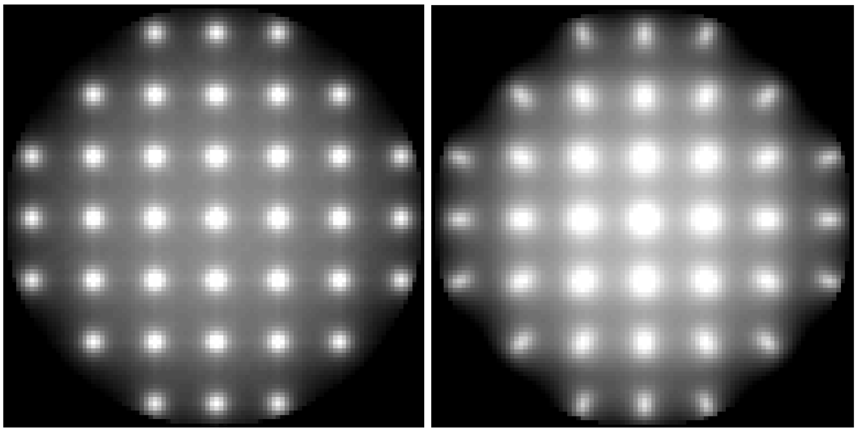

Figure 2.4: Comparison of space-invariant and space-variant geometric responses. A (left) - A set of space-invariant geometric responses for the strip integral model. B (right) - A set of space-variant geometric responses for the depth-dependent Gaussian model.

of the imaging volume. From Figure 2.4 we can see how the geometric response of the system is, in fact, shift-invariant; whereas the responses for the depth-dependent Gaussian model are shift-variant and anisotropic.

2.5 Tomographic Image Reconstruction

In equation (2.3) the measurement data $\boldsymbol{\nu}$ were defined in the form of projections through the true activity distribution $\boldsymbol{\lambda}$. These projection data can not be used directly to inform decision processes that arise in medical diagnosis. It would be extremely difficult for a physician to determine the exact position of an abnormality within the imaging volume by looking at the projection data only. Moreover, it is not possible to accurately quantify the radio-tracer uptake within a volume of interest directly from the measurement data. This implies the need for a reconstruction method. The reconstruction problem consists in recovering $\boldsymbol{\lambda}$ from its projections $\boldsymbol{\nu}$, therefore somehow returning the data in the projection space, back to the object space.

In the previous section, the concept of geometric response of a system was introduced, and we discussed how it depends on the collimator-detector response and its diverse physical characteristics. Nevertheless, the resolution of the reconstructed images can be significantly lower than this geometric response, as it is also affected by the reconstruction method applied to the data. This effect is undesirable, since important details may be lost if the resolution of reconstructed images is not fine enough. On the other hand, noise reduction is also an important feature of a reconstruction method, since noise artifacts degrade the quality of the image. Thus, since a trade-off between fine resolution and low noise level arises, the choice of the reconstruction method has a profound influence on the image quality.

Analytical reconstruction methods, such as Filtered Back Projection (FBP), have been extensively studied and widely clinically used in CT reconstruction. Such methods are based on the idealised continuos model. They calculate an estimate of the activity distribution, from the measurement data, ignoring measurement noise in the formulation of the problem and treating the arising noise-related problems by post-filtering operations [Natterer, 1986]. Moreover, analytical methods require certain standard geometries (e.g., parallel rays and complete sampling in radial and angular coordinates) and therefore incorporating a detailed SPECT system model in FBP methods is not straightforward. In this PhD work, we focus on statistical methods for image reconstruction, since they can overcome all these limitations.

In this section, some commonly-used statistical reconstruction methods are described, discussing how each of them deals with the trade-off that arises between resolution and noise properties.

2.5.1 Statistical Reconstruction

Statistical reconstruction methods are based on the discrete model presented in section 2.2 where both the measurement data and the object space are discretized.

Using a statistical reconstruction method we seek to estimate $\boldsymbol{\lambda}$ from $\boldsymbol{\nu}$ by maximizing an objective function $\Psi(\boldsymbol{\lambda})$ of the form

$$\Psi(\boldsymbol{\lambda}) = L(\boldsymbol{\nu}, \boldsymbol{H}\boldsymbol{\lambda}) - \beta R(\boldsymbol{\lambda}) \tag{2.9}$$

$$\hat{\boldsymbol{\lambda}} = \arg\max_{\boldsymbol{\lambda} \geq 0} \Psi(\boldsymbol{\lambda}) \tag{2.10}$$

where $L(\cdot,\cdot)$ and $R(\cdot)$ are referred to as the likelihood function and the penalty function respectively. $\beta \geq 0$ is a regularization parameter (or hyper-parameter) that controls the relative weight of the two terms. The non-negativity constraint $\boldsymbol{\lambda} \geq 0$ is needed only on physical grounds, rather than mathematical. In Emission Tomography, in fact, $\boldsymbol{\lambda}$ consists of photon emission densities and therefore can not include negative values.

Equation (2.9) defines a penalized likelihood reconstruction method when the choice for $L(\cdot,\cdot)$ is the log of the conditional likelihood $p(\boldsymbol{\nu}|\boldsymbol{\lambda})$ of the data $\boldsymbol{\nu}$ given the true activity distribution $\boldsymbol{\lambda}$. In the case of Emission Tomography, where the conditional probability distribution associated with measurement data is Poisson; the likelihood $p(\boldsymbol{\nu}|\boldsymbol{\lambda})$ is defined as in equation (2.6) and consequently the log-likelihood equals:

$$L(\boldsymbol{\lambda},\boldsymbol{\nu}) = \log p(\boldsymbol{\nu}|\boldsymbol{\lambda}) = \sum_d (-[\boldsymbol{H}\boldsymbol{\lambda}]_d + \nu_d \log([\boldsymbol{H}\boldsymbol{\lambda}]_d) - \log \nu_d!) \tag{2.11}$$

There is no closed analytical form that expresses explicitly $\hat{\boldsymbol{\lambda}}$ in terms of $\boldsymbol{\nu}$, when the probability distribution associated with measurement data is Poisson. In other words, the cost function (2.10) defines $\boldsymbol{\lambda}$ only implicitly as a function of $\boldsymbol{\nu}$, hence iterative algorithms are needed.

Assuming that the cost function in equation (2.9) has a unique maximiser and that we iterate the algorithm until the estimate $\hat{\boldsymbol{\lambda}}$ converges to the maximum, the choice of the algorithm has no effect on the quality of the estimate $\hat{\boldsymbol{\lambda}}$. A sufficient condition for the uniqueness of the solution is that the cost function in equation (2.9) is strictly concave and coercive [Ahn and Fessler, 2003].

Setting $\beta = 0$ in (2.9), we rely on an unregularised reconstruction method, which maximises the $L(\cdot,\cdot)$ function alone. This method is referred to as the Maximum Likelihood Estimation (MLE) . However, due to the fact that the problem is ill-conditioned, to the fact that the maximizer in (2.10) may be not unique, and to the noise in the measurement data; an unregularised reconstruction method yields unacceptably noisy estimates. A strategy to reduce the noise is therefore needed.

We can find in the literature a wide range of methods to reduce the noise in the reconstruction. Each method presents certain advantages and disadvantages. In the following we present a brief discussion of several, popular methods.

2.5.2 Stopping-Rule

For many algorithms, one of the simplest approaches to noise reduction is to initialise the algorithm with a smooth image and then to stop iterating well before convergence [Hebert, 1990]

[Veklerov and Llacer, 1987] [Barrett et al., 1994]. As an example, if we initialise the algorithm with a uniform image, and we look at the estimate after each iteration, we can see how the reconstructed image starts as very smooth and becomes visually more noisy with increasing iteration, as high frequency components begin to appear [Barrett et al., 1994]. Therefore, stopping the algorithm before convergence acts as a kind of regularisation and the number of iterations then becomes the parameter that controls the noise/resolution trade-off. Using this method for noise reduction, the quality of the estimate becomes dependent on the image used for initialisation. In section 2.5.5 we will discuss several algorithms that are used for performing the objective function maximisation in (2.9). Stopping rules are only meaningful for those algorithms, i.e. MLEM [Shepp and Vardi, 1982] and OSEM [Hudson and Larkin, 1994]. For other algorithms, i.e. coordinate ascent algorithms, stopping the algorithm before convergence has no noise regularisation effect, since the high frequencies converge just as fast as the low frequencies. Thus, the resolution properties become dependent on the iterative algorithm that is used and on how far it is iterated.

2.5.3 Post-Smoothed Maximum Likelihood

Another method is to allow the unregularised maximum-likelihood algorithm to iterate until convergence and then to achieve noise reduction by post-smoothing the noisy reconstructed estimate [Beekman et al., 1998] [Slijpen and Beekman, 1999]. It is also possible to filter the estimated image after each iteration [Jacobson et al., 2000] [Silverman et al., 1990]. The maximum-likelihood estimator iterated until convergence attempts to obtain an ideal reconstruction with a delta impulse response. The resolution properties of the estimator are therefore completely determined by the post-smoothing filter operator which can be customised in order to control the resolution/variance trade-off. An advantage of this method is that a variety of post-smoothing filters can be applied at the expense of a single iterative solution.

However this method presents also disadvantages. Since the problem is typically ill-conditioned, unregularised MLEs tend to take many iterations to converge to a solution. Moreover, in case where there are multiple maximisers, and therefore multiple solutions, applying a post-smoothing filter to the reconstructed estimate will not make the solution unique. If the system matrix $\boldsymbol{H}$ has full rank, then the ML estimate is unique in the case where there are no non-negative solutions to the system of equations $\boldsymbol{\nu} = \boldsymbol{H}\boldsymbol{\lambda}$. In practice, however, this full rank

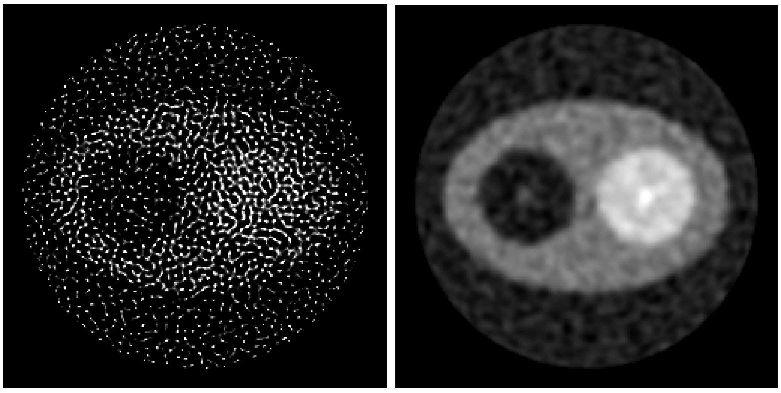

Figure 2.5: Statistical reconstruction with and without regularisation. Reconstruction of measurements obtained for a piecewise constant image are plagued by noise in a pure maximum-likelihood reconstruction (left), but the noise can be greatly reduced by regularising the reconstruction (right). A (left) - Maximum Likelihood Reconstruction. B (right) - Regularised Likelihood Reconstruction.

property is quite difficult to verify.

2.5.4 Penalised-Likelihood Estimation

The introduction of the regularisation function $R(\cdot)$ in equation (2.9) corresponds to the Penalised likelihood Estimator (PLE) estimator with a smoothness prior. The function $R(\cdot)$ is typically a roughness penalty that has a smoothing effect on the estimate by penalising the intensity differences between neighbouring voxels [Geman and McClure, 1985][Hebert and Leahy, 1989][Lange, 1990]. With the introduction of an appropriate penalty term $R(\cdot)$, we achieve not only noise reduction, but also convergence acceleration and we ensure the uniqueness of the solution.

Making use of the PLE, the resolution/variance trade-off in the reconstructed image is controlled by the regularisation parameter (or hyper-parameter) β. A smaller β in equation (2.9) gives more weight to the likelihood function $L(\cdot,\cdot)$, resulting in a noisier estimate, whereas a larger β gives more weight to the penalty function, resulting in a smoother estimate. Two examples of reconstructions of noisy measurement data iterated to nearly reach convergence using a unregularised ML Estimator and using a PL Estimator with a large value of β, are shown in Figure 2.5 - A and Figure 2.5 - B respectively.

The actual form of the penalty function itself also significantly affects the quality of the estimate.

The penalty function can take a wide variety of forms. However, we will restrict this discussion to the description of the most commonly used prior, which is the quadratic prior:

$$R(\boldsymbol{\lambda}) = \sum_i \frac{1}{2} \sum_{k \in K_i} w_{ik}(\lambda_i - \lambda_k)^2 \tag{2.12}$$

where K_i indicates the set of voxels whose intensities are to be combined . Typically, differences only between each element and its nearest neighbours are defined in (2.12) and this set of voxels is called a neighbourhood of voxel i (or clique).

A first order penalty includes voxel's vertical and horizontal neighbours, whereas a second order penalty included also the voxel's diagonal neighbours [Hebert and Leahy, 1989].

The penalty weights w_{ik} give the possibility to specify a different amount of regularisation for each neighbourhood, allowing for local resolution control. In order to assure that (2.9) has a unique solution when the likelihood function $L(\cdot, \cdot)$ is concave, we choose the quadratic penalty function (which is a symmetric concave function) to be a non-negative definite function. Therefore, the weights w_{ik} are constrained to be non-negative.

If the weights are identical, regardless of the voxel location in the object space, $R(\boldsymbol{\lambda})$ is called a shift-invariant penalty. If the penalty is shift-invariant and weights are identical for all pairs in the neighbourhood, the penalty is called a uniform penalty. For example, a conventional uniform first-order penalty is defined by $w_{ik} = 1$ for the horizontal and vertical neighbours and $w_{ik} = 0$ otherwise. A uniform second-order penalty includes $w_{ik} = 1/\sqrt{2}$, for the diagonal neighbours (where the $\sqrt{2}$ term is a distance scaling).

Penalised-likelihood (PL) methods present advantages over other regularisation techniques. Using PL methods the conditioning of the reconstruction problem is improved and the convergence rate of the iterative algorithm used for optimisation tends to increase.

However, the use of penalty functions present some possible disadvantages. The most notable disadvantage is the non-intuitive relation between the penalty function and the resolution properties of the reconstructed image. The use of a shift-invariant post-smoothing approach leads to a shift-invariant resolution response, whereas the use of a shift-invariant uniform penalty function results in a shift-variant resolution response (as we will discuss in chapter 3). However, such limitations may only represent problems with conventional penalty functions. Specially designed penalty functions to correct for space-variant resolution properties inherent in PL estimators have been introduced in [Stayman and Fessler, 2000].

2.5.5 Reconstruction Algorithms

Several algorithms are presented in the literature, that are used for performing the objective function maximisation in (2.9). In the case in which we are using a reconstruction method in order to maximise an objective function; the algorithm itself does not affect the solution. Only the speed at which the algorithm reaches convergences is affected by the characteristics of the algorithm.

The most commonly used iterative algorithm in emission tomography is the Maximum Likelihood Expectation Maximisation (ML-EM) algorithm, which is based on EM algorithm for the maximisation of the likelihood. The EM iteration alternates between performing an expectation (E) step, which creates a function for the expectation of the log-likelihood evaluated using the current estimate for the parameters, and a maximisation (M) step, which computes parameters maximising the expected log-likelihood found on the E step. First, ML-EM, when applied to the Poisson model, guarantees non-negativity of the solution; second, ML-EM guarantees convergence (for proof of convergence, see [Shepp and Vardi, 1982]).

Certain algorithms are not guaranteed to converge. For example, practical application of the expectation maximisation (EM) algorithm has been facilitated by the introduction of ordered subsets (OS) [Hudson and Larkin, 1994]. The OS method consists in accelerating the algorithm by processing only a subset of projection data at each iteration. However, how these subsets are chosen affects the image that is reconstructed, and therefore the algorithm may not be guaranteed to converge.

For PL estimation, the EM algorithm can not be applied (at least not in the formulation presented in [Shepp and Vardi, 1982]), since a closed form solution for the M-step does not exist for a generic choice of the penalty function. In order to calculate the PL estimate, an interesting possibility is to use the One Step Late EM (OSL-EM) introduced by Green [1990]. OSL-EM is not guaranteed to converge and does not impose non-negativity. However the algorithm behaves well in most practical cases as long as the prior has a relatively small importance. Several algorithms to calculate the PL estimate with guaranteed convergence have been developed. As an example, Pierro [1995] gave an alternative derivation of the MLEM algorithm using surrogate functions, which leads to a natural extension of the MLEM algorithm to penalised likelihood estimators, with guaranteed convergence. The algorithm presented in [Pierro, 1995] is no longer an expectation maximisation algorithm, but it is a convergent algorithm that becomes identical to MLEM if the weight of the penalty is set to zero.

In the previous paragraphs, a discussion on how the form of the Likelihood function $L(\cdot, \cdot)$ and of the Penalty function $R(\cdot)$ affects the quality of the estimate has been presented. All the algorithms described in the previous paragraphs contain a forward projection and a backprojection operation (although the backprojection is not always necessary, e.g. simulated anealing). Thus, another characteristic of the model in use, that affects the quality of the reconstructed image, is the accuracy of the system model $\boldsymbol{H}$. As an example, a less accurate backprojector $\boldsymbol{H_b} \neq \boldsymbol{H'}$ is sometimes substituted for adjoint operator in order to accelerate the reconstruction algorithm. However, since a typical iterative algorithm applies both the forward projector $\boldsymbol{H}$ and the backprojector $\boldsymbol{H'}$ once at each iteration, the error due to a mismatch between the two operators may accumulate, resulting in artifacts [Zeng and Gullberg, 2000]. Therefore a trade-off between accuracy of the system model and computational burden (associated with a more complex model) arises.

In this thesis, we have used for reconstruction an accelerated GPU (graphics processing unit) implementation of the ML-EM algorithm for ML estimation and of the OSL-EM algorithm for PL estimation, implemented as part of the NiftyRec toolbox Pedemonte et al. [2010]. In Nyftirec, the backprojector operator is carefully implemented to match the adjoint of the projector operator. The error between the two operators is given by:

$$\frac{\langle \boldsymbol{H}\boldsymbol{\lambda}, \boldsymbol{\nu} \rangle}{\langle \boldsymbol{\lambda}, \boldsymbol{H}^{*}\boldsymbol{\nu} \rangle} - 1 = 0.0016 \quad \forall \boldsymbol{\lambda}, \boldsymbol{\nu} \in \mathbb{R} \tag{2.13}$$

This mismatch of the 0.16% will be assumed in the rest of this thesis to not affect the reconstructed image quality.

2.6 Data Collection

In the next two chapters, different methods for the characterisation of the image quality and several Figures Of Merit (FOMs) will be presented.

The data used for image quality assessment in emission tomography can be collected by performing real measurements of phantoms, specially designed for different FOMs. As an example, the Jaszczak phantom, a phantom consisting of line sources with varying diameters, is often used for resolution studies. A phantom consisting of hot and cold spots on a uniform background is also used for evaluation of contrast recovery.

In the last two decades, several numerical simulation algorithms based on the Monte Carlo method have also been developed to simulate commercially available imaging systems for SPECT [Jan et al., 2004] [Toossi et al., 2010]. These Monte-Carlo simulators are capable of accommodating complex scanner geometry and imaging configurations, while including comprehensive physics modelling abilities. Millions of photons are simulated and tracked starting from their emission point, and taking into account the photon-electron interactions in the patient body tissues and in the detector materials. The use of a Monte-Carlo simulator therefore permits the implementation of a realistic system matrix, at the price of time consuming simulations.

In this PhD work we perform simulations using computer phantoms and the simplified convolution-based forward projector implemented as part of the NiftyRec toolbox [Pedemonte et al., 2010]. The use of computer phantoms and of a convolution-based projector algorithm to simulate data acquisition, is less realistic than either performing real measurements or using a Monte-Carlo simulator, but sufficiently accurate for our purposes and much less time consuming. The use of a convolution-based projector algorithm permits the generations of thousands of noise realisations from the same noiseless projection data. A process that would be impractical or extremely time consuming with one of the other two methods described above. As will be shown in the next chapter, being able to reconstruct thousands of noise realisation is in fact fundamental if one wants to perform a statistical calculation of the image quality.

Chapter 3

Statistical Evaluation of Image Quality

In this chapter we summarise the statistical approach for the calculation of the uncertainty in the reconstruction.

In chapter 2 section 2.5.1, the statistical reconstruction method to estimate the activity distribution $\boldsymbol{\lambda}$ from the measurement data $\boldsymbol{\nu}$ has been introduced. A reconstruction algorithm provides an image of the activity distribution that is most likely, according to the characteristics of the imaging system and to prior belief, to have generated the photon interactions measured by the imaging system. Such measurement is uncertain due to the limited amount of information that the scan may acquire. Thus, it is essential to estimate the uncertainty of the measurement produced by a given system set-up, in order to compare it with other set-ups, thereby enabling the optimisation of the design of an imaging system.

The estimation of the uncertainty in the reconstruction is an application-dependent problem, since it is highly influenced by many factors including the system design, the estimator, the specific reconstruction algorithm and the activity uptake and distribution. Therefore, as we will show in the following sections of this chapter, the calculation of the uncertainty is not only system- and estimator-dependent but it is also object-dependent.

Since this PhD work focuses on system optimisation, it is important to decouple the calculation of the uncertainty from the properties of the specific reconstruction algorithm in use. The Figure Of Merits (FOMs) presented in this chapter, such as mean and covariance of the estimator, are therefore derived using the assumption that the estimate is computed by completely maximising the objective function.

In order to (effectively) maximise the objective function, the reconstruction algorithm must be iterated to convergence. Therefore these FOMs are not applicable to unregularised methods that make use of a stopping rule to terminate the reconstruction algorithm at early iterations (before the maximiser is reached).

There has been much work in the literature, on analysing the statistical properties of an estimator as a function of the number of iterations. Barrett et al. [1994] introduced a method to estimate the variance by identifying how noise is propagated through EM iterations. These methods have been extended to identify the mean and variance properties of PL and OSEM estimates as a function of iteration [Wilson et al., 1994] [Soares et al., 2000] [Qi, 2003] [Soares et al., 2005]. The FOMs introduced in this chapter are somewhat easier to use because they are independent on the specific reconstruction algorithm, provided that sufficient iterations are used to maximise the cost function.

The FOMs presented in this chapter are also derived assuming that the estimator objective function has a unique solution and therefore a unique global maximiser. To ensure uniqueness of the solution, an appropriate regularisation penalty is included in order to obtain a strictly concave objective function. Bias is however unavoidable for Penalised-Likelihood (PL) estimators and an effective method for quantifying resolution properties of such estimators is needed.

A simple technique for investigating the resolution properties of an imaging system is to reconstruct a test phantom consisting of line sources with varying diameters. Imaging a test phantom we directly show the features that can be resolved. However, the resolution of these lines would depend not only on their dimensions, but also (non-linearly) on the intensity of the lines, on the intensity of the background the lines are immersed in and even on activity distribution in regions that are possibly far from the location of the lines. Resolution is therefore said to be spatially and object dependent.

In order to account coherently for the spatial and object dependency of the resolution, more specific tasks need to be considered. To this extent, task-specific FOMs, based on the performance of human or mathematical observers in classification, such as the detection of a certain class of tumours, have been defined and explored by Barrett et al. [1998], Barrett et al. [2006], Khurd and Gindi [2005], Gifford et al. [2005], Gilland et al. [2006], Yendiki and Fessler [2006].

Moreover, SPECT system optimisation can be evaluated based on the fundamental trade-off between bias and variance that can be achieved in the reconstruction of emission tomograms

[Nuyts, 2009] [Vunckx et al., 2008a] [Vunckx et al., 2008b] [Zhou et al., 2010]. In this chapter, a method to quantify resolution and variance for PL estimators will be discussed and the concept of Linearised Local Impulse Response(LLIR) will be introduced.

3.1 Statistical Method for the Calculation of the Uncertainty in the Estimation

In chapter 2 section 2.5.1, an objective function $\Psi(\boldsymbol{\lambda})$ has been introduced. This objective function depends on the unknown parameters $\boldsymbol{\lambda}$ and the noisy measurements $\boldsymbol{\nu}$ and can be expressed as:

$$\Psi(\boldsymbol{\lambda}, \boldsymbol{\nu}) = L(\boldsymbol{\nu}, \boldsymbol{H}\boldsymbol{\lambda}) - \beta R(\boldsymbol{\lambda}) \tag{3.1}$$

where $L(\cdot, \cdot)$ and $R(\cdot)$ are referred to as the likelihood function and the penalty function respectively. An estimator $\hat{\boldsymbol{\lambda}} = \hat{\boldsymbol{\lambda}}(\boldsymbol{\nu})$ was defined as the constrained maximizer of the following cost function:

$$\hat{\boldsymbol{\lambda}} = \arg\max_{\boldsymbol{\lambda} \geq 0} \Psi(\boldsymbol{\lambda}, \boldsymbol{\nu}) \tag{3.2}$$

In the following, it will be assumed that $\Psi(\cdot, \boldsymbol{\nu})$ has a unique solution and therefore a unique global maximiser $\hat{\boldsymbol{\lambda}} = \hat{\boldsymbol{\lambda}}(\boldsymbol{\nu})$.

In practice, since there is no closed form for (3.2), a reconstruction algorithm is iterated to convergence in order to maximize this objective function. A reconstruction algorithm therefore provides an estimate of the radioactivity distribution but such estimation is uncertain, due to the limited amount of information that the scan may acquire.

The reference method, or frequentist approach, used to calculate the uncertainty of the estimation involves the calculation of $\hat{\boldsymbol{\lambda}}$ for N_r independent experiments, where the expected measurement values $\overline{\boldsymbol{\nu}}$ are kept fixed, while the noise is sampled from the Poisson distribution.

Assuming the number of noise realisations N_r to be a large number, the mean of the estimate over the noise realisations

$$\check{\boldsymbol{\lambda}} = \frac{1}{N_r} \sum_{n=1}^{N_r} \hat{\boldsymbol{\lambda}}(\boldsymbol{\nu}_n) \tag{3.3}$$

can be considered as a good approximation of the expectation value $\mathbb{E}[\hat{\boldsymbol{\lambda}}]$.

The FOMs to characterise the statistical properties of the estimator are defined in the following.

Bias

The Bias is defined as the difference between the expectation value of the estimate and the true values of the parameters to be estimated (the true activity distribution):

$$Bias(\hat{\boldsymbol{\lambda}}) = \mathbb{E}[\hat{\boldsymbol{\lambda}}] - \boldsymbol{\lambda} = \check{\boldsymbol{\lambda}} - \boldsymbol{\lambda} \tag{3.4}$$

Variance

The variance of an estimator can be defined as follows:

$$\mathrm{Var}(\hat{\boldsymbol{\lambda}}) = \mathbb{E}[(\hat{\boldsymbol{\lambda}} - \mathbb{E}[\hat{\boldsymbol{\lambda}}])^2] = \mathbb{E}[(\hat{\boldsymbol{\lambda}} - \check{\boldsymbol{\lambda}}])^2] \tag{3.5}$$

If the variance have to be estimated from the data themselves, the sample variance can be calculated as:

$$\mathrm{Var}(\hat{\boldsymbol{\lambda}}) = \frac{1}{N_r}\sum_{n=1}^{N_r}\left(\hat{\boldsymbol{\lambda}}(\boldsymbol{\nu}_n) - \mathbb{E}[\hat{\boldsymbol{\lambda}}]\right)^2 = \frac{1}{N_r}\sum_{n=1}^{N_r}\left(\hat{\boldsymbol{\lambda}}(\boldsymbol{\nu}_n) - \check{\boldsymbol{\lambda}}\right)^2 \tag{3.6}$$

Covariance

The Covariance provides a measure of the strength of the correlation between one variable and all the other variables. If one is interested in how the variation in one voxel i is correlated to the variation in all the other voxels, it is possible to calculate the Covariance matrix for each pair of variables as:

$$\mathrm{Cov}(\hat{\lambda}_i, \hat{\lambda}_j) = \mathbb{E}[(\hat{\lambda}_i - \mathbb{E}[\hat{\lambda}_i])(\hat{\lambda}_j - \mathbb{E}[\hat{\lambda}_j])] = \mathbb{E}[(\hat{\lambda}_i - \check{\lambda}_i])(\hat{\lambda}_j - \check{\lambda}_j])] \tag{3.7}$$

If the variance have to be estimated from the data themselves, the sample covariance for a voxel i can be calculated as:

$$\begin{aligned}\mathrm{Cov}(\hat{\lambda}_i) &= \sum_{n=1}^{N_r}\sum_{j}^{N}\left(\hat{\lambda}_i(\boldsymbol{\nu}_n) - \mathbb{E}[\hat{\lambda}_i]\right)\left(\hat{\lambda}_j(\boldsymbol{\nu}_n) - \mathbb{E}[\hat{\lambda}_j]\right) \\ &= \sum_{n=1}^{N_r}\sum_{j}^{N}\left(\hat{\lambda}_i(\boldsymbol{\nu}_n) - \check{\lambda}_i\right)\left(\hat{\lambda}_j(\boldsymbol{\nu}_n) - \check{\lambda}_j\right)\end{aligned} \tag{3.8}$$

3.1.1 Variance as a Function of the Number of Iterations

The FOMs presented in the previous section are derived using the assumption that the estimate is computed by completely maximising the objective function.

In order to effectively maximise the objective function, the reconstruction algorithm must be iterated to nearly reach convergence. The number of iterations needed to reach the point of convergence depends on many factors; the system design, the object to be imaged, the specific reconstruction algorithm and the penalty function. For every experiment presented in chapters 5, 6 and 7, the objective function, $\Psi = L - \beta R$, will be plotted as a function of the number of iterations to visually show the rate of convergence for the specific system under investigation.

In this section, a simple experiment that visually shows the variance properties of an ML estimator as a function of EM iterations is presented.

The estimator variance is here calculated for a SPECT system. The SPECT system is based on a detector that rotates 360° around the centre of the imaging volume. The imaging volume dimensions are $128 \times 128 \times 1$ cubic voxels of 2.46 mm. The detector efficiency terms c_d, the attenuation factors a_{di} and the spatial variation in sensitivity terms s_i are set to 1. The geometric response is depth-dependent and therefore shift variant. The SPECT camera is placed at a distance of 162 mm from the centre of the imaging volume and is equipped with a Low Energy High Resolution collimator (LEHR). The intrinsic response is set to 3.6 mm, the collimator has a linearly varying depth-dependent Gaussian response that has a slope of 0.0533, which corresponds to 9.44 mm FWHM at the centre of the field of view.

The phantom used was a uniform disk positioned at the centre of the image space, with diameter $D = 157$ mm. The level of activity in the background was set to the 10% of the activity in the disk. From this activity distribution, projection data, $\overline{\boldsymbol{\nu}}$, has been simulated with 1 million total mean counts. A series of independent noise realisations was computed using a pseudo-random Poisson noise generator (from the IRT toolbox [Fessler, 2012]) The projection data have been reconstructed using the ML-EM algorithm included in the Niftyrec toolbox [Pedemonte et al., 2010]. Images of the variance for the ML estimator as a function of iterations are presented in Figure 3.1. From the image profiles in 3.1 - B we can remark that when the algorithm is stopped at 20 iterations the variance in the centre of the disk is lower than the variance at the borders. This effect is due to the fact that a SPECT system has a coarser resolution in the centre of the FOV rather then at the edges. In fact, even if the depth-dependent geometric response is modelled both in the projection and in the backprojection operator, if one stops the reconstruction algorithm at early iterations, a spatial dependent bias is introduced in the estimate.

When the reconstruction is stopped at 200 iterations the algorithm starts compensating for the depth-dependent geometric response and the variance has approximatively uniform values in the whole disk.

Finally, at 2000 iterations, if the system matrix $\boldsymbol{H}$ is non-singular and when the algorithm almost reaches the point of convergence, the off-centre voxels have a lower variance than the central voxels of the phantom. When the algorithm is iterated almost to convergence, it fully compensates for the geometric response, leading to an estimate that is approximately unbiased. Under the assumption that the estimate is approximately unbiased, the variance becomes the only FOM to quantify the uncertainty in the reconstruction. Therefore, Figure 3.1 shows a fairly intuitive result: a coarser geometric response in the centre of the FOV introduces higher uncertainty in the estimation and thereby a higher variance in that region of the phantom.

3.1.2 Estimated Variance as a Function of the Number of Noise Instances

In this section the estimated variance properties of a ML estimator, for a shift-variant SPECT system, are studied as a function of the number of noise instances.

In section 3.1, it has been stated that the frequentist approach for the calculation of the uncertainty involves the estimation of $\hat{\boldsymbol{\lambda}}$ for N_r independent experiments, where the expected measurement values $\overline{\boldsymbol{\nu}}$ are kept fixed, while the noise is sampled from the Poisson distribution. In this section we consider what happens when the number of noise instances N_r increases.

Images of variance for the ML estimator as a function of the number of noise instances are presented in Figure 3.1. The system under analysis is equivalent to the idealised SPECT system presented in the previous section.

The variance image obtained from the reconstruction of 20 noise instances is rather noisy, due to the finite number of repeated experiments. From Figure 3.1 it is possible to notice that the variance images become smooth with increasing number of noise realisations. In fact, the variance images, generally speaking, should be smooth because neighbouring voxels are affected by similar levels of noise.

In chapter 5, 6 and 7, it will be shown how, for photon counts typical of SPECT imaging, as the number of noise instances increases, the penalised maximum-likelihood estimator attains equality of the Uniform Cramér-Rao bound asymptotically and the estimated variance closely resembles the approximated deterministic calculation of the variance based on the Fisher Information method (presented in the next chapter).

3.2 Quantifying Resolution

In the two previous sections, the asymptotic properties of the ML estimator have been briefly discussed. Under the assumption that the system matrix $\boldsymbol{H}$ is non-singular and the positivity constraint is not active, the ML estimator is likely to be asymptotically efficient and asymptotically unbiased. Therefore, in the hypothetical case in which one performs an infinite sequence of experiments, iterating to convergence the algorithm used to maximise the ML objective function, the estimator attains equality of the Cramér-Rao bound (presented in section 4.4) and an expectation that equals $\boldsymbol{\lambda}$. However, in practice, the full rank property of the system matrix (non-singularity) seems quite difficult to verify and the conditions of asymptotic efficiency and asymptotic unbiasedness are not so certain. This problematic circumstance is addressed by including an appropriate regularisation penalty in equation (3.2) that leads to a strictly concave objective function (and therefore the maximiser in 3.2 to have an unique solution). Bias however is unavoidable for PL estimators, and an effective method for quantifying the resolution properties of such estimators is needed.

In this section, methods to quantify resolution are discussed and the concept of Linear Local Impulse Response (LLIR) for penalised-likelihood estimators is introduced.
Before the Local Impulse Response is discussed as a method for resolution quantification, some sample reconstructions using a PL estimator, where non-uniform resolution properties are readily apparent, are presented. The aim is to demonstrate that non-uniform resolution properties can arise even in reconstructions from noiseless projection data and from intrinsically shift-invariant systems. The resolution properties of an image are, in fact, not only system- and estimator-dependent but are also object-dependent. A simple technique for investigating the resolution properties of an imaging system is to reconstruct a test phantom. Imaging a test phantom, we directly show the features that can be resolved.

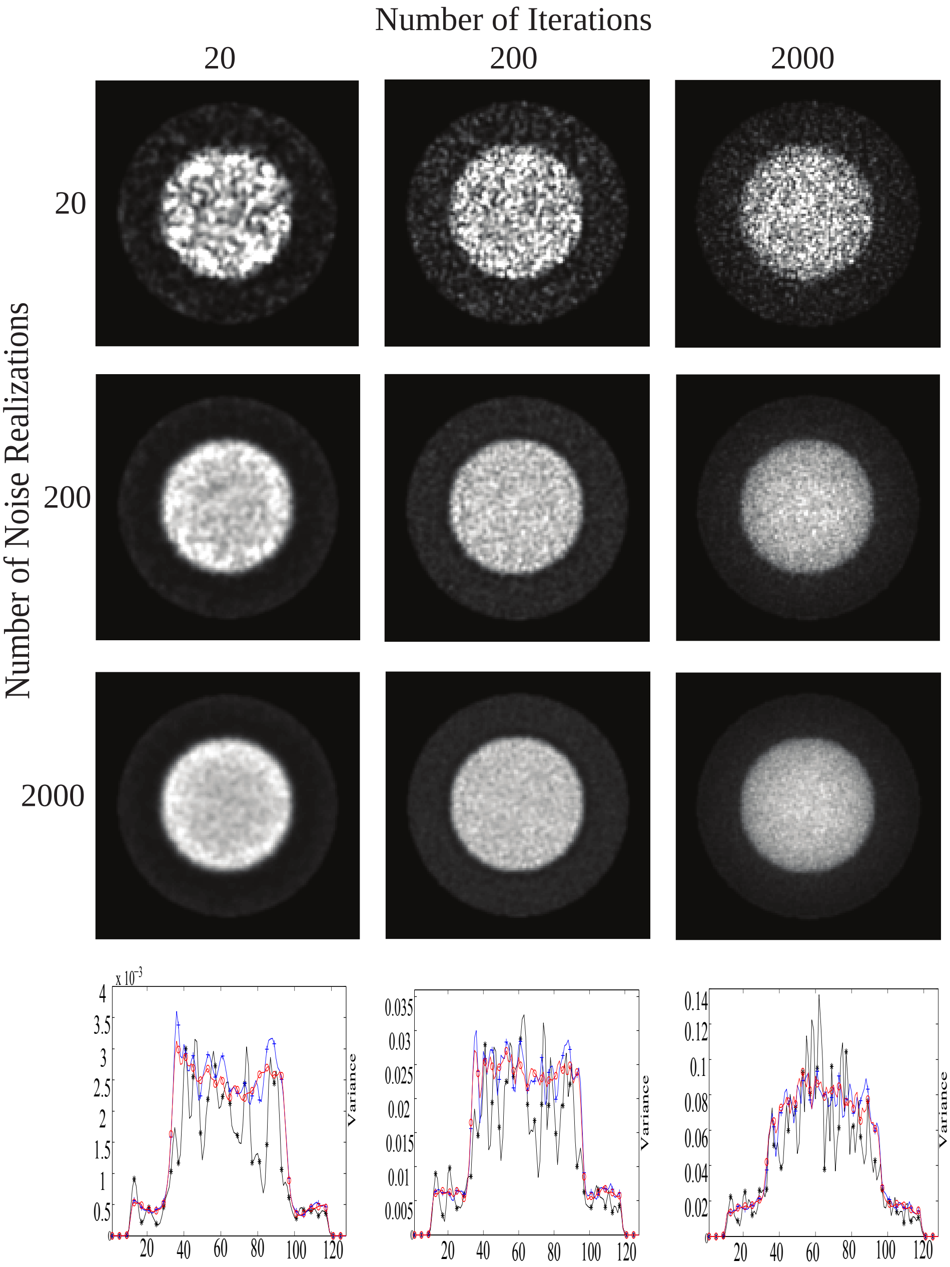

Figure 3.1: Top figure: A - Images of the variance for ML estimator as a function of the number of iterations and as a function of the number of noise instances. Bottom figure: Image profiles that show the differences in the variance images at increasing number of noise instances. B (left) - 20 iterations. C (centre) - 200 iterations. D (right) - 2000 iterations. Black (∗) - 20 noise instances. Blue (+) - 200 noise instances. Red (○) - 2000 noise instances.

3.2.1 Non-uniform Resolution in Ideal Emission Tomography

The resolution properties of a penalised-likelihood estimator are here studied for an idealised emission tomographic system.

The system is based on a detector that rotates 360° around the centre of the imaging volume. The imaging volume dimensions are $128 \times 128 \times 1$ cubic voxels of 2.46 mm. The detector efficiency terms c_d, the attenuation factors a_{di} and the spatial variation in sensitivity terms s_i are set to 1. The geometric response $\boldsymbol{B}'\boldsymbol{B}\mathrm{e}_i$ is assumed to be shift-invariant except for discretisation effects.

The phantom presents a hot circular region on the right, a cold circular region on the left, a background ellipse (see Figure 3.2 - A) . Additionally, a small hot spot is placed in the centre of each circular region. From this activity distribution noiseless projection data, $\overline{\boldsymbol{\nu}}$, has been simulated.

The noiseless projection data have been reconstructed using the Niftyrec algorithm for PL estimation [Pedemonte et al., 2010]. The algorithm has been iterated to nearly reach convergence in order to maximise the cost function in equation (3.2). A Maximum Likelihood estimate, where the hyperparameter is set to $\beta = 0$, is shown in Figure 3.2 - B . Since the system is shift-invariant, the ML estimate yields shift-invariant resolution properties and all the regions of the image are smoothed identically. A PL estimate where $\beta = 10^{-3}$ is shown in Figure 3.2 - C. Comparing the ML and PL reconstructions, a decreased contrast in the small hot spot (in the hot region) for the PL reconstruction can be noted. The non-uniform resolution properties are more evident looking at the image profiles in Figure 3.2 - D. The profiles of the two hot spots in the ML image are nearly identical, since both have the same contrast (height relative to local background) and nearly identical shape. However, the hot spots in the PL image have different contrast, since in the cold region the height relative to the local background is roughly 80% with respect to the true spot relative height, as opposed to 50% for the spot in the hot region. For this specific phantom, this is due to the resolution being much lower in the hot region than in the cold region. This experiment therefore demonstrates the well-known non-uniform resolution properties of a PL estimator even for an ideal imaging system [Fessler and Rogers, 1996]. Strictly speaking, however, the resolution properties in a voxel of interest do not depend only on the intensity of the background the voxels are immersed in; since, when the prior has a

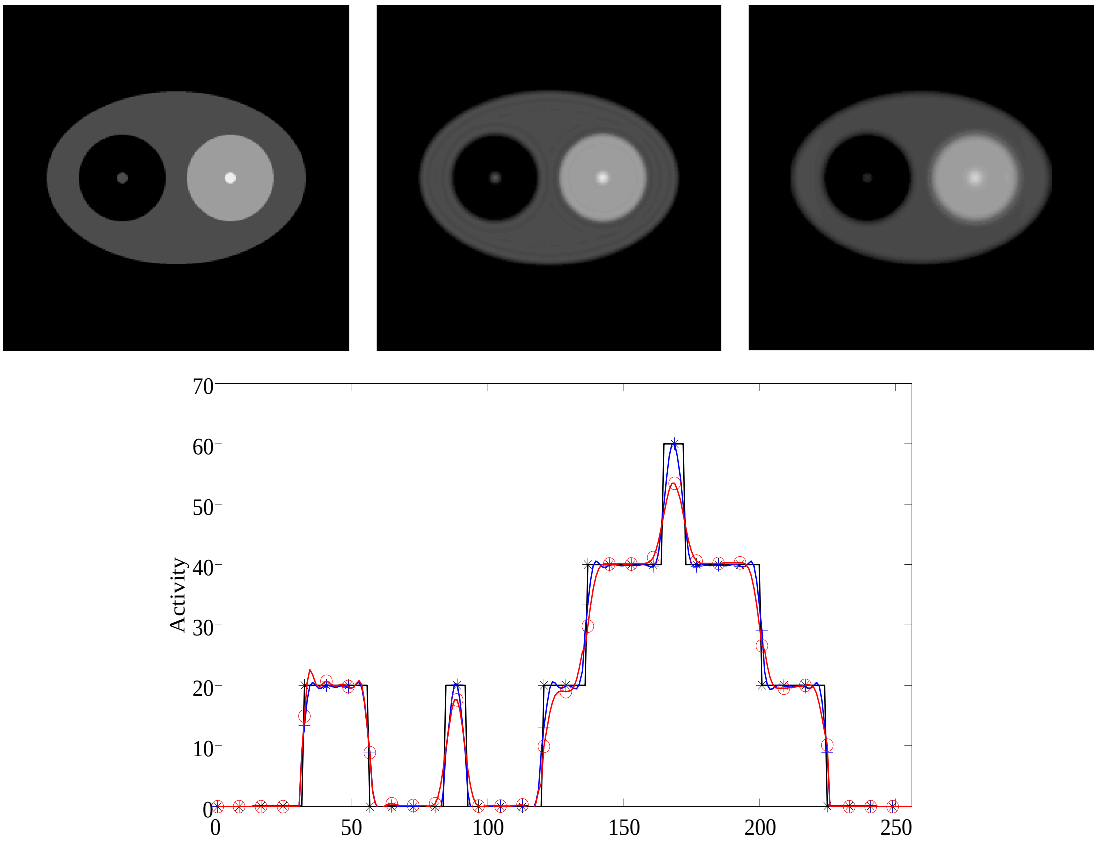

Figure 3.2: Non uniformities for PL estimator in ideal emission tomography reconstruction using a test phantom. Non uniform resolution properties arise even in the case of an idealised shift-invariant system model. A (top left) - True activity distribution. B (top centre) - ML estimate. C (top right) - PL estimate. D (bottom) - Image profiles: Black line ($*$) - true activity distribution. Blue line ($+$) - ML estimate. Red line ($\circ$) - PL estimate.

quadratic form (so that $\boldsymbol{R}$ is independent of $\boldsymbol{\lambda}$), the resolution properties depend on the object $\boldsymbol{\lambda}$ only through the projections $\boldsymbol{\nu}$. Due to the characteristics of tomographic acquisition, the estimated activity in a given voxel is affected by the counts acquired in all detector bins whose rays intersect that voxel. The uncertainty, and consequently the resolution, in the voxel of interest is therefore dependent on the uncertainty in all voxels that are located along the path of these lines.

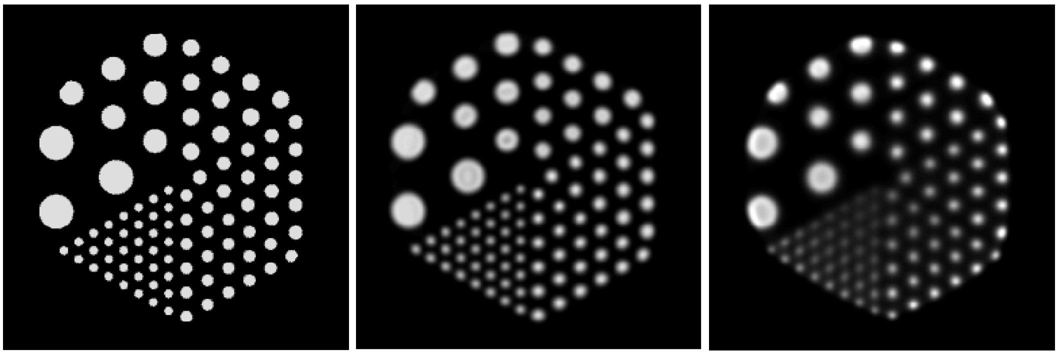

Figure 3.3: Non uniformities in SPECT reconstruction using a Jaszczak test phantom. A (left) - Activity distribution. B (centre) - ML estimate. C (right) - PL estimate.

3.2.2 Non-uniform Resolution in SPECT

In this section, the resolution properties of a Penalised-Likelihood estimator for a shift-variant SPECT system are studied by reconstructing a test phantom.

The SPECT system model is equivalent to the idealised model presented in the previous section except that the geometric response $\boldsymbol{B}'\boldsymbol{B}\mathrm{e}_i$ is depth-dependent and therefore shift variant. The SPECT camera is placed at a distance of 162 mm from the centre of the imaging volume and is equipped with a Low Energy High Resolution collimator (LEHR) . The intrinsic response is set to 3.6 mm, the collimator has a linearly varying depth-dependent Gaussian response that has a slope of 0.0533, which corresponds to 9.44 mm FWHM at the centre of the field of view. We simulated a 236 mm diameter Jaszczak phantom with rod diameters of 6.4, 9, 10.25, 12.8, 17.9, and 25.6 mm. The activity distribution for this phantom is shown in Figure 3.3 - A. Figure 3.3 - B and Figure 3.3 - C show reconstructions of noiseless data using an ML estimator and a PL estimator, respectively. For the PL estimator there is coarser resolution at the centre of the field of view than at the edges. Moreover Figure 3.3 - C shows that the nonuniform resolution properties are also anisotropic with increased radial blur. It should be noted, in fact, that the rods appear slightly elliptical. While these phantom reconstructions are helpful in demonstrating the possible anisotropy of the resolution properties of a PL estimator; it is difficult to identify the local properties quantitatively. Therefore, it is important to look at the local impulse response at various locations to quantify the local resolution properties.

3.3 The Linearised Local Impulse Response

In the previous section, the object-dependent resolution properties of an estimator have been demonstrated. Test phantoms for resolution quantification are often very different from the typical activity distributions that are being imaged in a real scan. Therefore, sample reconstructions of test phantoms are not necessarily reliable to investigate resolution properties of an estimator. Moreover because resolution properties are shift-variant for SPECT, it is difficult to fully investigate resolution properties with a single phantom.

It is possible to fully represent the resolution properties of a system, with a shift-invariant response, by means of a shift-invariant convolutional filter. This filter is referred to as the Impulse Response function, since it refers to the reaction of a shift-invariant system in response to an impulse function. This concept has been extended, for purposes of resolution investigation in Emission Tomography, by looking at reconstructions with impulses added to an object of interest [Fessler, 1996] [Stamos et al., 2010]. These responses therefore depend on the location of the impulse and are referred to as Local Impulse Response (LIR) .

The local impulse response for the i-th voxel is defined as:

$$LIR^{i}(\hat{\boldsymbol{\lambda}}) = \lim_{\delta \to 0} \frac{\mathbb{E}\left[\hat{\boldsymbol{\lambda}}(\boldsymbol{\nu}(\boldsymbol{\lambda} + \delta \mathrm{e}_i))\right] - \mathbb{E}\left[\hat{\boldsymbol{\lambda}}(\boldsymbol{\nu}(\boldsymbol{\lambda}))\right]}{\delta} \tag{3.9}$$

where e_i is the i-th unit vector. The Local Impulse Response characterises the estimator, system and object dependent resolution properties. The LIR measures the change in the mean reconstructed image, due to perturbation of a particular voxel in the noiseless object and it is local in two different senses. First, it is a function of the index i, reflecting the shift-variant system response. Second, it is dependent on the location in the N dimensional space through the assumed parameter $\boldsymbol{\lambda}$, reflecting the object dependency. It should be noted that $\mathbb{E}\left[\hat{\boldsymbol{\lambda}}\right] = \boldsymbol{\lambda}$ for unbiased estimators, in which case $LIR(\hat{\boldsymbol{\lambda}}) = \mathrm{e}_i$. PL estimators, however, are always biased, so the LIR will appear to have the shape of a bump-function as will be shown in the following sections.

The statistical calculation of the LIR implies the reconstruction until convergence of a set of noisy measurements for every pixel j under investigation. In the context of emission tomography, it has been observed by several investigators [Barrett et al., 1994] [Wilson et al., 1994] [Carson et al., 1994] that the mean over multiple noise realizations for a likelihood-based estimators is approximately equal to the value that one obtains reconstructing a noiseless data set:

$$\check{\boldsymbol{\lambda}} = \mathbb{E}\left[\hat{\boldsymbol{\lambda}}(\boldsymbol{\nu})\right] \approx \hat{\boldsymbol{\lambda}}(\boldsymbol{\nu}(\overline{\boldsymbol{\lambda}})) \tag{3.10}$$

This approximation is equivalent to make the assumption that the likelihood estimator is locally linear. Therefore substituting (3.10) in (3.9) it yields the definition of Linearised Local Impulse Response (LLIR) :

$$LLIR^{i}(\hat{\boldsymbol{\lambda}}) \approx \lim_{\delta \to 0} \frac{\hat{\boldsymbol{\lambda}}(\overline{\boldsymbol{\nu}}(\boldsymbol{\lambda} + \delta \mathrm{e}_i)) - \hat{\boldsymbol{\lambda}}(\overline{\boldsymbol{\nu}}(\boldsymbol{\lambda}))}{\delta} \tag{3.11}$$

which leads to a much less computational expensive calculation of the LIR.

3.3.1 Sample Linearised Local Impulse Response for SPECT

In this section, equation (3.11) is used to investigate the resolution properties of a PL estimator at different locations in the reconstructed image.

The shift-variant SPECT system model is equivalent to the one described in section 3.2.2. The digital phantom in Figure 3.4 - A is similar to the phantom presented in section 3.2.1. This $128 \times 128 \times 1$ phantom is composed of a warm background ellipse, a cold left disc, and a hot right disc, with relative emission intensities of 2, 1, and 3. From this activity distribution, noiseless projection data, $\overline{\boldsymbol{\nu}}$, have been simulated.

The sample locations are marked by the $\times$ symbols in figure 3.4 - A. Choosing a PL estimator with a Poisson noise model and a conventional uniform quadratic penalty, equation (3.11) has been used to estimate the Linearised Local Impulse Response at these locations. The regularisation parameter for this particular estimator was chosen to be $\beta = 10^{-3}$. The contours of these linear local impulse responses are presented in Figures 3.4 - B, C and D. Contours are formed at 25%, 50%, 75%, and 99% of the peak value. Moreover, the profiles of these linear local impulse responses are presented in Figure 3.4.

This investigation demonstrates the shift-variant resolution properties of a PL estimator for a conventional SPECT system. From the linear local impulse responses in Figure 3.4, it is possible to observe a broader response in the hot disc region with respect to the cold disc region.

For this specific experiment, we could deduce that a PL estimator introduce greater smoothing in high activity regions. However, when $\boldsymbol{R}$ has a quadratic form (so that $\boldsymbol{R}$ is independent of $\boldsymbol{\lambda}$), then remarkably the local impulse response given by (3.11) depends on the object $\boldsymbol{\lambda}$ only through its projections $\boldsymbol{\nu}$. Since a given voxel is primarily affected by the counts acquired in all detector bins whose rays intersect it, each voxel sees a different uncertainty and hence a different effective smoothing parameter. For voxels where data certainty is smaller, the PL estimate will give more weight to the prior, which (being a smoothness prior) will cause more smoothing. In emission tomography, voxels with higher activity yield rays with higher counts and hence more uncertainty. Thus, penalised-likelihood methods using the standard uniform penalty have lower spatial resolution in voxels that are intersected by rays with higher counts. Moreover, as one might expect, due to the depth-dependent detector response, the local impulse responses are broader near the centre of the field of view and narrower at the edges of the phantom. From Figure 3.4 the anisotropic resolution responses at different locations in the reconstructed image are also evident.

3.4 Figures Of Merit for Image Quality

For the calculation of the Image Quality, usually a trade-off between the resolution properties and the noise properties of the estimator arises. An approach widely used in emission imaging is to define a scalar measure for the bias and to consider the trade-off between bias and variance. Therefore bias and variance properties are often combined in one FOM.

3.4.1 Contrast Recovery Coefficient

Evaluating the local impulse response serves as an important tool for resolution investigation. Despite the fact that we can learn much by investigating the form of the local impulse response, a scalar measure for resolution quantification is needed. It is possible to reduce the function in equation (3.11) to a scalar measure by considering the local Contrast Recovery Coefficient (CRC) , which is defined as the i-th element of the local impulse response calculated for the i-th voxel:

$$CRC_i = LLIR_i^i(\hat{\boldsymbol{\lambda}}) \tag{3.12}$$

The CRC can be used as an alternative to the full-width at half maximum (FWHM) as a measure

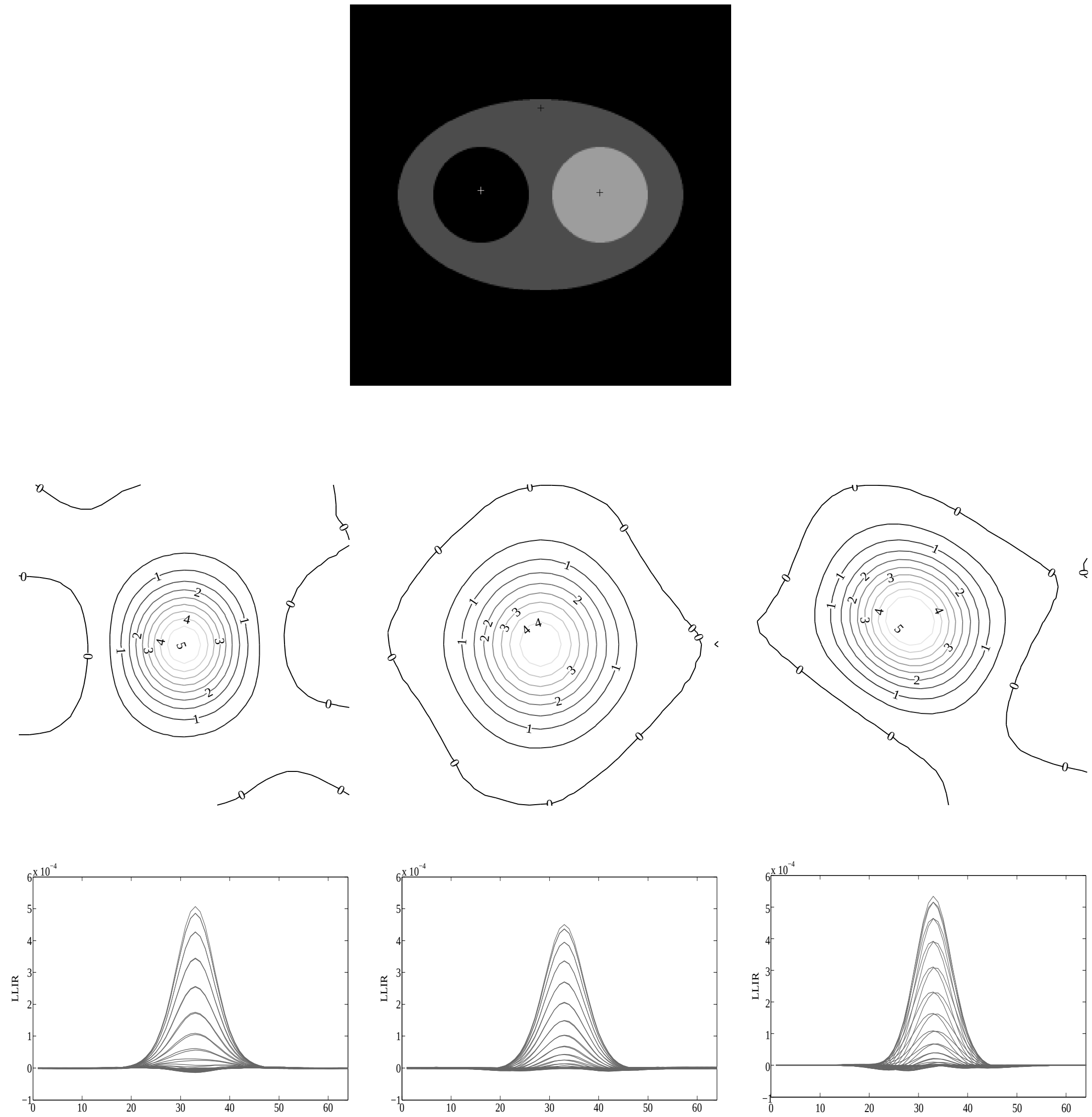

Figure 3.4: LLIRs contours and profiles, at different locations of the phantom, for a SPECT system and PL estimator. A (top) - Activity distribution and sample positions for the Linearised Local Impulse Response investigation. B (centre left) - LLIR contour for the left location in the cold disc. C (centre) - LLIR contour for the right location in the hot disc. D (centre right) - LLIR contour for the top location in the proximity of the phantom edge. E (bottom left) - LLIR profile for the left location in the cold disc. F (bottom centre) - LLIR profile for the right location in the hot disc. G (bottom right) - LLIR profile for the top location in the proximity of the phantom edge.

of resolution. There is a direct correspondence between the CRC-value and a certain resolution expressed in FWHM, if the Impulse Response is a Gaussian-shaped function, which is not always the case in emission tomographic imaging. To achieve some insight into the relationship between these two metrics, a study has been performed in [Qi and Leahy, 2000]. In this study, the FWHM and the CRC were compared for voxels at different radial distance from the scanning axis. Since the local impulse response is not symmetric, a mean FWHM in the transaxial plane was computed, using the mean FWHM area of the contour at half maximum. Even if this study indicates a monotonic relationship between the mean FWHM and the CRC for every voxel; the asymmetry of the LIR presupposes that any scalar measure of resolution at a point will be deficient in characterising the resolution response. However, for our purposes, the CRC has a distinct advantages over the FWHM in terms of the possibility to directly compute it from equation (3.11).

3.4.2 Contrast to Noise Ratio

In the rest of this thesis, we consider the Contrast to Noise Ratio (CNR) in voxel i as figure of merit for image quality calculation.

$$CNR_i = \frac{CRC_i}{\sqrt{Var_i}} \tag{3.13}$$

It accounts for the trade off between quantification accuracy and noise in the reconstructed image, combining these two properties as one FOM.

3.4.3 Signal to Noise Ratio

Another commonly used figure of merit is the Signal-to-Noise Ratio (SNR) which is

$$SNR^j = \frac{Signal}{\sqrt{Var}} \tag{3.14}$$

with Signal defined as the voxel value in the reconstructed image $\hat{\lambda}_i$.

3.5 Proposed Experimental Set-Up

For a systematic optimisation of a system, one would have to estimate the properties of the estimator (covariance and the LLIRs) not only considering a single phantom but a class of objects. The computational complexity of the statistical method, involving reconstruction to convergence of thousands of noise realisations, has therefore precluded a systematic optimisation of the system design for a wide range of system parameters.The problem is further complicated by the choice of the regularisation parameter β.

In this section, a criterion for the choice of β and description of purpose-made GPU accelerated reconstruction software are presented. The reconstruction software is capable of processing multiple reconstructions in parallel, enabling the estimation of the reference variance in a reasonably short time (see table 4.1).

A series of independent noise realisations was computed using a pseudo-random Poisson noise generator (of the IRT toolbox [Fessler, 2012]), based on the rejection sampling algorithm described at page 293 of [Vetterling et al., 2002]. The noisy data sets were reconstructed using an accelerated GPU implementation of the One Step Late algorithm for PL estimation, implemented as part of the NiftyRec toolbox [Pedemonte et al., 2010]. 10000 iterations were performed. A smoothing prior with a small weight $\beta = 10^{-12}$ was included in the cost function. The value of the regularisation parameter was chosen after trial and error, as a minimum value that nearly guarantees convergence within 10000 iterations. The calculation of the variance is based on 10240 noise realisations. The number of noise instances is a multiple of 1024 (10 times) as NiftyRec can process concurrently up to 1024 reconstructions in order to make efficient use of the GPU. 10 repetitions were chosen in order to obtain satisfactory images of variance. Though often variance is calculated with much smaller sample size and number of iterations, we found that such large numbers are necessary to obtain a good estimate of the variance.

Deterministic approximations for the mean and covariance of the PL estimator are presented in the next chapter. Moreover, in chapter 5, 6 and 7, the results obtained with the reference statistical method based on the reconstruction of multiple noise realisations will be compared with the results obtained from these deterministic approximations based on the Fisher Information method. It will be shown that such large numbers of noise realisations are necessary in order to meaningfully compare the results obtained with the statistical method with the results obtained with the deterministic FIM-based method.

Chapter 4

Deterministic Evaluation of Image Quality

4.1 Introduction

In this chapter the deterministic approach for the estimation of the uncertainty in the reconstruction is summarised. Closed form analytical expressions of the statistical properties (such as mean and variance) of the PL estimator, as defined in chapter 2, are unavailable. The absence of an explicit analytical expression makes it difficult to study the noise properties of the estimator $\hat{\boldsymbol{\lambda}}$. In contrast, one can easily analyse the statistical properties of linear reconstruction methods such as the FBP method [Chesler et al., 1977]. Thus, in chapter 3, it has been shown how it is possible to rely on numerical simulations to examine statistical properties of nonlinear estimators. Although empirical studies are important, deterministic expressions, even if approximate, can be convenient for designing imaging systems, and for comparing estimators. A SPECT system may be evaluated based on the fundamental bias/variance trade-off of the estimator and one would like to be able to easily study the estimator characteristics over a range of system parameters. In such cases, numerical simulations can be prohibitively expensive and therefore it is useful to have approximate expressions of the mean and covariance of the PL estimator.

These approximations have been derived by Fessler and Rogers [1996] using the implicit function theorem, the Taylor expansion and the chain rule of differentiation. The expressions are defined solely in terms of the partial derivatives of the cost function used for estimation. Since partial derivatives are used in the derivation, the approximations presented in this chapter are restricted to problems where $\boldsymbol{\lambda}$ is a continuous parameter vector. Therefore, strictly speaking, the approach is not applicable to problems where inequality constraints are imposed. For unregularised estimation methods, it is often necessary to impose a non-negativity constraint on $\hat{\lambda}$. However, for cost functions that include an appropriate penalty function for regularisation, non-negativity constraints are active relatively infrequently. Negative values can be avoided by keeping the weight of the prior beta low enough [Bruyant, 2002]. An approach that tackles this problem has been proposed by [Li et al., 2004], where they introduce a generalisation based on truncated Gaussian distributions.

The approximations presented in this chapter are derived using the assumption that the estimate is computed by completely maximising the cost function. Approximations for mean and covariance of the estimator will be defined at point of convergence.

The use of these deterministic approximations of the mean and the covariance of the estimator has been explored for the purpose of measuring image quality in [Fessler and Rogers, 1996] [Qi and Leahy, 2000] [Stayman and Fessler, 2000] and for the purpose of system design optimisation in [Nuyts, 2009] [Zhou et al., 2010] [Vunckx et al., 2008a]. SPECT systems may, in fact, be evaluated based on the fundamental trade-off between bias and variance that can be achieved in reconstructed images and such trade-offs may be derived analytically using the Cramer-Rao type bounds [Hero et al., 1996] [Meng and Clinthorne, 2004] [Meng and Li, 2000]. All these deterministic approximations imply the calculation and the inversion of the Fisher Information Matrix (FIM) and they are referred to as Fisher Information-based methods in the remainder of this manuscript.

Unfortunately, computing, storing and inverting the FIM is not feasible for the typical matrix size of 3D imaging systems. In order to tackle the problem of the computational load in inverting the FIM, an approximation has been previously proposed. Qi et al. [Qi and Leahy, 2000] argued that if we are only interested in calculating the properties of an estimator in a single voxel i, it is acceptable to ignore the non-stationarity of the FIM. The computations are done for voxel i, and therefore only the i-th row of the FIM needs to be calculated. This local approximation of the FIM, explained in more detail in section 4.5.1 of this chapter, is obtained by replacing all rows of the FIM with the shifted version of its $i - th$ row. Consequently the FIM simply reduces to a circulant matrix and this approximation is referred to as the *circulant approximation*.

However, due to the characteristics of tomographic acquisition, the estimated activity in a given voxel is affected by the counts acquired in all detector bins whose rays intersect that voxel. The uncertainty (or variance) in the voxel of interest is dependent on the uncertainty in all voxels that are located along the path of these lines. The uncertainties of all those other voxels are also dependent on the uncertainty in all voxels that are located along the path of the lines that intersect them. Therefore, the influence of the distribution of activity in distant voxels on the uncertainty in a voxel of interest depends, comprehensively, on the tomographic reconstruction approach. However, since in a SPECT system with parallel hole (PH) collimators, the counts in the detector bins are the expressions of the integral of the emitted photons originating from a conical volume (and not simply from a line), the interdependence between the voxels becomes even more complex. The full FIM accounts for such complex interdependence between all the voxels in the imaging volume; whereas the aforementioned circulant approximation makes use of a single row of the FIM and does not capture such interaction (see section 4.5.4).

In this chapter, a novel algorithm for efficient estimation of the uncertainty in the reconstruction, based on the Fisher Information, is introduced. This new formulation relies on subsampling the FIM. The FIM is calculated over a subset of voxels arranged in a grid that covers the whole volume. Every element of the FIM at the grid points is calculated exactly, accounting for the acquisition geometry and for the object, without further approximation. This new formulation, presented in section 4.5.2, reduces the computational complexity in inverting the FIM but nevertheless accounts for the global interdependence between the variables.

The main aim of this thesis is to emphasise the benefits of this new approximation of the FIM with respect to the aforementioned circular approximation for the optimisation of SPECT systems with highly shift-variant response (as a result of distance dependent resolution and adaptive data sampling).

4.2 Mean and Covariance of Implicit Estimators

In chapter 2 section 2.5.1 an objective function $\Psi(\boldsymbol{\lambda})$ has been introduced:

$$\Psi(\boldsymbol{\lambda}, \boldsymbol{\nu}) = L(\boldsymbol{\nu}, \boldsymbol{H}\boldsymbol{\lambda}) - \beta R(\boldsymbol{\lambda}) \tag{4.1}$$

that depends on unknown parameters $\boldsymbol{\lambda}$ and noisy measurements $\boldsymbol{\nu}$ and where $L(\cdot, \cdot)$ and $R(\cdot)$ are referred to as the likelihood function and the penalty function respectively. An estimator

$\hat{\boldsymbol{\lambda}} = \hat{\boldsymbol{\lambda}}(\boldsymbol{\nu})$ is here defined as the unconstrained maximiser of this cost function:

$$\hat{\boldsymbol{\lambda}} = \arg\max \Psi(\boldsymbol{\lambda}, \boldsymbol{\nu}) \tag{4.2}$$

In the following, $\Psi(\boldsymbol{\lambda})$ is assumed to be a suitably regular function, so that the partial derivatives used below exist. $\Psi(\cdot, \boldsymbol{\nu})$ is also assumed to have a unique maximum and therefore a unique global maximiser $\hat{\boldsymbol{\lambda}} = \hat{\boldsymbol{\lambda}}(\boldsymbol{\nu})$.

4.2.1 Mean Approximation

In this section an approximation for the calculation of the mean of the estimator $\hat{\boldsymbol{\lambda}}$ is described. In equation 3.3, we defined the mean over multiple noise realisations for a likelihood-based estimator. In general, calculation of the mean, $\mathbb{E}[\hat{\boldsymbol{\lambda}}]$, can be extremely computationally expensive. However, as already described in section 3.3, assuming that the cost function Ψ is locally linear in $\boldsymbol{\nu}$, the expectation of the estimate is approximately equal to the value produced by applying the estimator to noiseless data:

$$\check{\boldsymbol{\lambda}} = \mathbb{E}[\hat{\boldsymbol{\lambda}}(\boldsymbol{\nu})] \approx \hat{\boldsymbol{\lambda}}(\overline{\boldsymbol{\nu}}) \tag{4.3}$$

This approach requires modest computation, and it has been used extensively by investigators in emission tomography [Barrett et al., 1994] [Wilson et al., 1994] [Carson et al., 1994].

4.2.2 Covariance Approximation

In this section, the approximations for the deterministic calculation of the covariance of the estimator $\hat{\boldsymbol{\lambda}}$ are described. In the following, the covariance matrix will be derived as the inverse of the Hessian of the posterior distribution calculated at the PL estimate; where the Hessian is the sum of two matrices: the Fisher Information Matrix and the Hessian of the penalty function. These approximations follow the derivations of Fessler [1996].

Covariance Approximation 1

A necessary condition for $\hat{\boldsymbol{\lambda}}$ to be the maximiser of (4.1) is that the following equation must be satisfied for any i:

$$\frac{\partial}{\partial \lambda_i} \Psi(\boldsymbol{\lambda}, \boldsymbol{\nu})|_{\boldsymbol{\lambda}=\hat{\boldsymbol{\lambda}}(\boldsymbol{\nu})} = 0, \quad i = 1, \ldots, N. \tag{4.4}$$

This equation can be written concisely in vector form as:

$$\boldsymbol{\Sigma}(\hat{\boldsymbol{\lambda}},\boldsymbol{\nu}) = \nabla^{[1,0]}\Psi(\hat{\boldsymbol{\lambda}},\boldsymbol{\nu}) = 0 \tag{4.5}$$

where $\nabla^{[1,0]} = \left[\frac{\partial}{\partial\lambda_1},\ldots,\frac{\partial}{\partial\lambda_N}\right]$ is the row gradient operator, which returns a vector of partial derivatives with respect to the first argument of Ψ.

Equation (4.5) requires Ψ to be a suitably regular function, and therefore restricts the approximation to continuous parameters. This is the reason why this covariance approximation can not account for non-negativity constraints.

Performing the first-order Taylor series expansion of $\boldsymbol{\Sigma}$, the following equation is obtained:

$$\boldsymbol{\Sigma}(\hat{\boldsymbol{\lambda}}) \approx \boldsymbol{\Sigma}(\check{\boldsymbol{\lambda}},\boldsymbol{\nu}) + \nabla^{[1,0]}\boldsymbol{\Sigma}(\check{\boldsymbol{\lambda}},\boldsymbol{\nu})(\hat{\boldsymbol{\lambda}} - \check{\boldsymbol{\lambda}}) \tag{4.6}$$

where the linearisation point is chosen to be the approximated mean of the estimator, $\check{\boldsymbol{\lambda}}$, as in equation (4.3). The $N \times N$ matrix $\boldsymbol{\Gamma} = \nabla^{[1,0]}\boldsymbol{\Sigma}$ is the Hessian of Ψ

$$\boldsymbol{\Gamma} = \nabla^{[1,0]}\boldsymbol{\Sigma}(\check{\boldsymbol{\lambda}},\mathbb{E}[\boldsymbol{\nu}]) = \nabla^{[2,0]}\Psi(\check{\boldsymbol{\lambda}},\overline{\boldsymbol{\nu}}) \tag{4.7}$$

where the $\nabla^{[2,0]}$ operator yields a matrix whose (i,j) -th element is $\frac{\partial^2}{\partial\lambda_i\partial\lambda_j}$.

It is then possible to equate the approximation (4.6) to zero, using (4.5), yielding:

$$\boldsymbol{\Sigma}(\check{\boldsymbol{\lambda}},\boldsymbol{\nu}) \approx -\boldsymbol{\Gamma}(\hat{\boldsymbol{\lambda}} - \check{\boldsymbol{\lambda}}) = -\nabla^{[2,0]}\Psi(\check{\boldsymbol{\lambda}},\overline{\boldsymbol{\nu}})(\hat{\boldsymbol{\lambda}} - \check{\boldsymbol{\lambda}}) \tag{4.8}$$

Assuming that the Hessian $\boldsymbol{\Gamma}$ is invertible (i.e. if the cost function Ψ is strictly concave) and rearranging equation (4.8), the following linearised approximation for the estimator is obtained:

$$\begin{aligned}\hat{\boldsymbol{\lambda}} &\approx \check{\boldsymbol{\lambda}} - \left[\nabla^{[2,0]}\Psi(\check{\boldsymbol{\lambda}},\overline{\boldsymbol{\nu}})\right]^{-1}\boldsymbol{\Sigma}(\check{\boldsymbol{\lambda}},\boldsymbol{\nu})\\ &\approx \check{\boldsymbol{\lambda}} - \boldsymbol{\Gamma}^{-1}\boldsymbol{\Sigma}(\check{\boldsymbol{\lambda}},\boldsymbol{\nu})\end{aligned} \tag{4.9}$$

Taking the covariance of both sides of (4.9) yields the following covariance approximation

$$\begin{aligned}\mathrm{Cov}(\hat{\boldsymbol{\lambda}}) &\approx \left[\nabla^{[2,0]}\Psi(\check{\boldsymbol{\lambda}},\overline{\boldsymbol{\nu}})\right]^{-1}\mathrm{Cov}\{\boldsymbol{\Sigma}(\check{\boldsymbol{\lambda}},\boldsymbol{\nu})\}\left[\nabla^{[2,0]}\Psi(\check{\boldsymbol{\lambda}},\overline{\boldsymbol{\nu}})\right]^{-1}\\ &\approx \boldsymbol{\Gamma}^{-1}\mathrm{Cov}\{\boldsymbol{\Sigma}(\check{\boldsymbol{\lambda}},\boldsymbol{\nu})\}\boldsymbol{\Gamma}^{-1}\end{aligned} \tag{4.10}$$

Covariance Approximation 2

As a further approximation, it is possible to linearise $\boldsymbol{\Sigma}$ around the mean of the measurements data $\overline{\boldsymbol{\nu}}$:

$$\boldsymbol{\Sigma}(\check{\boldsymbol{\lambda}}, \boldsymbol{\nu}) \approx \boldsymbol{\Sigma}(\check{\boldsymbol{\lambda}}, \overline{\boldsymbol{\nu}}) + \nabla^{[0,1]}\boldsymbol{\Sigma}(\check{\boldsymbol{\lambda}}, \overline{\boldsymbol{\nu}})(\boldsymbol{\nu} - \overline{\boldsymbol{\nu}}) \tag{4.11}$$

where

$$\nabla^{[0,1]}\boldsymbol{\Sigma}(\boldsymbol{\lambda}, \boldsymbol{\nu}) = \nabla^{[1,1]}\Psi(\boldsymbol{\lambda}, \boldsymbol{\nu}) \tag{4.12}$$

and the $\nabla^{[1,1]}$ operator yield a matrix whose the (i, d) -th element is $\frac{\partial^2}{\partial \lambda_i \partial \nu_d}$.
This linearisation leads to the second approximation:

$$\mathrm{Cov}\{\boldsymbol{\Sigma}(\check{\boldsymbol{\lambda}}, \boldsymbol{\nu})\} \approx \left[\nabla^{[1,1]}\Psi(\check{\boldsymbol{\lambda}}, \overline{\boldsymbol{\nu}})\right] \mathrm{Cov}\{\boldsymbol{\nu}\} \left[\nabla^{[1,1]}\Psi(\check{\boldsymbol{\lambda}}, \overline{\boldsymbol{\nu}})\right]' \tag{4.13}$$

which, substituted into (4.10), yields the following covariance approximation:

$$\begin{aligned}\mathrm{Cov}\{\hat{\boldsymbol{\lambda}}\} \approx \left[\nabla^{[2,0]}\Psi(\check{\boldsymbol{\lambda}}, \overline{\boldsymbol{\nu}})\right]^{-1} \left[\nabla^{[1,1]}\Psi(\check{\boldsymbol{\lambda}}, \overline{\boldsymbol{\nu}})\right] \mathrm{Cov}\{\boldsymbol{\nu}\}\cdot \\ \cdot \left[\nabla^{[1,1]}\Psi(\check{\boldsymbol{\lambda}}, \overline{\boldsymbol{\nu}})\right]' \left[\nabla^{[2,0]}\Psi(\check{\boldsymbol{\lambda}}, \overline{\boldsymbol{\nu}})\right]^{-1}\end{aligned} \tag{4.14}$$

To summarize, (4.10) and (4.14) are expressions for the estimator covariance that depend only on the partial derivatives of the cost function Ψ, and do not require an expression for the implicit function $\hat{\boldsymbol{\lambda}}(\boldsymbol{\nu})$. These approximations do depend on $\check{\boldsymbol{\lambda}}$, which one usually computes using (4.3) by applying the reconstruction algorithm to the noise free data $\overline{\boldsymbol{\nu}}$.

Covariance Approximation for Penalised Likelihood Estimators

In this section, the covariance approximation is derived for the penalised likelihood estimators of the form expressed in equation (4.1). Substituting (4.1) into (4.14), we obtain:

$$\begin{aligned}\mathrm{Cov}\{\hat{\boldsymbol{\lambda}}\} \approx \left[-\nabla^{[2,0]}L(\check{\boldsymbol{\lambda}}, \overline{\boldsymbol{\nu}}) + \nabla^{[2,0]}R(\check{\boldsymbol{\lambda}})\right]^{-1} \left[\nabla^{[1,1]}L(\check{\boldsymbol{\lambda}}, \overline{\boldsymbol{\nu}})\right] \mathrm{Cov}\{\boldsymbol{\nu}\}\cdot \\ \cdot \left[\nabla^{[1,1]}L(\check{\boldsymbol{\lambda}}, \overline{\boldsymbol{\nu}})\right]' \left[-\nabla^{[2,0]}L(\check{\boldsymbol{\lambda}}, \overline{\boldsymbol{\nu}}) + \nabla^{[2,0]}R(\check{\boldsymbol{\lambda}})\right]^{-1}\end{aligned} \tag{4.15}$$

For the Poisson model in (2.6) the partial derivatives of the log-likelihood are:

$$\begin{aligned}\frac{\partial L}{\partial \lambda_i}(\boldsymbol{\lambda},\boldsymbol{\nu}) &= \sum_{d}^{M} h_{di}\left(\frac{\nu_d}{\sum_k^N h_{dk}\lambda_k}\right)\\ \frac{\partial^2 L}{\partial \lambda_i \partial \lambda_j}(\boldsymbol{\lambda},\boldsymbol{\nu}) &= -\sum_{d}^{M} \frac{h_{di}h_{dj}\nu_d}{(\sum_k^N h_{dk}\lambda_k)^2}\\ \frac{\partial^2 L}{\partial \lambda_i \partial \nu_d}(\boldsymbol{\lambda},\boldsymbol{\nu}) &= \frac{h_{di}}{\sum_k^N h_{dk}\lambda_k}\end{aligned} \tag{4.16}$$

Making use of a uniform quadratic penalty function $R(\boldsymbol{\lambda})$ of the form presented in (2.12), and noting that this smoothing penalty may be written in matrix form as:

$$R(\boldsymbol{\lambda}) = \frac{1}{2}\boldsymbol{\lambda}'\boldsymbol{R}\boldsymbol{\lambda}, \tag{4.17}$$

the $N \times N$ matrix, $\boldsymbol{R}$, is the Hessian of the quadratic penalty and has elements defined by:

$$\boldsymbol{R}_{jk} = \begin{cases} \sum_{l=1}^{p} \frac{1}{2}(w_{lj} + w_{jl}) & k = j \\ -w_{jk} & k \neq j \end{cases} \tag{4.18}$$

Therefore the partial derivatives can be expressed in matrix form as:

$$\begin{aligned}\nabla^{[2,0]}\Psi(\check{\boldsymbol{\lambda}},\overline{\boldsymbol{\nu}}) &= \boldsymbol{H}'D\left[\frac{\overline{\nu}_d(\boldsymbol{\lambda})}{\overline{\nu}_d^2(\check{\boldsymbol{\lambda}})}\right]\boldsymbol{H} - \beta\boldsymbol{R}\\ \nabla^{[1,1]}\Psi(\check{\boldsymbol{\lambda}},\overline{\boldsymbol{\nu}}) &= -\boldsymbol{H}'D\left[\frac{1}{\overline{\nu}_d(\check{\boldsymbol{\lambda}})}\right]\end{aligned} \tag{4.19}$$

It should be noted that the measurements data have independent Poisson distributions and therefore that $\mathrm{Cov}\{\boldsymbol{\nu}\} = D\left[\overline{\nu}_d(\boldsymbol{\lambda})\right]$. Substituting (4.19) into (4.14) we obtain the following approximation to the covariance of the estimator:

$$\begin{aligned}\mathrm{Cov}(\hat{\boldsymbol{\lambda}}) = &\left[\boldsymbol{H}'D\left[\frac{\overline{\nu}_d(\boldsymbol{\lambda})}{\overline{\nu}_d^2(\check{\boldsymbol{\lambda}})}\right]\boldsymbol{H} - \beta\boldsymbol{R}\right]^{-1}\boldsymbol{H}'D\left[\frac{1}{\overline{\nu}_d(\check{\boldsymbol{\lambda}})}\right]D\left[\overline{\nu}_d(\boldsymbol{\lambda})\right]\cdot\\ &\cdot D\left[\frac{1}{\overline{\nu}_d(\check{\boldsymbol{\lambda}})}\right]\boldsymbol{H}\left[\boldsymbol{H}'D\left[\frac{\overline{\nu}_d(\boldsymbol{\lambda})}{\overline{\nu}_d^2(\check{\boldsymbol{\lambda}})}\right]\boldsymbol{H} - \beta\boldsymbol{R}\right]^{-1}\end{aligned} \tag{4.20}$$

and then, simplifying, the following expression is obtained:

$$\mathrm{Cov}\{\hat{\boldsymbol{\lambda}}\} \approx \left[\boldsymbol{F} - \beta\boldsymbol{R}\right]^{-1}\boldsymbol{F}\left[\boldsymbol{F} - \beta\boldsymbol{R}\right]^{-1} \tag{4.21}$$

where $\boldsymbol{F}$ is an approximation of the Fisher Information Matrix defined as:

$$\boldsymbol{F}(\boldsymbol{\lambda}) = \boldsymbol{H}' D \left[\frac{\overline{\nu}_d(\boldsymbol{\lambda})}{\overline{\nu}_d(\check{\boldsymbol{\lambda}})^2} \right] \boldsymbol{H} \tag{4.22}$$

whereas the Fisher Information Matrix is normally defined as the minus of the matrix of expectations of the second partial derivatives of the log-likelihood (as described in section 4.4).

4.3 Linear Local Impulse Response

Recalling from section 3.3, the local impulse response describes how the estimate $\hat{\boldsymbol{\lambda}}$ would change due to a point-like perturbation of the true object at a given spatial location.

As already stated in section 4.2.1, the mean over multiple noise realisations, for a penalised likelihood estimator, is approximately equal to the value that one obtains reconstructing a noiseless data set. This approximation is equivalent to making the assumption that the likelihood estimator is locally linear. Thus, as already stated in section 3.3, this approximation results in the definition of Linearised Local Impulse Response:

$$\begin{aligned} LLIR^i(\hat{\boldsymbol{\lambda}}) &\approx \lim_{\delta \to 0} \frac{\hat{\boldsymbol{\lambda}}(\overline{\boldsymbol{\nu}}(\boldsymbol{\lambda} + \delta \mathrm{e}_i)) - \hat{\boldsymbol{\lambda}}(\overline{\boldsymbol{\nu}}(\boldsymbol{\lambda}))}{\delta} \\ &= \nabla \hat{\boldsymbol{\lambda}}(\overline{\boldsymbol{\nu}}(\boldsymbol{\lambda})) \nabla \overline{\boldsymbol{\nu}}(\boldsymbol{\lambda}) \mathrm{e}_j \end{aligned} \tag{4.23}$$

where $\nabla \hat{\boldsymbol{\lambda}} = \frac{\partial}{\partial \lambda_i} \hat{\boldsymbol{\lambda}}$ is a $N \times M$ matrix and $\nabla \overline{\boldsymbol{\nu}} = \frac{\partial}{\partial \lambda_i} \overline{\boldsymbol{\nu}}$ and is a $M \times N$ matrix.
This formulation leads to a much less computational expensive calculation of the LIR. However, even the calculation of the $LLIR$ is computationally expensive since it implies a noiseless reconstruction for each voxel i. Noting, from (4.23), that the LLIR is equivalent to the gradient of $\hat{\boldsymbol{\lambda}}$; to further reduce the computational burden, an approximated deterministic formulation of the LLIR is introduced in the following.

The definition of local impulse response given in (4.23) leads to expressions that depend on the gradient of the estimator $\hat{\boldsymbol{\lambda}}$. We apply the chain rule to differentiate (4.5) with respect to $\boldsymbol{\nu}$ as follows:

$$\nabla^{[2,0]} \Psi(\hat{\boldsymbol{\lambda}}(\boldsymbol{\nu}), \boldsymbol{\nu}) \nabla \hat{\boldsymbol{\lambda}}(\boldsymbol{\nu}) + \nabla^{[1,1]} \Psi(\hat{\boldsymbol{\lambda}}(\boldsymbol{\nu}), \boldsymbol{\nu}) = 0 \tag{4.24}$$

Solving the above equation yields the following general expression for the estimator gradient:

$$\nabla\hat{\boldsymbol{\lambda}}(\boldsymbol{\nu}) = \left[\nabla^{[2,0]}\Psi(\hat{\boldsymbol{\lambda}}(\boldsymbol{\nu}),\boldsymbol{\nu})\right]^{-1}\left[-\nabla^{[1,1]}\Psi(\hat{\boldsymbol{\lambda}}(\boldsymbol{\nu}),\boldsymbol{\nu})\right] \tag{4.25}$$

Combining (4.25) with (4.23) yields the following expression for the Linearised Local Impulse Response for penalised likelihood estimators:

$$LLIR^i(\hat{\boldsymbol{\lambda}}) = \left[\nabla^{[2,0]}L(\check{\boldsymbol{\lambda}},\overline{\boldsymbol{\nu}}) + \nabla^{[2,0]}R(\check{\boldsymbol{\lambda}})\right]^{-1}\left[-\nabla^{[1,1]}L(\check{\boldsymbol{\lambda}},\overline{\boldsymbol{\nu}})\right]\nabla\overline{\boldsymbol{\nu}}(\boldsymbol{\lambda})\mathrm{e}_i \tag{4.26}$$

Substituting (4.19) into (4.26) and noting that $\nabla\overline{\boldsymbol{\nu}}(\boldsymbol{\lambda})\mathrm{e}_j = \boldsymbol{H}\mathrm{e}_j$, we obtain the following expression:

$$\begin{aligned} LLIR^i(\hat{\boldsymbol{\lambda}}) &= \left[\boldsymbol{H}'D\left[\frac{\overline{\nu}_d(\boldsymbol{\lambda})}{\overline{\nu}_d^2(\check{\boldsymbol{\lambda}})}\right]\boldsymbol{H} - \beta\boldsymbol{R}(\check{\boldsymbol{\lambda}})\right]^{-1}\boldsymbol{H}'D\left[\frac{1}{\overline{\nu}_d(\check{\boldsymbol{\lambda}})}\right]\boldsymbol{H}\mathrm{e}_i \\ &= \left[\boldsymbol{F} - \beta\boldsymbol{R}\right]^{-1}\boldsymbol{F}\mathrm{e}_i \end{aligned} \tag{4.27}$$

where

$$\boldsymbol{F} = \boldsymbol{H}'D\left[\frac{1}{\overline{\boldsymbol{\nu}}(\check{\boldsymbol{\lambda}})}\right]\boldsymbol{H} \tag{4.28}$$

It should be noted that, in the above equations, we assume $\overline{\nu}_d(\boldsymbol{\lambda})$ to equal $\overline{\nu}_d(\check{\boldsymbol{\lambda}})$ and that the above equation therefore differs from (4.22). However, when the system model $\boldsymbol{H}$ closely approximates the actual system, and when the regularisation term approaches zeros, usually $\check{\boldsymbol{\lambda}}$ approaches $\boldsymbol{\lambda}$ and this approximation becomes reasonable.

The local impulse response therefore depends on the regularisation term through its Hessian, $\boldsymbol{R}$. In equation (4.27) it is clear that $LLIR^i(\hat{\boldsymbol{\lambda}})$ approaches e_i as $\boldsymbol{R}$ approaches zero, for well-conditioned problems.

4.4 The Fisher Information Matrix

In the previous section, the Fisher Information Matrix (FIM) of the likelihood has been introduced in the definition of the deterministic approximations for covariance and the LLIR of penalised likelihood estimators. Generally speaking, this matrix plays a key role in the analysis of both spatial resolution and noise properties of image reconstruction methods based on the Poisson model.

The FIM, $\boldsymbol{F}(\boldsymbol{\lambda})$, is a $N \times N$ symmetric matrix whose (i,j)-th element are defined by the expected values of the second partial derivatives of the log-likelihood:

$$\begin{aligned} F_{ij}(\boldsymbol{\lambda}) &= -\mathbb{E}\left[\frac{\partial^2}{\partial\lambda_j\partial\lambda_i}\ln p(\boldsymbol{\nu}|\boldsymbol{\lambda})\right] \\ &= -\mathbb{E}\left[\frac{\partial^2}{\partial\lambda_j\partial\lambda_i}L(\boldsymbol{\nu},\boldsymbol{\lambda})\right] \end{aligned} \tag{4.29}$$

where $p(\boldsymbol{\nu}|\boldsymbol{\lambda})$ is the conditional probability distribution associated with photon counting, that, in Emission Tomography, is Poisson distributed (see equation 2.6). It can be shown that $\boldsymbol{F}$ is positive semi-definite (positive definite if non-singular).

Considering an unbiased estimator as a maximiser of the log-likelihood function; a necessary condition for the solution of such estimator is to set the first derivatives of the log-likelihood to zero (see equation (4.5)). The second derivatives of the log-likelihood give us instead information about the quality of the estimator and the FIM can be considered as a measure of the sharpness of the log-likelihood function, around the ML estimate $\hat{\boldsymbol{\lambda}}$. A sharp curve around the value of the estimate $\hat{\boldsymbol{\lambda}}$ corresponds to a well-posed problem, whereas a flat curve around $\hat{\boldsymbol{\lambda}}$ corresponds to an ill-posed problem.

The second partial derivatives of the log likelihood have been defined in equation (4.19). Therefore for the Poisson model, the (i,j)-th element of the Fisher information matrix is calculated as:

$$\begin{aligned} \boldsymbol{F}(\boldsymbol{\lambda}) = -\mathbb{E}\left[\frac{\partial^2 L}{\partial\lambda_i\partial\lambda_j}\right] &= \sum_d^M \frac{h_{di}h_{dj}\overline{\nu}_d}{(\sum_i^N h_{di}\lambda_i)^2} \\ = \sum_d^M \frac{h_{di}h_{dj}}{\overline{\nu}_d} &= \sum_d^M \frac{h_{di}h_{dj}}{\sum_i^N h_{di}\lambda_i} \end{aligned} \tag{4.30}$$

The matrix form of (4.30) is given by

$$\boldsymbol{F}(\boldsymbol{\lambda}) = \boldsymbol{H}' D\left[\frac{1}{\overline{\nu}_d(\boldsymbol{\lambda})}\right]\boldsymbol{H} \tag{4.31}$$

The first assumption we make in (4.30), in order to calculate the FIM, is that the reconstruction is locally linear, meaning that the mean of the noisy reconstruction can be well estimated by the reconstruction of noiseless data. This in turn means that to calculate the FIM we need to know the activity distribution in advance. The exact FIM is a function of the reciprocal of the mean measurement data at individual sinogram bins, which are unknown for real datasets. Fessler and Rogers in [Fessler and Rogers, 1996] argued that, even for real noisy measurements, it is possible to predict the variance simply by replacing $\overline{\nu}$ with ν in (4.30). While this approximation works well for high count studies, it is heavily biased in low count situation, and needs further investigation. Moreover, another problem that could arise is the fact that the real measurement data may present zero values at individual sinogram bins, $\nu_d = 0$. In this case, the FIM would present infinite values, clearly leading to an incorrect estimation of the variance. This problem is avoided in the experiments presented in the rest of this thesis by always setting the background of the object to a non-zero value, therefore imposing $\nu_d \neq 0$.

It is also important to remark that the inverse of the Fisher information defines the lower bound on the variance of an unbiased estimator.

The Cramér-Rao Bound

For any unbiased estimator $\hat{\boldsymbol{\lambda}}$ of $\boldsymbol{\lambda}$, assuming that $\boldsymbol{F}$ is non-singular, the variance is bounded by the following inequality:

$$\mathrm{Var}(\hat{\lambda}_i) \geq \left[\boldsymbol{F}^{-1}(\boldsymbol{\lambda})\right]_{ii} \tag{4.32}$$

which is known as the Cramér-Rao Bound (CRB) [Cramér, 1946]. Placing a bound on the covariance matrix of the estimator, we obtain

$$\mathrm{Cov}(\hat{\boldsymbol{\lambda}}) \geq \boldsymbol{F}^{-1}(\boldsymbol{\lambda}) \tag{4.33}$$

where this inequality means only that $\mathrm{Cov}(\hat{\boldsymbol{\lambda}}) - \boldsymbol{F}^{-1}(\boldsymbol{\lambda})$ is positive semi-definite. An estimator that achieves this lower bound is called efficient. Therefore, if an estimator $\hat{\boldsymbol{\lambda}}$ is unbiased and efficient, the covariance matrix of the estimate is the Inverse of the Fisher Information Matrix.

Covariance bounds are useful for establishing performance limits of estimators, and for imaging system design. In the presence of regularisation, the estimators can be biased so the inverse of the Fisher information is not an accurate approximation to the estimator covariance.

The Cramér-Rao Bound for Biased Estimators

In order to present a succinct derivation of the Cramér-Rao bound for biased estimators, the definition of mean gradient of an estimator is introduced in the following:

$$\boldsymbol{U}_{ij} = \frac{\partial}{\partial \boldsymbol{\lambda}}\mathbb{E}[\hat{\boldsymbol{\lambda}}] = \left\{ \begin{matrix} \frac{\partial}{\partial \lambda_1}\mathbb{E}[\hat{\lambda_1}] & \frac{\partial}{\partial \lambda_2}\mathbb{E}[\hat{\lambda_1}] & \cdots & \frac{\partial}{\partial \lambda_N}\mathbb{E}[\hat{\lambda_1}] \\ \frac{\partial}{\partial \lambda_1}\mathbb{E}[\hat{\lambda_2}] & \frac{\partial}{\partial \lambda_2}\mathbb{E}[\hat{\lambda_2}] & \cdots & \frac{\partial}{\partial \lambda_N}\mathbb{E}[\hat{\lambda_2}] \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial}{\partial \lambda_1}\mathbb{E}[\hat{\lambda_N}] & \frac{\partial}{\partial \lambda_2}\mathbb{E}[\hat{\lambda_N}] & \cdots & \frac{\partial}{\partial \lambda_N}\mathbb{E}[\hat{\lambda_N}] \end{matrix} \right\} \tag{4.34}$$

where each column of $\boldsymbol{U}$ is the local-impulse response (LIR) as defined in (4.27) and therefore

$$\boldsymbol{U} = [\boldsymbol{F} - \beta \boldsymbol{R}]^{-1}\, \boldsymbol{F} \tag{4.35}$$

For an arbitrary estimator $\hat{\boldsymbol{\lambda}}$, whose mean gradient matrix is $\boldsymbol{U}$, its covariance matrix must satisfy:

$$\mathrm{Cov}(\hat{\boldsymbol{\lambda}}) \geq \boldsymbol{U} \cdot \boldsymbol{F}^{-1} \cdot \boldsymbol{U}' \tag{4.36}$$

Another form for this bound uses the bias:

$$\mathrm{Bias}(\boldsymbol{\lambda}) = \mathbb{E}[\hat{\boldsymbol{\lambda}}] - \boldsymbol{\lambda} \tag{4.37}$$

with its corresponding bias gradient matrix

$$\boldsymbol{Q} = \boldsymbol{U} - \boldsymbol{I} \tag{4.38}$$

Thus we can write the biased CRB as follows:

$$\mathrm{Cov}(\hat{\boldsymbol{\lambda}}) \geq [\boldsymbol{Q} + \boldsymbol{I}] \cdot \boldsymbol{F}^{-1} \cdot [\boldsymbol{Q} + \boldsymbol{I}]' \tag{4.39}$$

It should be noted that, if $\hat{\boldsymbol{\lambda}}$ is an unbiased estimator for $\boldsymbol{\lambda}$, then $\mathbb{E}[\hat{\boldsymbol{\lambda}}] = \boldsymbol{\lambda}$ so $\boldsymbol{U} = \boldsymbol{I}$ and $\boldsymbol{Q} = 0$, leading to the conventional CRB for unbiased estimators.

Moreover substituting (4.35) in (4.36) we obtain

$$\mathrm{Cov}\{\hat{\boldsymbol{\lambda}}\} \geq [\boldsymbol{F} - \beta \boldsymbol{R}]^{-1}\, \boldsymbol{F}\, [\boldsymbol{F} - \beta \boldsymbol{R}]^{-1} \tag{4.40}$$

which is the Cramer-Rao bound for penalised likelihood estimators when the probability associated with the data measurements is Poisson distributed.

4.5 Efficient Calculation of the FIM

In the previous section, it has been described how the Fisher Information Matrix can be employed to characterise the statistical properties (such as mean and covariance) of a PL estimator. Unfortunately, computing the FIM inverse is intractable, since we are dealing with a large matrix of size $N \times N = (N_x \times N_y \times N_z)^2$.

4.5.1 The Circulant Approximation of the FIM

A computationally efficient approximation in calculating the inverse of the FIM has been previously proposed for the design of space-variant penalties that yield space-invariant impulse response functions [Stayman and Fessler, 2000] [Stayman and Fessler, 2004a] [Qi and Leahy, 2000].

Qi and Leahy [2000] argued that if it is reasonable to assume that the FIM varies slowly with position and if one is interested in calculating the effects of a quadratic prior in terms of bias gradient and variance in a voxel i, then it is acceptable to ignore the shift-variance of the FIM. The basic idea is to use a local invariance assumption in the context of shift-variant system modelling.

The computations are done for voxel i and therefore only the i-th row of the FIM needs to be calculated. For representation purposes, the elements of the i - th column of the Fisher Information Matrix can be re-ordered in order to represent an “image” associated with the i - th voxel. The assumption implies that these “images” vary smoothly as one moves to the column of the FIM associated with the neighbouring voxels. A second assumption implies that these “images” have a local support; meaning that, for the i - th column, the non-zero values are concentrated in the proximity of the (i, i) location in the FIM. Therefore when estimating the $LLIR^i$ and the variance at the voxel i, we assume that these are largely determined by the i - th column of the FIM.

This local approximation of the FIM is obtained by replacing all rows of the FIM with the shifted version of its $i - th$ row so that the resulting matrix $\boldsymbol{F}^i$ has a block Toeplitz structure. This shift-invariant matrix is then inverted in order to estimate the variance in each voxel i. This approximation is referred to as the *circulant approximation*, since it simply reduces the FIM to a circulant matrix. The computations in (4.21) therefore become tractable as a circulant matrix can be diagonalised using a Discrete Fourier Transform (DFT) . It is then possible to rewrite the formulas of the covariance and of the $LLIR$ for a voxel i, in the Fourier domain, as [Stayman and Fessler, 2000]:

$$\mathrm{Cov}^{i}_{circ}(\hat{\boldsymbol{\lambda}}) \approx \mathcal{F}^{-1}\left\{\frac{\mathcal{F}\{\mathrm{e}_i\}\cdot\mathcal{F}\{\boldsymbol{H}^T D[1/\overline{\nu}_d]\boldsymbol{H}\mathrm{e}_i\}}{|\mathcal{F}\{\boldsymbol{H}^T D[1/\overline{\nu}_d]\boldsymbol{H}\mathrm{e}_i\}+\mathcal{F}\{\beta\boldsymbol{R}\mathrm{e}_i\}|^2}\right\} \tag{4.41}$$

$$LLIR^{i}_{circ}(\hat{\boldsymbol{\lambda}}) \approx \mathcal{F}^{-1}\left\{\frac{\mathcal{F}\{\mathrm{e}_i\}\cdot\mathcal{F}\{\boldsymbol{H}^T D[1/\overline{\nu}_d]\boldsymbol{H}\mathrm{e}_i\}}{\mathcal{F}\{\boldsymbol{H}^T D[1/\overline{\nu}_d]\boldsymbol{H}\mathrm{e}_i\}+\mathcal{F}\{\beta\boldsymbol{R}\mathrm{e}_i\}}\right\} \tag{4.42}$$

where $\cdot$ denotes element-by-element multiplication and the division is an element-by-element division. The term e_i is the unit basis vector for the voxel i. The function $\mathcal{F}\{\cdot\}$ takes the DFT of its argument and $D[\cdot]$ produces a diagonal matrix whose diagonal entries are the reciprocal of the noiseless projection data $\overline{\nu}_d$.

The true Fisher Information $\boldsymbol{F}$ must be a positive semi-definite matrix, or equivalently, its eigenvalues must be real and non-negative. The circulant matrix used as an approximation is not guaranteed to have this property. Consequently, the row of the FIM, $[\boldsymbol{F}\cdot\mathrm{e}_i]$, is modified in order to impose the symmetry condition. The elements of the i-th row of $\boldsymbol{F}$ are first re-ordered as a 3-D matrix f. For an $N_x \times N_y \times N_z$ imaging volume, this 3-D matrix is shifted so that the i-th voxel is moved to the centre voxel $(N_x/2+1, N_y/2+1, N_z/2+1)$. In order to ensure that the 3 - D Fourier coefficients are real, the following symmetry is introduced:

$$f(i,j,k) = \max\{f(i,j,k), f(N_x-i+1, N_y-j+1, N_z-k+1)\} \tag{4.43}$$

Finally, any negative coefficient of the DFT of the resulting matrix are truncated to zero.

For a uniform quadratic prior, its hessian $\boldsymbol{R}$ already has a block Toeplitz structure. However, if a spatially variant prior is used, it is possible to apply the locally invariant approximation $\boldsymbol{R}^i$ in a similar manner to that described above for $\boldsymbol{F}$.

From (4.41) we can see how the approximated estimate of $\mathrm{Cov}^i(\hat{\boldsymbol{\lambda}})$ for a single voxel position i can be simply computed with a projection, a backprojection and a few fast Fourier transforms [Nuyts, 2009]. However, it should be noted that the aforementioned method must be repeated for each voxel under investigation.

The circulant approximation method is well suited for the calculation of the uncertainty for systems whose response can be approximated as shift-invariant; or in the case one wants to estimate the local effects that a penalty function has on bias gradient and variance. However this method does not account for the global interdependence between the estimates in all the voxels; it therefore can not incorporate the effects of long-range correlations (e.g., evaluation of the effects of data truncation or missing data).

4.5.2 A Subsampled version of the FIM

In this section we propose a different approach for calculating the inverse of the FIM. The FIM is calculated over a subset of voxels arranged in a grid that covers the whole volume. This formulation reduces the computational complexity in inverting the FIM but nevertheless accounts for the global interdependence between the variables. The FIM is calculated over a subset of the voxel indexes $\boldsymbol{G} \subset [1, \ldots, N]$ arranged in a grid that covers the whole volume. We define a subsampled version of the FIM calculated over a subset $\boldsymbol{\lambda_G}$ of the full set of parameters $\boldsymbol{\lambda}$:

$$F_{ij}^{G} = \sum_{d=1}^{M} \frac{h_{di} h_{dj}}{\sum_{b=1}^{N} h_{bd} \lambda_b} \quad \text{with } i, j \in G \tag{4.44}$$

This is equivalent to saying that, in the estimation of the covariance, we are accounting for the interdependence between a subset of voxels $i, j \in \boldsymbol{G}$ only, assuming that, for the remaining voxels $i, j \notin \boldsymbol{G}$, the PL estimate $\hat{\boldsymbol{\lambda}}$ is equal to the true value of $\boldsymbol{\lambda}$.

The approximate deterministic calculation of the covariance is simply obtained by substituting the FIM with its subsampled (4.44) version in the definition of the covariance matrix presented in (4.21):

$$\mathrm{Cov}(\hat{\boldsymbol{\lambda}}_{\boldsymbol{G}}) \approx [\boldsymbol{F}^{\boldsymbol{G}} - \beta \cdot \boldsymbol{R}^{\boldsymbol{G}}]^{-1} \cdot \boldsymbol{F}^{\boldsymbol{G}} \cdot [\boldsymbol{F}^{\boldsymbol{G}} - \beta \cdot \boldsymbol{R}^{\boldsymbol{G}}]^{-1} \tag{4.45}$$

and equivalently, the approximate deterministic calculation of the $LLIR$ is simply obtained by substituting (4.44) in (4.27):

$$LLIR^{i}(\hat{\boldsymbol{\lambda}}_{\boldsymbol{G}}) \approx [\boldsymbol{F}^{\boldsymbol{G}} - \beta \cdot \boldsymbol{R}^{\boldsymbol{G}}]^{-1} \cdot \boldsymbol{F}^{\boldsymbol{G}} \mathrm{e}_i \tag{4.46}$$

The number of elements in the full FIM equals N^2, whereas the number of elements of the subsampled FIM equals N_G^2; therefore reducing the computational burden in inverting the subsampled FIM.

The Hessian of the quadratic penalty $\boldsymbol{R}$ is not dependent on $\boldsymbol{\lambda}$ and therefore can be precalculated. Analogously to $\boldsymbol{F}$, the subsampled version of the penalty function $\boldsymbol{R}^G$ is obtained by selecting the elements of the matrix $\boldsymbol{R}$ that correspond to the points in the grid. In the methodology presented in the next section, a quadratic penalty function with a small weight has been included with the only purpose of enforcing the estimator to have a unique solution and therefore to guarantee the regularised FIM ($[\boldsymbol{F} - \beta \boldsymbol{R}]$) to be invertible. However, since in practice we subsample the matrix, the Hessian of the quadratic prior $\boldsymbol{R}$ reduces to a diagonal matrix $\boldsymbol{R}^G = \boldsymbol{R}_{ii} * \boldsymbol{I}$ (where $\boldsymbol{R}_{ii}$ is the value of the diagonal elements of the uniform quadratic prior $\boldsymbol{R}$ and $\boldsymbol{I}$ is an identity matrix of size $N_G \times N_G$). Thus, the addition of the term $\beta \boldsymbol{R}^G$, to aid the inversion of $[\boldsymbol{F}^G - \beta \boldsymbol{R}^G]$, can be seen as a sort of Tikhonov regularisation.

Two examples of grids are pictured in Figure 4.1, for a small imaging volume of 6 × 6 × 6 voxels. In Figure 4.1-A the grid accounts for the interdependence between every point in the imaging volume, in Figure 4.1-B, the grid accounts for the interdependence between half of the voxels in the imaging volume. This model allows the user to design the grid and therefore to define the degree of approximation in the calculation of the FIM. In section 4.5.4, visual representations of the FIM and of the Covariance matrix are presented for the full FIM, the subsampled FIM and the circulant approximation. A discussion on how the missing FIM entries between the grid points affect the accuracy of the results is also presented in section 4.7.

4.5.3 GPU Accelerated Implementation

Every element of the FIM at the grid points is calculated exactly, accounting for the acquisition geometry and the object without further approximation. If the grid has N_G nodes, the FIM is of size $N_G \times N_G$ and symmetrical, so filling the matrix requires the computation of $\frac{1}{2}N_G^2 + \frac{1}{2}N_G$ elements. Naive computation of the FIM requires one projection for the denominator of (4.44) and M sums of products (SOPS) for each of the $\frac{1}{2}N_G^2 + \frac{1}{2}N_G$ elements of the half FIM. The proposed algorithm is inspired by the rotation-based algorithm proposed by Zeng and Gullberg [Zeng and Gullberg, 1992]. The collimator-detector response is captured by a depth dependent Point Spread Function (PSF). Information being additive over the detector bins, the FIM element F_{ij} is the sum of F_{ij}^m contributions from the M camera positions indexed with

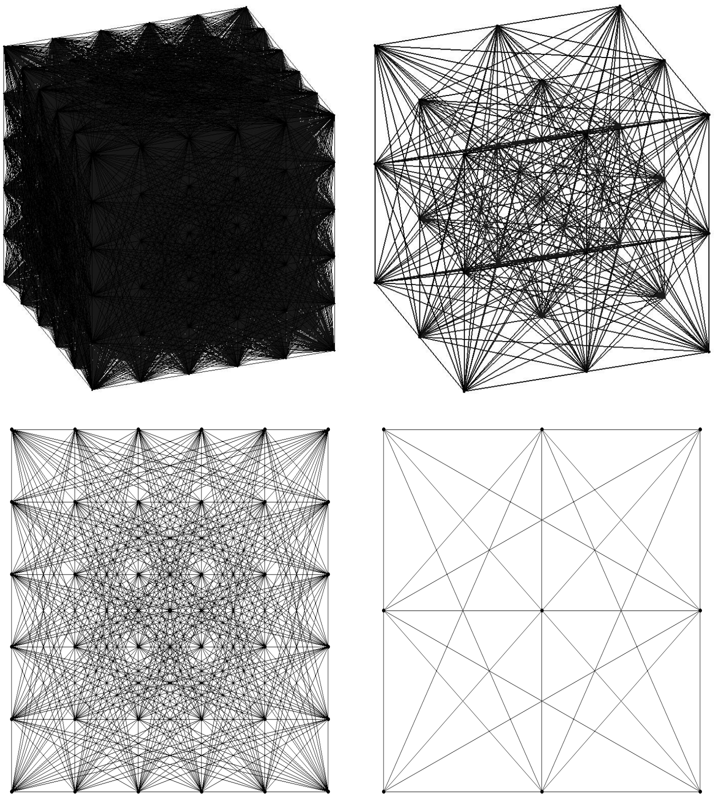

Figure 4.1: Example of grids for the estimation of the uncertainty. A (top left) the grid accounts for the correlation between every point in the imaging volume (full FIM). B (top right) the grid accounts for the correlation between $1/8$ of the voxels in the imaging volume (subsampled FIM). C (Bottom left) The central plane of the grid displayed in A. D (Bottom right) The central plane of the grid displayed in B.

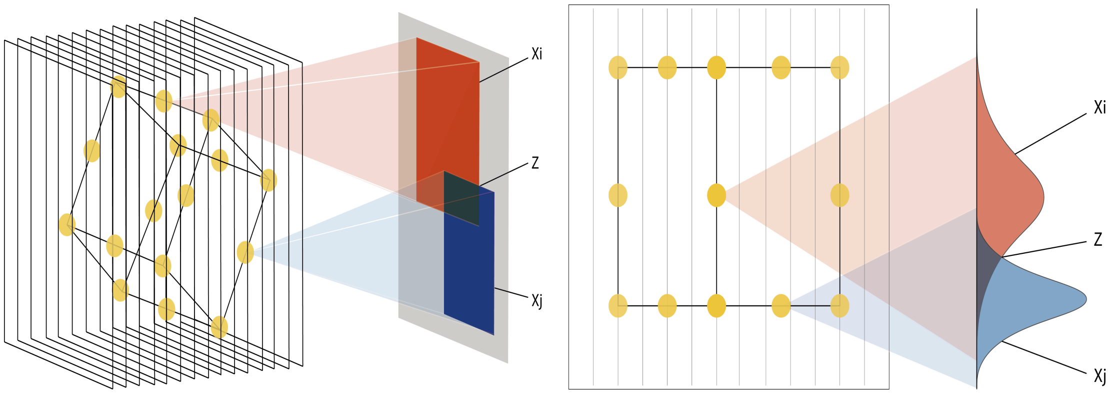

Figure 4.2: Rotation-based algorithm for fast computation of the Fisher Information Matrix: 3-D schematic representation. Algorithm: 1) For each Gamma camera (grey plane) position, the activity $\boldsymbol{\lambda}$ is resampled on a regular grid parallel to the camera plane and projected; 2) The grid points (yellow spheres) for the Fisher Information Matrix (FIM) are reinterpolated on the same parallel grid; 3) For each pair of points in the FIM grid, the FIM element is updated with the information relative to the current camera, which only depends on the region of the projection Z (black square) where the two PSFs X_i (red square) and X_j (blue square) intersect.

$m = 1, \cdots, M$. The algorithm is based on interpolation of the activity and of the FIM grid on a regular grid aligned with each camera. By re-interpolating the activity and the FIM grid on a regular grid, the PSF can be applied more efficiently in the frequency domain as all points that are at a given distance lie on the same plane. The PSF is non-zero within a box X (see Figure 4.2). The algorithm for the evaluation of the elements of the FIM consists of the following steps:

1) Compute projection of $\boldsymbol{\lambda}$ for each camera position.

2) Compute FIM elements for each camera position:

a. Re-sample the FIM grid positions on the voxel grid parallel to the camera by tri-linear interpolation.

b, For each pair of points i, j in the FIM grid

   i. Compute coordinates of the box Z, on the camera plane, where the two PSFs X_i X_j intersect (if they intersect).

Table 4.1: Computation times for the calculation of the FIM and its inverse.

	g1	g2	g3
Grid Size	27648	6912	3072
NVidia GeForce GTX-285	348s	30s	19s

Reference method	18 hours

ii. If X_i and X_j intersect, update the FIM element by integrating (4.44) over the intersection box Z.

The algorithm is implemented in the CUDA (aka Compute Unified Device Architecture) programming language for parallel execution on Graphics Processing Units (GPU). Tri-linear resampling is performed in hardware by the texture fetch unit of the GPU at the cost of a single memory access. Coalesced memory access is achieved by partitioning the memory transfers in blocks. The convolutions are calculated with the 2D-FFT (Fast Fourier Transform) and IFFT (Inverse Fast Fourier Transform) routines included in the NVidia CUFFT library. A tailor made GPU kernel computes the projection (sum of planes) with high device occupancy and maximises memory coalescing. A second kernel computes the integral in each intersection box (2-b-ii): each GPU thread computes the integral (for the current camera position), for a pair of points in the grid, so that the integrals for multiple pairs are evaluated concurrently on the multi-processors of the GPU. Each thread decides if the two PSFs intersect, then it loads from the global memory of the GPU device the sections of the PSFs that intersect and the projection data in the area of intersection Z (see Figure 4.2). Finally, the thread computes the integral (2-b-ii) in the intersection box. After completion of the partial FIM for a single camera position, the process is repeated for another camera position, accumulating the elements of the FIM, as, according to (4.44), information is additive. Computation times are reported in table 4.1. The GPU-accelerated algorithm for the computation of the FIM has been implemented by Stefano Pedemonte. This algorithm has been integrated in the Niftyrec reconstruction software toolbox [Pedemonte et al., 2010] and has Matlab and Python interfaces which enable real time scripting interaction and full flexibility in the definition of the grid.

4.5.4 Subsampled FIM vs Circulant FIM

In this section, the link between the shift-variance of the system and the different approximations of the FIM is illustrated with an example. Figure 4.3 show the full Fisher Information Matrix (Figure 4.3 A-B), the subsampled Fisher Information Matrix (Figure 4.3 C-D) and the Circulant Fisher Matrix (Figure 4.3 E-F) for a simple experiment where the system is a standard SPECT camera equipped with a Low Energy High Resolution (LEHR) collimator. One element of the FIM is, by definition (equation (4.29)), the second order mixed partial derivative of the log-likelihood with respect to the activity in two locations. These figures use “linear indices” through the volume to which grey values proportionally to the FIM elements are assigned. The indexing pattern of choice is the somewhat natural raster scanning pattern, where we start in the top left corner of the image and we run through the rows and columns.

The SPECT system is based on a detector that rotates 360% around the centre of the imaging volume. The imaging volume dimensions are $96 \times 96 \times 1$ cubic voxels of 2.46 mm. The detector efficiency terms c_d, the attenuation factors a_{di} and the spatial variation in sensitivity terms s_i are set to 1. The geometric response is depth-dependent and therefore shift variant. The SPECT camera is placed at a distance of 123 mm from the centre of the imaging volume and is equipped with a LEHR collimator. The intrinsic response is set to 3.6 mm, the collimator has a linearly varying depth-dependent Gaussian response that has a slope of 0.0562, which corresponds to 7.89 mm FWHM at the centre of the field of view.

The phantom used was a uniform disk positioned at the centre of the image space, with diameter $D = 24.6$ mm. The level of activity in the background was set to the 10% of the activity in the disk. The uniform background was a disk positioned in the centre of the image space, with radius $106.3\ mm$ and height $29.5\ mm$.

It should be noticed that solving equation (4.41), for a voxel of interest i, is equivalent to the inversion of a column of the FIM as if the full FIM were a block circulant matrix with circulant blocks (for the 2-D case), which we refer to as $\tilde{\boldsymbol{F}}$. This is equivalent to creating a new matrix by extracting the i-th column from the full FIM, $\tilde{\boldsymbol{F}}^i = \boldsymbol{F}^i$, and then obtaining from this vector the remaining columns of $\tilde{\boldsymbol{F}}$ by an appropriate circulant shift in 2D so that the peak of $\boldsymbol{F}^i$ becomes centered at the voxel corresponding to each column index.

Therefore we show, in Figure 4.3 E-F, the circulant Fisher Matrix for the calculation of the variance of a voxel of interest i in the centre of the FOV; whereas we show in Figure 4.3 A-B the FIM for all the points in the FOV and in Figure 4.3 C-D the FIM for $1/4$-th of the points. As already described in section 4.5.3, all elements of the subsampled FIM at the grid points are calculated exactly, accounting for the acquisition geometry and for the object. We can see, in fact, how in the subsampled FIM we account for the system response and for the object dependency, whereas with the circulant approximation method we make the assumption that the FIM (and therefore the system response) is shift invariant. The number of elements in the full FIM and in the circulant FIM equals N^2, whereas the number of elements of the subsampled FIM equals $(N/2)^2$; therefore highly reducing the computational burden in inverting the subsampled FIM. Since the FIM is a very large matrix and therefore difficult to display, we show the FIMs for the 2-D case. Therefore the full FIM has 9216 points and the subsampled FIM has 2304 points.

However, what is ultimately of interest is the inverse of $(\boldsymbol{F} - \beta\boldsymbol{R})$ (the Covariance matrix). Figure 4.4 shows the Covariance matrix calculated from the full FIM (Figure 4.4 A), the Covariance matrix calculated from the subsampled FIM (Figure 4.4 D) and the Covariance matrix calculated using the circulant approximation method (Figure 4.4 B). Figure 4.4 B displays the Covariance matrix obtained by row-by-row inversion of the circulant FIM, where each column of the matrix is evaluated separately using (4.41). Symmetry is then imposed on the circulant Covariance matrix by performing $\mathrm{Cov}_{circ} = (\mathrm{Cov}_{circ} + \mathrm{Cov}'_{circ})/2$. The resulting Covariance matrix obtained using the circulant approximation is spatially variant (non circulant) but clearly does not show the same structure as the full FIM inverse, $(\boldsymbol{F} - \beta \cdot \boldsymbol{R})^{-1}$, in Figure 4.4 A. We can therefore deduce that the circulant FIM can not incorporate the effects of shift-variancy, since it does not account for a depth-dependent collimator response. A direct visual comparison between the inverse of the full FIM, $(\boldsymbol{F} - \beta \cdot \boldsymbol{R})^{-1}$, and the inverse of the subsampled FIM, $(\boldsymbol{F}_G - \beta \cdot \boldsymbol{R}_G)^{-1}$, is arduous, because of the different size of the two matrices. Hence we show, in Figure 4.4 C, a matrix which are obtained selecting the voxels at the locations of the covariance matrix obtaind from the full FIM that correspond to locations of the elements of the subsampled matrix and then rebinning the selected voxels in a smaller matrix of size $(N/2)^2$. Clearly, the inverse of the subsampled FIM, $(\boldsymbol{F}_G - \beta \cdot \boldsymbol{R}_G)^{-1}$, will not be exact at the grid points because of the missing off-diagonal FIM entries between the grid points, however the matrices in 4.4 C and 4.4 D exhibit the same structure. This encompass the capability of the method to incorporate non-stationary system functions and effects of long-range correlations.

4.6 Image Quality Quantification for System Design

In this paragraph, a figure of merit for SPECT system design based on the trade-off between the bias and the variance that can be achieved in the reconstruction of emission tomograms, is introduced.

Under the assumption that the system matrix $\boldsymbol{H}$ is non-singular and imposing $\beta = 0$ in (3.2), the maximum likelihood estimator is asymptotically efficient and asymptotically unbiased. One approach to system design, for the unbiased estimator, is to choose the parameters of the imaging system that would produce the least error (minimum variance) in the asymptotic case. This simply involves the inversion of the FIM, to obtain the covariance of the estimator, and is referred to as the Cramer-Rao bound (see equation (4.32)).

However such approach is problematic because, in practice, the full rank property of the system matrix (non-singularity) is quite difficult to verify. This problem is addressed by including the regularisation penalty of equation (3.2) that leads to a strictly concave cost function and makes $\boldsymbol{F} - \beta \cdot \boldsymbol{R}$ in (4.21) invertible. However bias is unavoidable for penalised estimators, so the unbiased Cramer-Rao bound is not-applicable.

The approach that is widely used in emission imaging is to obtain a local measure of the bias gradient, using the Linearised Local Impulse Response(LLIR) for the i-th voxel as in equation (4.27), and to consider the trade-off between bias and variance for the optimisation of the system. In [Hero et al., 1996], a particular type of PL estimator including an appropriate space-variant quadratic smoothing prior has been shown to achieve the Uniform Cramer Rao Bound (UCRB). However the space-invariant prior further contributes to space-variance in the bias gradient, as it has been shown in section 3.2.1. Therefore including and designing an appropriate space-variant penalty function lacks practical justification for the optimisation of the design of the imaging system.

To enable comparison between different systems at equal bias gradient, we rely on an adaptation of (4.45) and (4.46) where a post-smooth filter $\boldsymbol{P}$ is added to the equations:

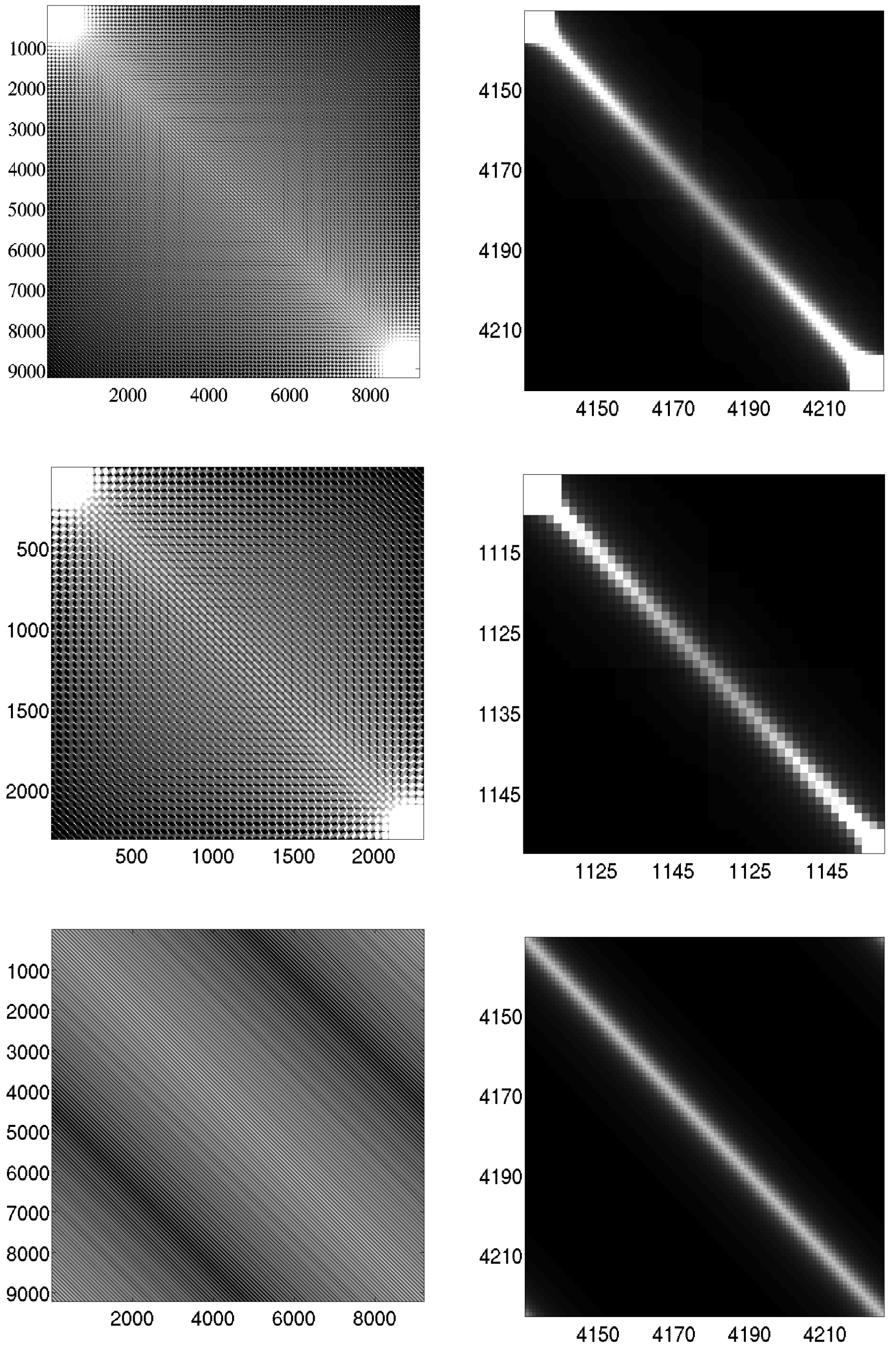

Figure 4.3: Fisher Information Matrix for the experiment in section 4.5.4. Comparison between full FIM, subsampled FIM and Circulant FIM. A (top left) - Full FIM, grid 9216 with points. B (top right) - A zoom-in part of A. C (central left) - subsampled FIM, grid with 2304 points. D (central right) - A zoom-in part of C. E (bottom left) - Circulant FIM for a voxel of interest i in the centre of the FOV only. F(bottom right) - A zoom-in part of E

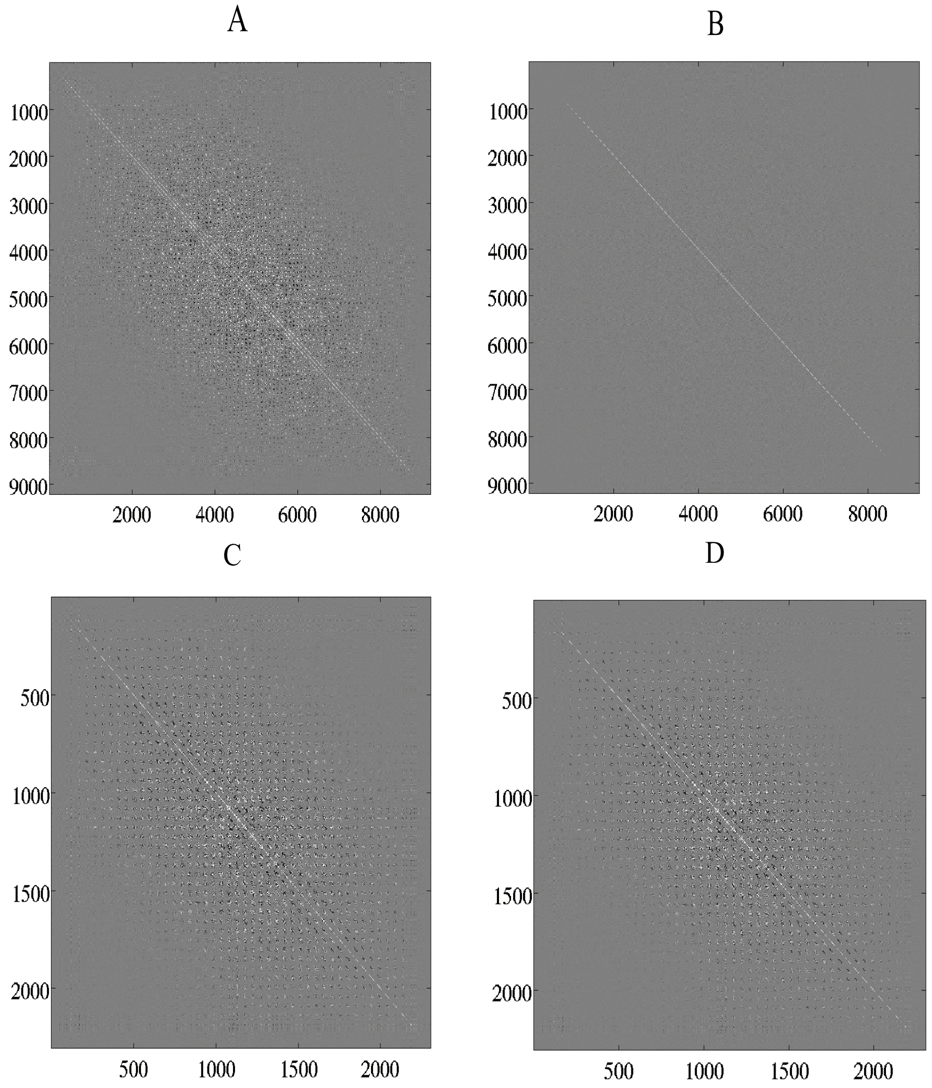

Figure 4.4: Inverse of $(\boldsymbol{F}_G - \beta \cdot \boldsymbol{R}_G)^{-1}$ for the experiment in section 4.5.4. Comparison between the Covariance matrix obtained from the full FIM, the subsampled FIM and obtained by row-by-row inversion of the circulant FIM, where each column of the matrix is evaluated separately using (4.41). A - Covariance matrix obtained from the Full FIM, grid with 9216 points. B - Covariance matrix obtained by row-by-row inversion of the circulant FIM. C - Matrix obtained selecting the voxels at the locations of the full FIM that corresponds to locations of the elements of the grid for the subsampled FIM. D - Covariance matrix obtained from the subsampled FIM, grid with 2304 points.

$$\mathrm{Cov}(\hat{\boldsymbol{\lambda}}_{\boldsymbol{G}}) \approx \boldsymbol{P}^T \cdot [\boldsymbol{F^G} - \beta \cdot \boldsymbol{R^G}]^{-1} \cdot \boldsymbol{F^G} \cdot [\boldsymbol{F^G} - \beta \cdot \boldsymbol{R^G}]^{-1} \cdot \boldsymbol{P} \tag{4.47}$$

$$LLIR^i(\hat{\boldsymbol{\lambda}_G}) \approx \boldsymbol{P}^T \cdot \boldsymbol{F^G} \cdot [\boldsymbol{F^G} - \beta \cdot \boldsymbol{R^G}]^{-1} \cdot \boldsymbol{P} \cdot \mathrm{e}_i \tag{4.48}$$

where $\boldsymbol{F^G}$ and $\boldsymbol{R^G}$ are the subsampled version of the FIM and the subsampled version of the hessian of quadratic penalty, respectively. It should be noted that $\boldsymbol{F^G}$ and $\boldsymbol{R^G}$ can be substituted with circulant approximation of the FIM, $\boldsymbol{F^{circ}}$, and with the circulant approximation of the Hessian of the penalty function, $\boldsymbol{R^{circ}}$, in the above equation, if one wants to rely on the circulant approximation method for the calculation of bias and variance in a voxel of interest i.

In equations (4.47) and (4.48), the penalty function $\boldsymbol{R}$ with a small regularisation parameter is included with the only purpose of making the cost function strictly concave. Therefore the bias property of the estimator is mainly determined by the filter function $\boldsymbol{P}$. In order to compare different systems, first, a target bias gradient function is defined as an isotropic Gaussian $\boldsymbol{P}_t$ (described by its Full Width at half Maximum $FWHM_t$). Consequently, for every system, an anisotropic post-smooth filter $\boldsymbol{P}$ is designed, so that the $LLIR^i$ in (4.48) matches the target isotropic Gaussian function $\boldsymbol{P}_t$. Designing a specific post-smooth filter for every system under investigation, the noise properties of the estimator can be compared at equal bias gradient. For more details on the achievability of the bound for this method and on how it compares with UCRB, see [Meng and Clinthorne, 2004].

The method presented in equations (4.48) and (4.47) corresponds to iterating the algorithm used to maximise the PL objective function (3.2) to convergence and then convolving the solution with an anisotropic filter $\boldsymbol{P}$ in order to impose a fixed target bias gradient $\boldsymbol{P_t}$.

Relying on this method a post-smooth filter $\boldsymbol{P}$ has to be specifically designed for every system under investigation. A method for the design of this filter was first introduced by Zhou et al. [2010] and is described in the following.

Design of the Post-Smooth Filter

Firstly, for every system η, the Local Impulse Response $LLIR^i_\eta(\hat{\boldsymbol{\lambda}})$ (which is described by its $FWHM_\eta$) is calculated as in (4.46). An isotropic Gaussian target function $\boldsymbol{P}_t$ is defined so that $FWHM_t \geq FWHM_\eta$ (which implies that post-smoothing is always needed to achieve the target bias gradient). The post smoothing filter $\boldsymbol{P}$ is finally defined taking into account the deviation between $LLIR^i_\eta(\hat{\boldsymbol{\lambda}})$ and the target isotropic Gaussian $\boldsymbol{P}_t$:

$$\mathcal{P}[k] = \begin{cases} \mathcal{P}_t[k]/\mathcal{B}^i_\eta[k] & \text{if } \Re(\mathcal{B}^i_\eta[k]) \geq 0.005 \\ 0 & \text{if } \Re(\mathcal{B}^i_\eta[k]) < 0.005 \end{cases} \tag{4.49}$$

where $\mathcal{P}_t$, $\mathcal{P}$ and $\mathcal{B}^i_\eta$ are the Fourier transforms of $\boldsymbol{P}_t$, $\boldsymbol{P}$ and $LLIR^i_\eta(\hat{\boldsymbol{\lambda}})$, k is the index of the elements in the Fourier domain and $\Re(\cdot)$ denotes the real part of a complex number. Applying the filter $\boldsymbol{P}$ to the LLIR in (4.46) ensure that the final Local Impulse Response $LLIR^i(\hat{\boldsymbol{\lambda}})$ in (4.48) equals $\boldsymbol{P}_t$. The only parameter to optimize then, is the variance $\mathrm{Var}_i = \mathrm{Cov}_{ii}$ with Cov_{ii} calculated as in equation (4.47).

4.6.1 Figure Of Merit: the Contrast to Noise Ratio

We can now reduce (4.47) and (4.48) to a scalar measure by taking into account only the variance and the Contrast Recovery Coefficient (CRC) for the voxel i, which are defined as $Var_i = \mathrm{Cov}_{ii}$ and $CRC_i = LLIR_{ii}$ (the CRC can be seen as an alternative to the FWHM as a measure of bias, see section 3.4.1). Thanks to the fixed resolution after post-smoothing, the CRC should be more or less constant; the only parameter to optimise then is the variance. In the rest of this thesis, however, we consider the Contrast to Noise Ratio (CNR) in voxel i as figure of merit for image quality:

$$CNR_i = \frac{CRC_i}{\sqrt{Var_i}} \tag{4.50}$$

4.7 Discussion

In this chapter, approximate expressions of the mean and the covariance of the PL estimator has been obtained via the FIM, as expressed in (4.3) and (4.21).

In order to tackle the problem of the computational load in calculating and inverting the FIM, it has been proposed to approximate it with a circulant matrix (see section 4.5.2). The circulant FIM is generally considered to yield a good approximation of the covariance matrix for nearly shift-invariant systems, however 3-D imaging systems are inherently shift variant, presenting a block-circulant FIM even in case of an ideal uniform object in the FOV.

In section 4.5.2, we introduced a new approximation which relies on a subsampled version of the FIM that addresses the shortcomings of the circulant approximation (see section 4.5.4).

The sub-sampled FIM trades off computational complexity and accuracy of the estimation, enabling the adaptation of the accuracy of the estimation based on the available computational resources. When sufficient resources are available, the GPU-accelerated software described in section 4.5.3 can compute the covariance matrix exactly using the full FIM, $\boldsymbol{F}$. One important advantage of the scalable sub-sampled FIM approximation is that the algorithm provides an estimate of the full covariance matrix, though sub-sampled, accounting for the global interdependence between the variables of the tomogram.

Evaluation of the effect of sub-sampling, however, is complicated by the trade-off that arises. It is not possible to define an absolute criterion for the choice of the sub-sampling scheme. The contributions to the FIM at a given camera position, for a given pair of grid points, arise only from overlap in the projected PSF from those points (as shown graphically in Figure 4.2). The implication is that the grid points must be close enough to ensure there is overlap between the projected PSFs. This condition depends on many factors, including: the image volume size, the voxel size, the size of the PSF, the camera trajectory etc. This condition refers to accuracy of the FIM entry for those two points, but does not apply to accuracy of its inverse, which will suffer from missing points even if the "overlapping PSF" condition is met. Therefore, a general criterion to define a relationship between the subsampling and the reliability of the variance estimation, can not be provided. This criterion depends in fact on the properties of the specific system.

The sub-sampled FIM formulation and the software tool described in this paper may be employed for the optimisation of a range of design parameters of emission imaging systems. However only three guidelines can be given, so far, for the choice of the subsampling scheme. The first is trivially to adopt the most dense grid for the available computational resources. The second is to restrict the grid volume to a specific region of the FOV, in case we know in advance that activity is present only in that region of interest. The third refers to adaptive imaging systems, where the system adapts during acquisition, in response to the projection data, and therefore where the computational resources are limited by the real-time requirements. For a specific adaptive imaging system and for the specific parameter we want to modify during acquisition, a sufficient condition of optimality needs to be defined. This condition accounts for the trade-off between accuracy of the estimate and computational complexity. Once a sufficient condition of optimality is defined, the subsampling model should be chosen prospectively by comparing the estimates of the optimum scanning parameters for different subsampling models, with the parameters obtained from the reference method.

4.8 Proposed Experimental Set-Up

The model presented in section 4.5.2 allows the user to design the grid and therefore to define the degree of approximation in the calculation of the FIM. For the deterministic method, the subsampled version of the FIM has been calculated over three different grids of $g_1 = 27648$, $g_2 = 6912$ and $g_3 = 3072$ points equally distributed over the slice intersecting the point (or region of interest) of interest and the two neighbouring slices. It should be noted that grid g_1 is fully sampled over the three slices of interest and therefore, in the following, this FIM will be referred to as the full FIM.

The variance images displayed in the following chapters (5, 6 and 7), are obtained by re-ordering the diagonal of the covariance matrix $\mathrm{Cov}(\hat{\boldsymbol{\lambda}})$ calculated as in (4.45). For the full FIM (g_1), the diagonal of $\mathrm{Cov}(\hat{\boldsymbol{\lambda}})$ is simply reshaped to a three dimensional matrix, whereas for grids g_2 and g_3 every point of the diagonal of $\mathrm{Cov}(\hat{\boldsymbol{\lambda}})$ is allocated to the respective points of the grid in the imaging volume. A trilinear interpolation is then performed in order to facilitate the visual comparison between the variance images obtained with the different subsamplings. It should be noted that a direct interpolation on $\mathrm{Cov}(\hat{\boldsymbol{\lambda}})$ can not be performed. However in order to calculate the CNR in (4.50), which is the figure of merit for the optimisation of the system design, we rely directly on the subsampled $\mathrm{Cov}_G(\hat{\boldsymbol{\lambda}})$, without performing any interpolation.

To clarify this point, Figure 4.5 is included, which refers to a 2-D imaging system. In Figure 4.5 - A we show an example of grid g_2 for the estimation of the uncertainty. The image voxel indices are in red and the points for a grid g_2 are in grey color. Figure 2 - B shows the respective covariance matrix. Grey pixels are the locations where the elements of the Covariance obtained with the subsampled FIM are non-zero. Figure 4.5 - C shows a compact representation of the Covariance matrix obtained with the subsampled FIM. The number of cells in B is equal to the square of the number of cells in A and the number of grey cells in B is equal to the square of the number of grey cells in A. This example points out that direct bilinear interpolation of $\mathrm{Cov}_G(\hat{\boldsymbol{\lambda}})$ can not be performed and perhaps filling of the covariance matrix is material for

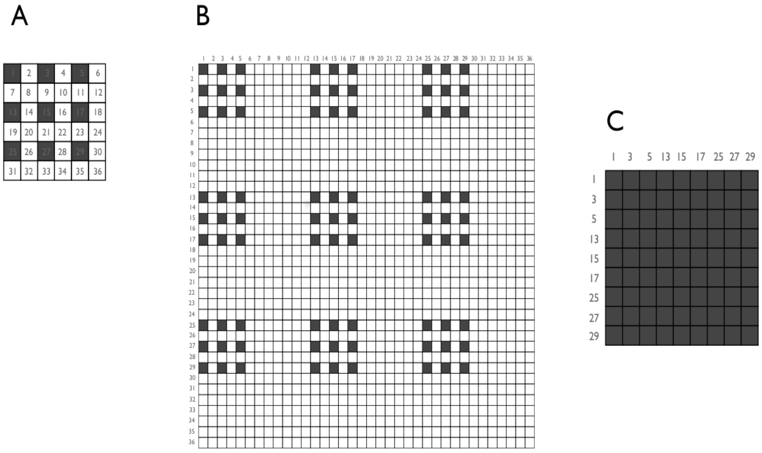

Figure 4.5: Schematic representation of a grid and respective subsampled covariance matrix. A - An example of grid g_2 is shown. The image voxel indices are in red and the grid points are grey. B - Respective covariance matrix. Grey pixels are the locations where the elements of the covariance matrix obtained with the subsampled FIM are non-zero. C - Compact representation of the covariance matrix obtained with the subsampled FIM.

further investigation.

A more closely spaced grid gives a more precise estimation of the variance but at the cost of increased computational complexity of the estimation. The computation time needed to calculate the FIMs and their inverse are presented in table 4.1. Variance images obtained with the circulant approximation will be also presented for comparison. Every pixel of the images is calculated according to equation (4.41).

Chapter 5

Collimator Design: Resolution-Sensitivity trade-off

In SPECT, the design of the collimator determines the detection efficiency and the system resolution. Hence it is one of the key factors influencing the system response.

For an image to be formed, ideally, there must be a one-to-one correspondence between the direction of emission of a γ-ray from a radioactive source and its point of detection. As already described in 2.1.3, the purpose of the collimator is to achieve such spatial localisation of the photon emissions, in the imaging space.

Although a large number of collimator designs is available, in this chapter we focus on the optimisation of Parallel Holes (PH) collimators, which are the most widely used collimators in clinical nuclear medicine.

For parallel hole collimators, each collimator element allows the detection only of photons originating from a specific range of angles around the direction parallel to the collimator hole. Generally speaking, a narrow hole leads to a better resolution but also it leads to a decrease in the number of photons that reach the scintillation crystal, therefore lowering the detector sensitivity.

The detection sensitivity and the system resolution are both monotonically increased with the opening of the collimator aperture and decreased with the septa length of the collimator. Hence, there is no straightforward way to improve these two properties simultaneously by adjusting the collimator geometry, and a trade-off between resolution and sensitivity arises.

In the last few decades, optimisation of the collimator has essentially proceeded by trial and error: certain designs that have proven good image quality for certain imaging conditions, such as a certain energy range, have been designed and standardised [Keller, 1994]. Only in the last few years the problem has been treated systematically, driven by the need to compare and optimise collimator designs prospectively, by computer simulation. The developments of tools for the estimation of the uncertainty allows us to accelerate the design process by exploring the design space in software simulations. Optimisation of the design of the collimator is still an open and challenging problem and several evaluation strategies have been presented in the literature.

Task-specific Figures Of Merit (FOMs), based on the performance of human or mathematical observers in classification tasks, such as the detection of a certain class of tumours, have been defined and explored by Barrett et al. [1998]. Measures of performance on such tasks [Barrett and Myers, 2004] can be used to optimise collimator properties.
The ideal observer [Barrett and Myers, 2004] has been applied in the context of optimising pinhole SPECT in planar emission imaging. Other attempts to optimise apertures in SPECT have used approaches that differ from the ideal-observer method. For example, Zheng and Gullberg [2002] applied a human-performance emulating mathematical observer, a channelized hotelling observer, in the reconstruction domain, to optimise a parallel-hole collimator. In other studies more realistic tasks have been used. As an example, a joint detection and localisation task is medically more realistic than the pure detection task, since a physician must localise a lesion in an image. A general ideal observer for this task was proposed by Khurd and Gindi [2005]. In [Zhou and Gindi, 2008] this task has been applied for the optimisation of a parallel hole collimator.

Moreover, SPECT collimator designs can be evaluated based on the fundamental trade-off between bias and variance that can be achieved in the reconstruction of emission tomograms [Nuyts, 2009] [Vunckx et al., 2008a] [Vunckx et al., 2008b] [Zhou et al., 2010]. Such tradeoffs may be derived analytically using the Cramer-Rao type bounds [Hero et al., 1996] [Meng and Clinthorne, 2004] [Meng and Li, 2000] which imply the calculation and the inversion of the Fisher Information Matrix (FIM). In this chapter, the Fisher Information Matrix method will be used to characterise the uncertainty in the reconstruction for the purposes of parallel hole collimator optimisation.

Computing, storing and inverting the FIM is not practical for the typical matrix size of 3D SPECT imaging systems. In order to tackle the problem of the computational load in inverting the FIM, two approximations have been presented in chapter 4. The circulant approximation has been presented in section 4.5.1 and the subsampled FIM approximation has been presented in section 4.5.2.

The aim of this chapter is to explore the use of the subsampled FIM approximation for the optimisation of parallel hole collimators in SPECT; emphasising how it enables us to explore the design of highly shift invariant systems as a result of distance dependent resolution. The results obtained with the approximate subsampled FIM method are compared with the circulant approximation method and with the reference method based on reconstruction of multiple noise instances (see chapter 3).

Before the subsampled FIM approximation can be applied routinely to evaluate and optimise a collimator design (see appendix A), two sets of experiments are performed for validation purposes. The subsampled FIM and the circulant FIM methods are applied for the calculation of the variance of a NCAT phantom and a uniform phantom. The first experiment is performed simply to validate the deterministic method with a realistic phantom. The second experiment is performed to investigate the relation between the optimal collimator aperture and the target resolution.

5.1 Collimator Design Parameters

Nuclear medicine imaging collimators are characterised mainly by their imaging properties: geometric resolution and detection sensitivity. Although physical measurements are required for the determination of system response, collimators are affected by numerous secondary problems that are only indirectly related to the geometry. For example, the penetration of the radiation through the collimator septa, the visibility of the collimator hole pattern in the images, weight constraints imposed by the camera gantry, the minimal septal thickness imposed by the limitations of fabrication, and the effects of scattering within the collimator. If a collimator is properly designed, the secondary effects should not be noticeable. However, naive designs based exclusively on geometry without careful attention to the secondary problems may introduce either images that are significantly degraded or collimators that are not feasible for clinical use. Therefore collimator-detector system optimisation is a fundamental part of the design process. In this chapter we will concentrate on optimising the geometrical parameters of the design that affects the sensitivity and resolution. The secondary design consideration are briefly covered in this chapter since, in future investigations (see chapter 8), they will be taken into account for the design of new collimator geometries.

Parallel Hole (PH) collimator is the most widely used collimator in clinical nuclear medicine. The collimator usually consists of a two-dimensional array of parallel holes. The collimator allows only those photons nearly perpendicular to the detector surface to pass through, without being absorbed in the collimator material, and produces a planar image the same size as the source object. The variation between different parallel hole collimators is confined to the hole shape and to the geometric dimensions that describe the holes. Many hole shapes and patterns have been tested over the past few decades; these include: hexagonal, square, triangular and circular holes (a review of this can be found in Henkin et al. [1996])

The most important design parameters of a parallel hole collimator are: the length of the holes l; the diameter of the holes (or, for hexagonal holes, the perpendicular ”face-to-face” distance) v; the thickness of the septa q; and, finally, the material used to make the collimator.

In order to understand the relationship between these parameters and collimator properties, such as spatial resolution, geometric efficiency and penetration fraction, three types of events should be considered. The desirable, geometrically collimated gamma-rays traverse the collimator entirely within a hole, without contacting the septal material. Penetrating gamma-rays go through one or more septal walls without interacting, while scattered photons are deflected into the detector by a Compton interaction in a septum. The collimator penetration and scatter components are generally undesirable because their point of origin in the source is uncertain compared to their point of detection in the scintillation crystal. Large numbers of penetrating or scattering photons can contribute a substantial background to the image, thereby degrading the contrast of important image features. The spatial resolution, efficiency and penetration fraction of a parallel, multi-hole collimator were first discussed by Anger [1964].

5.1.1 Collimator Spatial Resolution

The spatial resolution of a parallel-hole collimator is expressed as the full width at half-maximum (FWHM) of the point spread function (PSF). The PSF for the parallel hole collimator is here described with the analytical depth-dependent model described by Anger [1964]. This model expresses the FWHM of the Gaussian at location x, y, z as:

$$FWHM(x, y, z) = \frac{v(CD(x, y, z) + l_2)}{l_2} \tag{5.1}$$

where $CD(x, y, z)$ is the distance between (x, y, z) and the detector plane, $l_2 = l - 2/\mu$ and μ is the total linear attenuation coefficient of the collimator material ($\mu = 22.7cm^{-1}$ for lead at a photon energy of $140keV$). This formula was modified from that presented by Anger [1964], by considering that the collimator septa length, l, should be reduced on both ends by approximately $1/\mu$ due to penetration effects [Jaszczak et al., 1986]. We use the acronym FWHM (without argument (x, y, z)) to denote the collimator aperture that corresponds to the centre of the image space.

It is clear from (5.1) that the spatial resolution of the collimator deteriorates with increasing distance from the collimator face. The resolution increases with reduction in hole diameter and increase in hole length but both these parameterisations reduce geometric efficiency. Also, it should be noted that the resolution changes less over a given range of collimator-source distances for larger values of collimator thickness. In other words, the slope of the collimator response (meaning the slope of the resolution as a function of distance line) is less for thicker collimators. It should be noted that the single scalar parameter in (5.1) does not offer a complete description of the collimator's PSF or its Fourier transform, whose detailed shape depends upon the shape of the collimator holes.

As already described in chapter 2 section 2.1.3, the overall resolution of a gamma-camera image is given by the 2-D convolution of the collimator PSF with the intrinsic detector's PSF. The intrinsic PSF is usually well approximated by a radially symmetric Gaussian function; with a FWHM of 3.6 mm is typical for a current conventional γ-camera. For the sake of simplicity, if we also approximate the collimator's PSF by a Gaussian function, then the combined system spatial resolution is approximately given by the following quadrature sum:

$$FWHM_{system} = \sqrt{FWHM^2_{collimator} + FWHM^2_{intrinsic}} \tag{5.2}$$

For most collimators used to date in nuclear medicine studies, the overall system resolution is dominated by the collimator resolution.

5.1.2 Geometric Efficiency

The geometric efficiency of a collimator system, E, is closely related to the collimator aperture and can be defined as the fraction of photons, emitted by a point source in the volume, that are detected in the detector bins if there is no attenuation. This efficiency is effectively independent of the source to collimator distance (under usual imaging conditions) and for the PH collimator with square holes, E is estimated as:

$$E = \frac{v^4}{4\pi l_2^2 (v + q)^2} \tag{5.3}$$

The geometric efficiency is then dependent on the FWHM of the PSF and is assumed to degrade approximately as the square of the geometric spatial resolution.

The collimator's efficiency determines the number of gamma-ray counts that may be recorded, for a given activity distribution, in a given scan time.

5.1.3 Septal penetration

The geometric response is further degraded owing to the penetration of some photons through the collimator septa; but no analytical treatment of this effect appears to exist in the literature.

Most collimator designers have used an ad-hoc "rule", allowing a certain small fraction of gamma-rays to penetrate along the minimum path length through a single septum (e.g. Keller [1994]), in order to account for penetration.

Several investigators have successfully used numerical ray-tracing methods [Muehllehner, 1973] [Beck and Redtung, 1985] [Han et al., 1999] or Monte Carlo simulation techiniques [Jan et al., 2004] [Cot et al., 2002] [Staelens et al., 2007] to examine the penetration component.

In appendix A a method for the investigation of new collimator design is introduced and in section A.2 the implementation of a ray-tracing algorithm is described. This algorithm has been used to model the geometric response and the septal penetration of novel collimator geometries.

As a general consideration for the design of the collimator, it should be noted that when the collimator hole spacing, $v+q$, becomes large compared to the intrinsic resolution of the gamma-camera, then the hole pattern can become visible in the projection images, therefore leading to artifacts [Newiger and Jordan, 1985]. System designers should take this into consideration, especially when designing collimators for high-energy isotopes which present high penetration factors and therefore need collimators with thicker septa.

5.1.4 Collimator Scatter

An analytic description of the collimator scatter component has eluded researchers. This effect is caused by photons that scatter in the collimator septa and still remain within the detection energy window.

A practical measurement of the effect of the collimator on gamma-camera performance, showed sensitivity changes on the order of 1%-2% due to scattering from the collimator [Moore et al., 1992]. Similar to septal penetration, collimator scatter can be evaluated using Monte Carlo simulation techniques.

5.2 Experiments

5.2.1 System Description

The SPECT system is based on a detector of size $236.16\ mm \times 236.16\ mm$. The detector rotates over 360° at a regular angular step of 2°. The axis of rotation is the line parallel to the detector surface through the centre of the image volume. The imaging volume dimensions are $96 \times 96 \times 12$ cubic voxels of $2.4\ mm$. Photon counts are binned on a grid of 96×12 pixels of $2.46\ mm$ and the intrinsic resolution is set to 3.6 mm. The distance $CD = 123\ mm$ between the centre of the image volume and the detector surface is constant during the tomographic acquisitions.

We consider a parallel hole collimator consisting of a two-dimensional array of square holes with septa thickness $q = 0.2\ mm$, hole diameter v, and length l. The diameter and the length of the hole are left unknown since they define the collimator aperture. The collimator aperture is characterised by the Point Spread Function (PSF). The PSF for the parallel hole collimator is described by the analytical depth-dependent Gaussian model in (5.1).

5.2.2 Statistical method based on Reconstruction of Multiple Noise Realisations

For the reference statistical method, the noisy data sets are reconstructed using an accelerated GPU implementation of the PL estimation, implemented as part of the NiftyRec toolbox [Pedemonte et al., 2010]. The depth-dependent Gaussian that characterises the collimator response is modelled both in the projector and in the backprojector operator. 10000 iterations are performed in order to maximise the cost function.

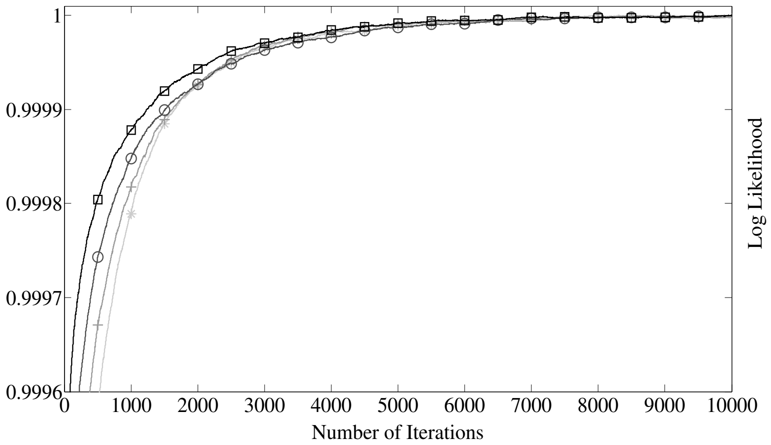

Figure 5.1: Objective function Ψ curves as a function of number of iterations. The different curves represent the objective function for different collimator apertures: $FWHM = 5.9\ mm$ to $FWHM = 11.08\ mm$ (from dark grey $\Box$ to light gray $*$). All the curves are scaled with respect to their maximum value.

A smoothing prior with a small weight $\beta = 10^{-12}$ is included in the cost function. The value of the regularisation parameter is chosen after trial and error, as a minimum value that nearly guarantees convergence within 10000 iterations. In Figure 5.1, the objective function $\Psi = L - \beta R$, for the digital phantom described in 5.2.5, is plotted as a function of the number of iterations. The different curves represent the objective function Ψ for different collimator apertures.

The calculation of the LLIR implies one reconstruction for each voxel i as in equation (3.11). A first reconstruction of the object from noiseless projection data is performed and iterated until convergence. The LLIR for each voxel i under investigation is then computed by performing the reconstruction (until convergence) of the noiseless projection of the object perturbed by an impulse in i.

In order to impose a fixed target bias gradient $\boldsymbol{P}_t$, the estimated image, obtained iterating the reconstruction algorithm to convergence, is convolved with an anisotropic filter $\boldsymbol{P}$. The post-smooth filter $\boldsymbol{P}$ has do be specifically designed for every system under investigation (i.e., for every collimator parameterisation) as in 4.6.

A series of independent noise realisations is computed using a pseudo-random Poisson noise generator. The variance is calculated as in 3.1 and is based on 10240 noise realisations. The reconstruction algorithm in use relies on the discrete model presented in 2.2, where both the image space and the detection space are digitised into voxels with finite size. However, since the variance is calculated for a post-smoothed image with a predefined target resolution; as long as the voxel size is sufficiently small compared to the target resolution (which is the case in our study), the effect of the discretisation on the variance is small and can be considered negligible.

5.2.3 Deterministic Method based on the Fisher Information

For the deterministic method, the subsampled version of the FIM has been calculated over three different grids of $g_1 = 27648$, $g_2 = 6912$ and $g_3 = 3072$ points equally distributed over the slice intersecting the point (or region of interest) of interest and the two neighbouring slices. It should be noted that grid g_1 is fully sampled over the three slices of interest and therefore, in the following, the respective FIM will be referred to as the full FIM.

The LLIR for a voxel of interest i is calculated as in equation (4.46) for the subsampled FIM and as in equation (4.42) for the circulant approximation of the FIM.
In order to impose a fixed target bias gradient $\boldsymbol{P}_t$, as for the reference statistical method, a post-smooth filter $\boldsymbol{P}$ has to be specifically designed for every collimator parameterisation (see section 4.6). The calculation of the covariance is then performed as in equation (4.47).

The variance images in Figures 5.2 (C-D-E) and 5.3 (C-D-E), are obtained by re-ordering the diagonal of the covariance matrix Cov calculated as in (4.47). As described in section 4.8, a trilinear interpolation is then performed in order to facilitate the visual comparison between the variance images obtained with the different grid models. Variance images obtained with the circulant approximation method are also presented for comparison. Every pixel of the images in Figures 5.2 (F) and 5.3 (F), are calculated according to equation (4.41).

5.2.4 Experiment A: NCAT Phantom

The first experiment is performed simply to validate the deterministic method with a realistic phantom. The phantom used for this experiment was a heart phantom (NCAT) [Segars, 2001]. The activity within the phantom was $\lambda = 8kBq/cm^3$ in the left and right ventricle myocardium, $\lambda = 3.4kBq/cm^3$ in the left and right ventricle chamber, $\lambda = 0.9kBq/cm^3$ in the lungs and $\lambda = 0.6kBq/cm^3$ in the background. The collimator hole diameter is $v = 1\ mm$, collimator hole length $l = 35\ mm$ and collimator aperture has $FWHM = 7.89\ mm$.

5.2.5 Experiment B: Collimator Aperture Optimisation

This experiment is performed to derive the relation between the optimal collimator aperture $FWHM_{opt}$ and the target resolution P_t for a uniform phantom. The CNR, defined in equation (4.50), is used as the figure of merit for collimator aperture optimisation and therefore the collimator aperture which corresponds to the maximal CNR will be considered as optimal.

The phantom used was a uniform sphere positioned at the centre of the image volume, with diameter $D = 39\ mm$. The activity concentration in the sphere was set to $\lambda = 8kBq/cm^3$ and in the background, to $\lambda = 2.4kBq/cm^3$. During the experiment, the collimator aperture varies from $FWHM = 5.9\ mm$ to $FWHM = 11.08\ mm$. The target resolutions are set to $FWHM_t = 12, 14$ and $16\ mm$.

The CNR is calculated for the central point of the sphere.

5.3 Results

In the following sections 5.3.1 and 5.3.2, the results for the two different digital phantoms and different collimator parameterisations are shown and discussed. The results are obtained with the presented subsampled approximation of the FIM and, for validation and comparison, with the circulant approximation of the FIM and with the reference statistical method. In section 5.4, all validation points are gathered to get a global overview of the agreement between the new method and the reference method.

5.3.1 Experiment A: NCAT Phantom

The calculated variance images for the NCAT phantom are shown in Fig. 5.2. Fig. 5.2-B shows the variance image obtained from the reconstruction of 10240 noisy projection data sets. Fig. 5.2 C-D-E show the corresponding images calculated with the full FIM for grid $g_1 = 27648$ and with the subsampled Fisher Information for grid $g_2 = 6912$ and $g_3 = 3072$ respectively. Fig. 5.2-F shows the variance image calculated with the circulant approximation of the FIM. Fig. 5.2-G shows the horizontal profiles.

From these images we can infer that both the method based on the subsampled Fisher Information Matrix and the method based on the circulant approximation of the FIM approximately predict the variance of the PL estimator, presenting minor, but obvious, differences with respect to the variance obtained with the reference method. The variance image obtained from the reconstruction of multiple noise realisations is rather noisy, due to the finite number of repeated experiments (10240). The variance images predicted with the FIM method, generally speaking, are smooth because neighbouring voxels are affected by similar levels of noise. However a sparser grid gives a more approximated estimation, as we can see in Fig. 5.2-F. In fact, a fundamental limitation of the subsampled FIM approach is that fine detail is being lost as the grid becomes more sparse. We noticed that, for such a complex phantom, performing the calculation with a grid $g_4 = 1728$ points or less, would lead to incorrect results.

Both the method based on the subsampled FIM and the method based on the circulant approximation of the FIM are somewhat less accurate near the edge of the finite support used in image reconstruction, for unknown reasons. This effect (which has been reported also in another study [Zhang-O'Connor and Fessler, 2007]) can lead to a discrepancy exceeding 10% and it is more noticeable, in case of low level of activity, in the off-center voxels of the phantom [Li et al., 2004].

5.3.2 Experiment B: Collimator Aperture Optimisation

The calculated variance images for a uniform sphere phantom at a collimator aperture $FWHM = 7.89\ mm$ are displayed in Figure 5.3. Figure 5.3-B shows the variance image obtained from the reconstruction of 10240 noisy projection data sets. Figure 5.3 C-D-E show the corresponding images calculated with the full FIM for grid $g_1 = 27648$ and with the subsampled Fisher Information for grid $g_2 = 6912$ and $g_3 = 3072$ respectively. Figure 5.3-F shows the variance image calculated with the circulant approximation of the FIM. Figure 5.3-G shows the horizontal profiles.

Once again, for this experiment, we can see how both the method based on the subsampled Fisher Information Matrix and the method based on the circulant approximation of the FIM approximately predict the variance of the PL estimator, presenting minor differences with respect to the variance obtained with the reference method.

These results show that the off-centre voxels have a lower variance than the central voxels of the phantom. This well-known observation [Alpert et al., 1982] is explained by the fact that, with a SPECT system, some of the planes through off-centre voxels have less intersection area with the phantom and are less multiplexed with neighbouring voxels than planes through the central voxels.

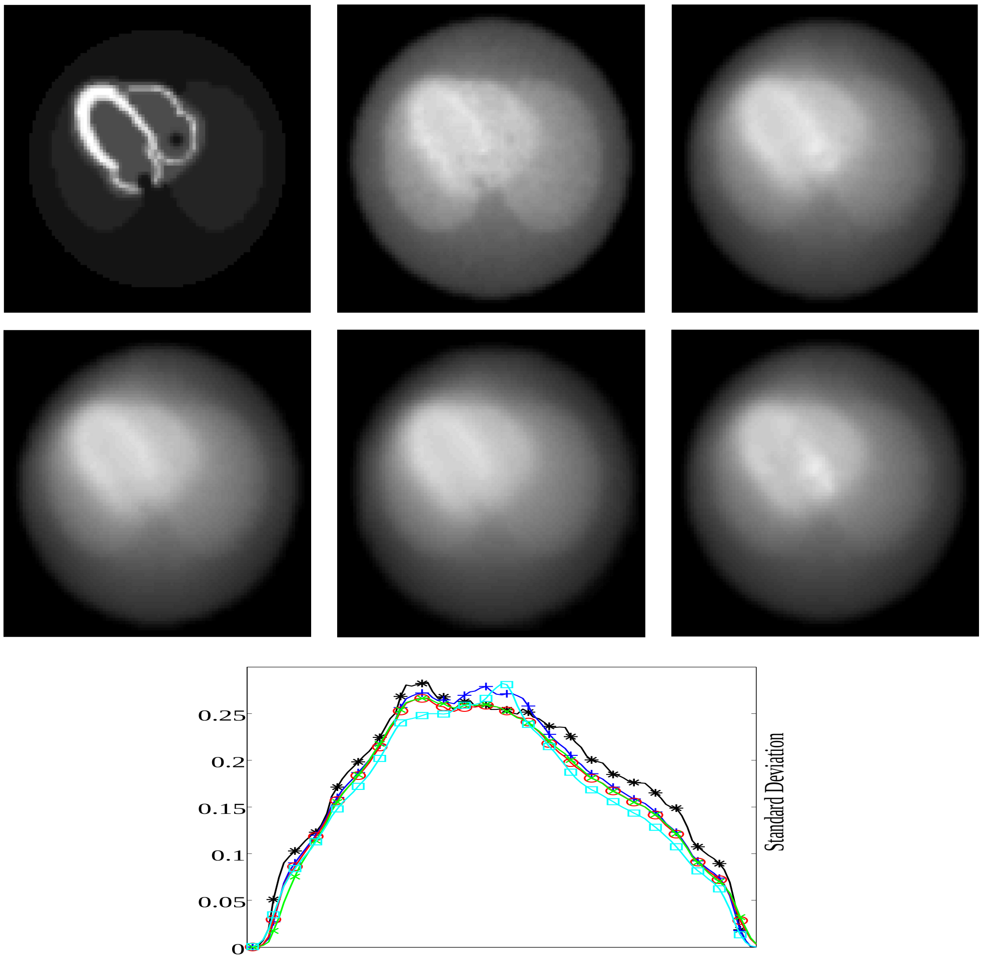

Figure 5.2: Variance images for the NCAT phantom obtained with a standard SPECT system. A (top left) - NCAT phantom. B (top central) - reference method (variance image obtained from the reconstruction of 10240 noisy projection data sets), C (top right) - Full Fisher Information-based method with grid g_1, D (central left) - Subsampled Fisher Information based method with grid g_2, E (central) - Subsampled Fisher Information based method with grid g_3, F (central right)- Variance image obtained with the circulant approximation method. G (bottom)- Variance profiles at the centre of the heart: reference method (black line - $*$), grid g_1 (blue line - $+$), grid g_2 (red line - $\circ$), grid g_3 (green line - $\times$), circulant (cyan line - $\square$).

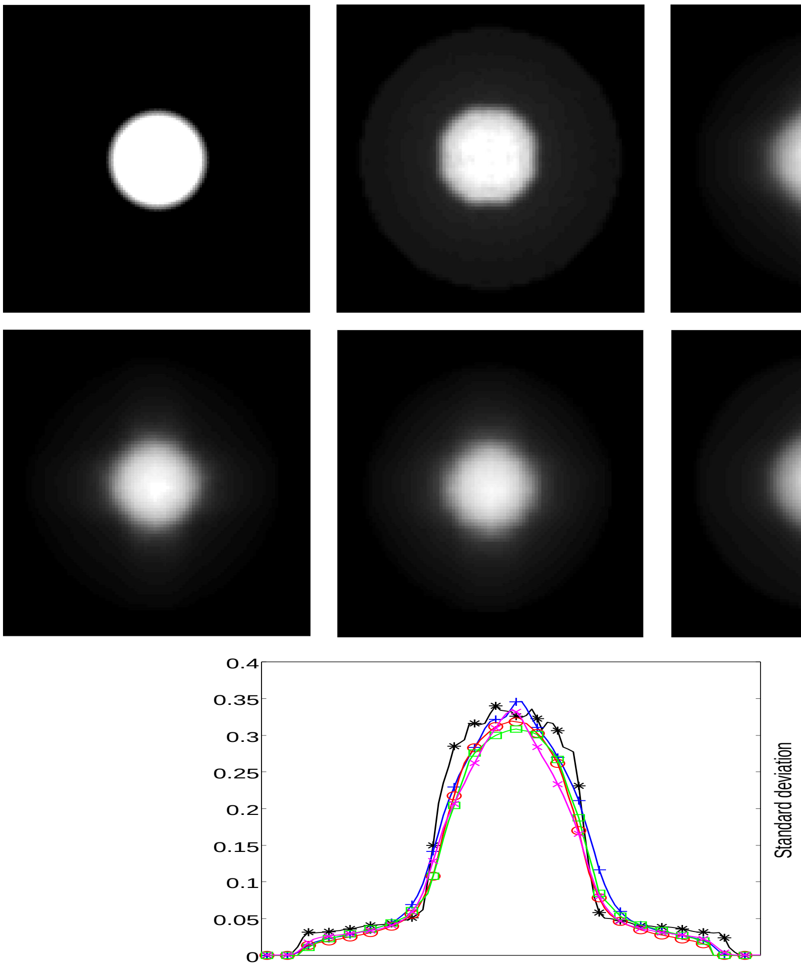

Figure 5.3: Variance images of a uniform sphere obtained with a standard SPECT system. A (top left) - Sphere phantom. B (top centre) - reference method (variance image obtained from the reconstruction of 10240 noisy projection data sets), C (top right) - Full Fisher Information-based method with grid g_1, D (central left) - Subsampled Fisher Information based method with grid g_2, E (central) - Subsampled Fisher Information based method with grid g_3, F (central right)- Variance image obtained with the circulant approximation method. G (bottom)- Variance profiles at the centre of the sphere: reference method (black line - $*$), grid g_1 (blue line - $+$), grid g_2 (red line - $\circ$), grid g_3 (green line - $\times$), circulant (cyan line - $\Box$).

Figure 5.4 shows the optimal collimator aperture for the central point of the sphere obtained with the reference method, with the full FIM (grid g_1), with the subsampled FIM method for grid g_2 and grid g_3, and with the method based on the circulant approximation. The optimal apertures are plotted for the three different target resolutions $P_t = 12\,mm$, $14\,mm$, and $16\,mm$. From the three plots, we can see how we obtain the same maximum CNR (so in turn, minimum variance) with the reference method, with the full FIM, with the two different subsamples of the FIM and with the circulant approximation method. We can see also how the optimal aperture varies almost linearly in relation with the target resolution imposed in the reconstruction. Similar results were presented in another study [Zhou et al., 2010] which claims that the FWHM of the parallel hole collimator aperture that yields the minimum variance equals, for volume imaging, the spatial resolution divided by $\sqrt{3}$.

However, from Figure 5.4 we can also notice that a sparser grid (i.e. grid g_3) gives a more approximated estimation of the CNR for small collimator apertures ($FWHM = 5.9$ to $7.8\,mm$) The minimum grid spacing is in fact determined by the size of the PSF, which implies that an higher spatial resolution requires more closely spaced grid points. It therefore seems that the subsampling approach would be less sensitive to the effects of spatial resolution than is the full FIM approach, for small collimator apertures, when the grid becomes sparse. The effects of changing spatial resolution on the CRB are reflected locally, and these effects will be lost because of the missing points between the grid points if the grid is too sparse.

The way the algorithm has been designed permits the degree of approximation in the estimation to be defined by the user. Therefore a trade-off between computational complexity and reliability of the estimation of the CNR arises. We noticed that, in this case, performing the calculation with a grid $g_4 = 1728$ points or less, would lead to incorrect results. The minimum number of grid points necessary to obtain a reliable estimation of the CNR depends on the characteristics of the system under investigation (for a discussion on the selection of the number of grid points, see section 4.7).

5.4 Validation

A validation for the subsampled Fisher Information-based variance calculation method is presented in the following.

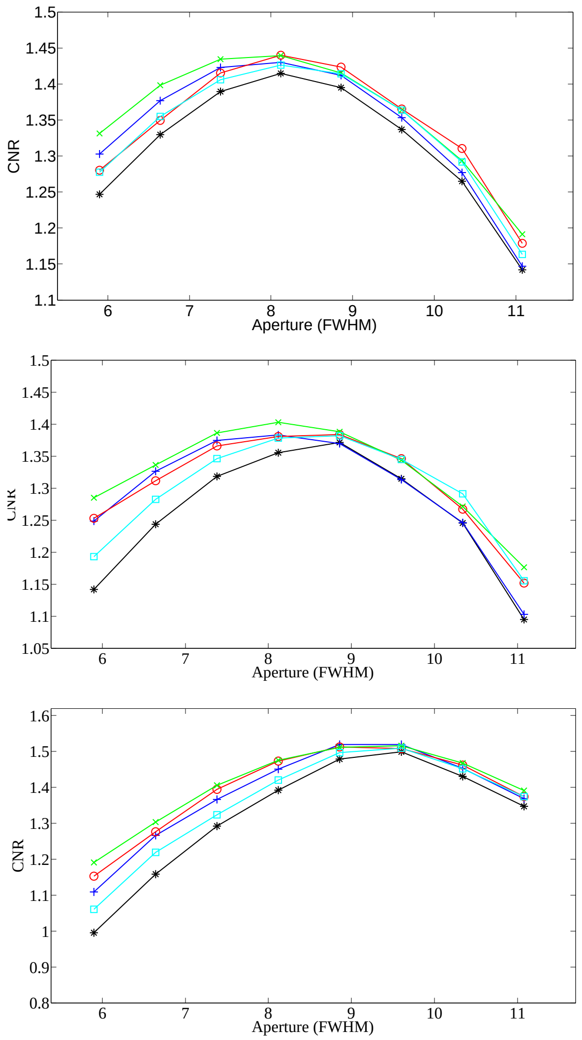

Figure 5.4: CNRs for different collimator apertures (from $FWHM = 5.9\ mm$ to $FWHM = 11.08\ mm$), obtained with the reference method (black line - $*$), with the method based on the Fisher Information with grid g_1 (blue line - $+$), with grid g_2 (red line - $\circ$), grid g_3 (green line - $\times$) and with the circulant approximation method (cyan line - $\Box$). Optimal collimator apertures are calculated for target resolutions $P_{target} = 12\ mm$ (A - top figure), $14\ mm$ (B - central figure) and $16\ mm$ (C - bottom figure).

Table 5.1: Validation of the Subsampled FIM for the NCAT phantom.

Ncat	CC	SEE	int	RC
g1	0.9933	2.03 e-04	-0.00413	0.97292
g2	0.9420	2.63 e-04	-0.00433	0.94042
g3	0.9392	3.04 e-04	-0.00454	0.93924
g4	0.9289	0.0046	-0.01069	0.91321
g5	0.9028	0.0061	-0.01803	0.89534

Validation of the Subsampled FIM in comparison with the Reference Statistical Method for the experiment presented in section 5.3.1. CC: Correlation Coefficent. SEE: Standard Error of the Estimate. int:intercept of the least squares fit. RC: regression coefficient of the least squares fit.

For the experiments presented in section 5.3.1 and 5.2.5, the variance obtained with the reference statistical method is plotted with respect to the variance predicted with the Fisher Information-based method, for grid $g_1 = 27648$ points, grid $g_2 = 6912$ points, grid $g_3 = 3072$ points, grid $g_4 = 1728$ points and grid $g_5 = 1106$ points. A least squares fitting is performed through the data. The regression coefficients, the intercepts of the line, the correlation coefficients and the standard error of the estimate are presented in Table 5.1 and Table 5.2. In the experiment presented in section 5.2.5, the variance is calculated for a range of collimator apertures, the validation results presented in Table 5.2 encompass the least squares fitting through all this data.

All validation points for the experiment in section 5.3.1 for those three grid models and for the circulant approximation method are shown in Figure 5.5. In these figures, the standard deviation calculated with the reference method with respect to the standard deviation predicted with the subsampled FIM method and with the circulant approximation method are plotted. The solid line was fitted to minimise the least squares distance between these points.

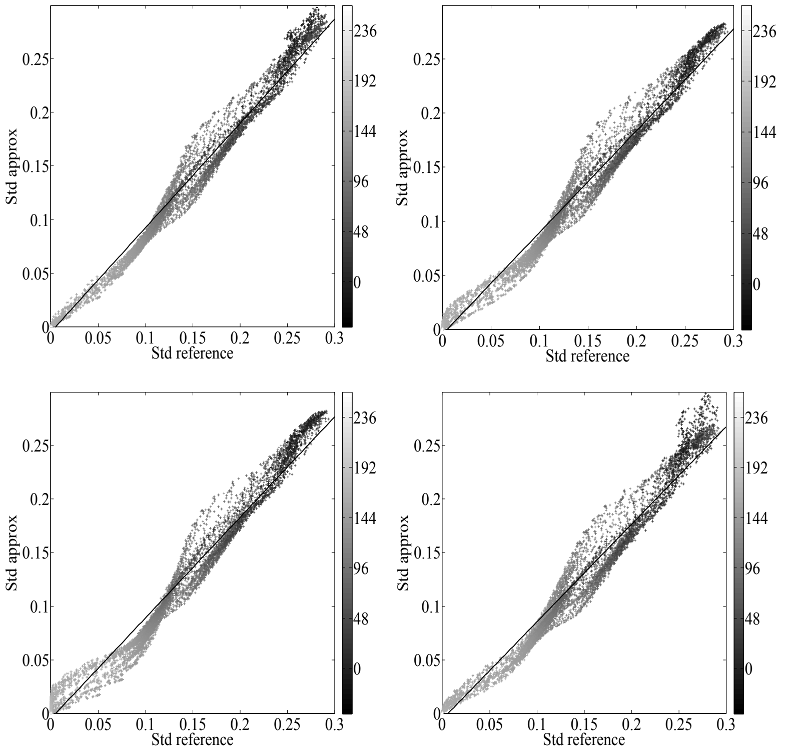

Figure 5.5: Validation of the Fisher information-based standard deviation calculation method for different grids for the experiment presented in section 5.3.1. The standard deviation obtained with 10240 repeated simulations is plotted with respect to the standard deviation predicted with the Fisher information-based method. A (top left)- grid g_1, B (top right) - grid g_2, C (bottom left) - grid g_3 and D (bottom right) - circulant approximation. Spatial information is given by the colour of the dots. The light grey dots represent voxels which are more distant from the rotation axis (off-centre voxels); whereas the dark grey dots represent voxels which are closer to the rotation axis (central voxels). A colour bar with the respective distance from the rotation axis (in mm) is displayed for every plot.

Table 5.2: Validation of the Subsampled FIM for the uniform phantom.

Sphere	CC	SEE	int	RC
g1	0.9871	1.62 e-04	-0.00158	0.97934
g2	0.9741	4.47 e-04	-0.00438	0.96172
g3	0.9470	5.72 e-04	-0.00868	0.93471
g4	0.9084	0.0018	-0.01422	0.89297
g5	0.8900	0.0025	-0.01896	0.82733

Validation of the Subsampled FIM in comparison with the Reference Statistical Method for the experiment presented in section 5.2.5. CC: Correlation Coefficent. SEE: Standard Error of the Estimate. int:intercept of the least squares fit. RC: regression coefficient of the least squares fit.

5.5 Discussion

Although it has not been thoroughly tested yet, the method for the quantification of the uncertainty in the estimate, presented in section 4.5.2, should also be valid for the investigation of other collimator configurations in SPECT (i.e. slit-slat or multi-pinhole collimators), given its capability to account for highly shift-invariant system responses. In addition this method can be applied to PET as well.

Our method differs from the methods presented in [Nuyts, 2009] by the use of post-smoothed PL instead of post-smoothed MLEM. With our method, a target resolution can be easily imposed, which has the advantage of enabling comparison of different designs at the same resolution and that only one parameter has to be considered for optimisation (the variance in the voxel of interest, or equivalently its CNR). At the same time, including a regularisation penalty in the cost function we guarantee the invertibility of $[\boldsymbol{F} - \beta \boldsymbol{R}]$ and therefore the uniqueness of the solution. Even if a small regularisation parameter is used with the only purpose of making the cost function strictly concave; the introduction of a penalty function $\boldsymbol{R}$, leads to shift-variant resolution properties of the estimator. This effect hinders the possibility of using a global figure of merit and therefore the uncertainty of the estimate needs to be calculated locally.

Given this consideration, it is important to discuss the computational performance of the subsampled FIM method with respect to the circulant approximation method. With the circulant approximation method, for each examined voxel, 1 projection, 1 backprojection, 1 Fourier transformation, 2 inverse Fourier transformations and 2 convolutions (for post-smoothing) are required. It is obvious that the efficiency is much higher than that of the reference method, if only a few points (up to several 100s) are under investigation. Some previously published methods, like [Qi and Leahy, 1999], go even further in their approximations. They try to isolate shift-invariant factors, like the geometric projection matrix for positron emission tomography (PET) imaging. Since it is the same for all voxels, it only has to be calculated once. This saves a lot of computation time. For a SPECT, however, it seemed impossible to distinguish and isolate shift-invariant factors. Another possibility is to isolate the matrices that are independent of the object, like the camera geometry matrix in [Qi and Leahy, 2000] [Stayman and Fessler, 2004b], such that it is sufficient to calculate them in advance. However, this is not useful for collimator design or for the examination of design parameters, since each camera geometry is only tested once.
The computation times for the calculation of the subsampled FIM and its inverse are presented in table 4.1. Thanks to an efficient GPU implementation, our novel algorithm for the estimation of the uncertainty, drastically reduces the computational complexity in calculating the FIM and its inverse, making the algorithm a good candidate for fast design optimisation. Even if the proposed approximation implies the calculation and the inversion of the FIM only once for each set of system parameters; it should be noted that, for the optimisation of the collimator design at a fixed target resolution, 2 convolutions are also required for each voxel under examination.

From the experiment presented in section 5.3.2, it can be seen that the subsampled FIM method and the circular FIM method give very similar results for the central point of a uniform sphere. It has been found that the optimal collimator aperture is proportional to the target resolution that is imposed in the reconstruction. Similar results were presented in a study of pinhole imaging system [Fessler, 1998], which claims that, in order to minimise the variance of the emission-rate density estimate at a particular spatial resolution, the pinhole size should be proportional to that resolution. Furthermore, another study [Zhou et al., 2010] shows that for the PH collimator, the FWHM of the aperture that yields the minimal variance equals, for volume imaging, the desired spatial resolution divided by $\sqrt{3}$.

From the results shown in section 5.3.2, we can also notice that a sparser grid gives a more approximated estimation of the CNR for small collimator apertures. It therefore seems that the subsampling approach would be less sensitive to the effects of spatial resolution than is the full FIM approach, for small collimator apertures, when the grid becomes too sparse.

Evaluation of the effect of subsampling is therefore complicated by the trade-off that arises between the subsampling and the reliability of the variance estimation. A general criterion to define a relationship between the subsampling and the reliability of the variance estimation, can not be provided. This criterion depends in fact on the properties of the specific system.

Chapter 6

Region Of Interest Reconstruction

6.1 Introduction

In single-photon emission computed tomography (SPECT), measurement projection data can be truncated when the camera's field of view is smaller than the object to be imaged.

In the last few years, several companies have developed small dedicated cardiac SPECT systems with different designs. These dedicated systems have a smaller field of view (FOV) than a full-size clinical system. Thus, data truncation has become the norm rather than the exception in these systems. If small detectors are used, truncated projections must be used to reconstruct a region of interest (ROI) within the boundary of the object.

The Region-Of-Interest (ROI) reconstruction using truncated projections data, may be also referred to as the interior problem. This problem does not have an analytical, closed form, solution [Natterer, 1986]. This mathematical conclusion is, in general, correct; however, many researchers are seeking solutions to the interior problem if some constraints or some additional information about the object to be imaged can be enforced during reconstruction [Defrise et al., 2006] [Clackdoyle et al., 2004] [Clackdoyle and Noo, 2004] [Zou et al., 2005].
These studies results in several sufficient conditions for exact ROI reconstruction, where incomplete projection can be used to uniquely and reliably determine the unknown image of a region of the object. One of the limitations of analytical ROI reconstruction methods is that they are specifically designed for one type of truncation and for one type of imaging geometry. An inversion formula developed for a particular type of truncation situation often can not be applied to other cases. So far, there is no ROI reconstruction theory that can handle all the truncation cases and imaging geometries in a unified framework.

In contrast, statistical reconstruction methods (see 2.5.1), which are based on discrete models, are more flexible in handling truncated projection data. If some data in the projections are missing due to truncation effects, the corresponding rows in the system matrix are simply set to zero, and no further modification is required. Therefore, iterative reconstruction methods entail reconstruction of the whole object regardless of the size of the ROI, while analytical methods are capable of produce an image only of voxels inside an ROI.

Data truncation results in an underdetermined system of imaging equations, which may lead to non-unique solutions. In order to study the uniqueness of the solution, the concept of singular value decomposition is introduced and is used to evaluate the characteristics of the system operator.

[Zhang and Zeng, 2007] demonstrated that generic common statistical reconstruction algorithms are able to exactly reconstruct an ROI, under the conditions that the convex ROI is fully sampled and a single image value in a sub-region within the ROI is known. If the ROI includes a sub-region that is outside the patient body, then these conditions can be easily satisfied. We must point out, however, that an “exact” reconstruction is only possible in a theoretically ideal situation where data are noiseless, sampled on the detector with an infinitesimal sampling interval, and sampled angularly with an infinitesimal angular interval.

Given these considerations, it is necessary to study the noise properties of statistical reconstruction methods in the presence of truncated projection data. In this chapter, simulation experiments are designed to investigate the statistical properties of the estimator under two different truncation cases. In the first case, the ROI includes peripheral regions that satisfy the data sufficiency condition [Defrise et al., 2006]. In the second case, an internal ROI that does not qualify for exact and stable reconstruction [Natterer, 1986] is considered. For each case, we simulate different amounts of missing data.

A PL estimation algorithm has been used to reconstruct the two types of ROI, namely, peripheral ROI and interior ROI. One potential benefit of statistical ROI reconstruction over analytical ROI reconstruction formule is that the PL estimation algorithm can handle the two classes of problems in a unified framework, and no prior information about the ROI is needed. Moreover, statistical reconstruction methods account for measurement noise in the problem formulation.

In this chapter, it will be shown that a peripheral ROI, which satisfies the analytical data sufficiency condition, can be accurately reconstructed using the PL estimation algorithm, while the interior ROI reconstruction results in images that suffer from artifacts. These results are consistent with corresponding analytical investigations.

In order to study the noise properties of the estimator under the two different truncation cases, the statistical method (see chapter 3.1), based on the reconstruction of multiple noise instances, and the deterministic FIM-based method have been used for the calculation of the uncertainty. For the deterministic method, a comparison of the variance obtained from the full FIM, from the subsampled FIM and from the circulant FIM method is performed. One main contribution of this chapter is that it exposes certain pitfalls of the circulant approximation, when the effects of data truncation are investigated.

6.2 Singular Value Decomposition

The figures of merit for the calculation of the image quality discussed in the preceding chapters have focused on the statistical properties of the estimator used for reconstruction. The system operator $\boldsymbol{H}$ itself, can also provide information that is useful for investigating the stability of an ROI solution. In this section, a framework for analysing the system operator, using the tool of singular value decomposition (SVD) , is presented.

The singular value decomposition (SVD) method has been used to study the 2D interior problem of the Radon transform by Maass [1992]; while others used numerical evaluation of the SVD to study the performance of medical imaging systems [Zeng and Gullberg, 1997] [Jorgensen and Zeng, 2008] [Clarkson et al., 2010] and to study the interior problem [Zeng and Gullberg, 2012].

Ideally, the system matrix $\boldsymbol{H}$ would be invertible. However, this condition is quite difficult to verify and it is generally found that this matrix is singular; even when the system operator only approximates the true imaging system. One way to explore the singularity of $\boldsymbol{H}$ is to use the singular-value decomposition (SVD) theory.
The SVD of $\boldsymbol{H}'\boldsymbol{H}$ is given by

$$\boldsymbol{H}'\boldsymbol{H} = \boldsymbol{V}\boldsymbol{\Sigma}\boldsymbol{V}' \tag{6.1}$$

The $N \times N$ matrix $\boldsymbol{V}$ is a non-singular orthonormal matrix with each row of $\boldsymbol{V}$ being a singular vector of $\boldsymbol{H}$. $\boldsymbol{\sigma} = D\left[\sigma_1, \sigma_2, \ldots, \sigma_t, 0, \cdot, 0\right]$ is a diagonal matrix whose values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_t > 0$ are the singular values of $\boldsymbol{H}'\boldsymbol{H}$, or equivalently, the square of the singular values of $\boldsymbol{H}$.

The SVD provides a condition number to diagnose the system of linear equations. The condition number is defined as the ratio of the largest singular value and the smallest singular value of $\boldsymbol{H}'\boldsymbol{H}$. The condition number indicates how sensitive the solution is to the noise in the projection data $\boldsymbol{\nu}$. A system of linear equations with a low condition number is said to be well-conditioned, while a system with a high condition number is said to be ill-conditioned.

An ill-conditioned system of equations presents some small or zero singular values. Singular vectors with very small or zero singular values can hardly be measured by the imaging system.

Since missing data increases the ill-conditioning of the system matrix; many singular vectors often have very small singular values if the projection data are truncated or incomplete.

These singular vectors form a set of basis of the null space of the system matrix:

$$\operatorname{null}(\boldsymbol{H}) = \operatorname{span}\left\{\boldsymbol{v}_{t+1}, \boldsymbol{v}_{t+2}, \cdot, \boldsymbol{v}_N\right\} \tag{6.2}$$

where t is a threshold index chosen such that σ_j's values drop abruptly to zero for $j > t$ on the singular value spectrum of the system matrix.

6.3 Experiments

6.3.1 System Description

The SPECT system is based on a detector of size $236.16\ mm \times 236.16\ mm$. The detector rotates over 360° at a regular angular step of 2°. The axis of rotation is the line parallel to the detector surface through the centre of the image volume. The imaging volume dimensions are $96 \times 96 \times 12$ cubic voxels of $2.4\ mm$. Photon counts are binned on a grid of 96×12 pixels of $2.46\ mm$ and the intrinsic response is set to 3.6 mm. The distance $CD = 123\ mm$ between the centre of the image volume and the detector surface is constant during the tomographic acquisitions.

We consider a parallel hole collimator consisting of a two-dimensional array of square holes with septa thickness $q = 0.2\ mm$, hole diameter $v = 1.2\ mm$, and length $l = 35\ mm$. The collimator has a linearly varying depth-dependent Gaussian response that has a slope of 0.0562, which corresponds to 7.89 mm FWHM at the centre of the field of view.

In this experiment we investigate the effect on image variance for region-of-interest reconstruction from truncated projection data. Truncation is caused by a limited detector size. Only a certain number of detector bins u are used to measure data. Note that the Field Of View in this truncation situation forms cylinder of specific radius depending on the level of truncation. During the experiment, the ROI diameter varies from $u = 96$ to $u = 16$ (from $236.16\ mm$ to $39.36\ mm$). Truncation is simply modelled by setting to zero the rows of the system matrix $\boldsymbol{H}$, that corresponds to the missing projection data ν_d (due to truncation).

6.3.2 Singular Value Decomposition for ROI Acquisition

The singular value decomposition (SVD) method has been used to study the properties of the imaging system in case of truncation of the projection measurements.

The SVD of $\boldsymbol{H}'\boldsymbol{H}$ has been calculated as in equation (6.1).

It should be noted that $\boldsymbol{H}'\boldsymbol{H}$, which is here in matrix form, can also be formulated as:

$$\text{Crosstalk} = \boldsymbol{H}'\boldsymbol{H} = \sum_{d}^{M} h_{di}h_{dj} \tag{6.3}$$

This matrix will be referred to, in the following, as the crosstalk matrix (see Barret [Barrett et al., 1995]) and it corresponds to the nominator of equation (4.29), that defines the Fisher Information Matrix. It is therefore straightforward to calculate the crosstalk matrix, using the purpose-made GPU accelerated software for the calculation of the FIM described in 4.5.3, by simply imposing $\overline{\nu}_d = 1$ in the denominator of equation (4.29).

In section 4.5.3, a method to sparsify the FIM has been introduced. This sparsified matrix was referred to as the subsampled Fisher Information Matrix. In the same way, it is possible to calculate the crosstalk matrix in equation (6.3) over a subset of the voxel indexes $\boldsymbol{G} \subset [1, \ldots, N]$ arranged in a grid that covers the whole volume. The subsampled version of the crosstalk matrix is calculated over three different grids of $g_1 = 27648$, $g_2 = 6912$ and $g_3 = 3072$ points equally distributed over the slice intersecting the point (or ROI) of interest and the two neighbouring slices. It should be noted that grid g_1 is fully sampled over the three slices of interest and therefore, in the following, it will be referred to as the full crosstalk matrix; whereas grid g_2 and g_3 will be referred to as subsampled crosstalk matrices.

The singular value decomposition of $\boldsymbol{H}'\boldsymbol{H}$ has been calculated for the different levels of truncation. The SVD is calculated using MATLAB's built-in function. The singular value spectra for the full and the subsampled crosstalk matrices are presented, and the condition numbers for different levels of truncation are calculated to give a scalar measure for the accuracy of the solution.

6.4 A Data Sufficiency Condition for The Interior Problem

As described in the previous sections, in many applications of Emission Computed Tomography there is only a limited region of the object to be imaged that is of interest. It is desirable to identify the smallest set of line integrals required to accurately reconstruct this ROI. However, a general criterion to determine whether a given family of line integrals is sufficient for exact and stable reconstruction of a ROI has, to date, eluded researchers.

A data sufficiency condition for region of interest reconstruction from truncated projection data has recently been introduced in [Defrise et al., 2006]. This study shows that analytical reconstruction algorithms are able to exactly reconstruct the ROI under the conditions that the convex ROI is fully sampled and the image value in a sub-region within the ROI is known. If the ROI includes a sub-region that is outside the patient body, then the conditions can be easily satisfied. Although the derivation of Defrise's algorithm [Defrise et al., 2006] is based on the concept of differentiated backprojection and finite Hilbert transform inversion [Noo et al., 2004] [Pan et al., 2005] [Sidky and Pan, 2005]; applying these tools is not mandatory when effectively reconstructing the ROI, especially since a closed form analytic inversion formula for the truncated Hilbert transform has not been derived. A statistical reconstruction method can be applied to the truncated projection data set, taking into account the entire support of the object and any additional prior information.

Zhang and Zeng [2007] investigated the use of statistical iterative algorithms for the ROI reconstruction from truncated projection data. However this study considers only bias in the reconstruction from noiseless data, disregarding noise. Evaluation of the effect of noise in case of truncated emission data is one of the contributions of this chapter.

Experiment A - A Subregion Within the ROI is Known

In this experiment, the statistical properties of the estimator, for ROI reconstruction from truncated projection data, are investigated in the case that the ROI satisfies the data sufficiency conditions described above.

Figure 6.1-A shows the software phantom where the sphere in the centre of the FOV represents the heart and the off-centre sphere represents the liver. The intensity of the heart and that of the liver are equal and are set to $\lambda = 8kBq/cm^3$, whereas the activity in the background is set to zero. The Field-Of-View in this truncation situation forms cylinders of different radii depending on the level of truncation. The size and the position of the central sphere has been chosen in order to have the sphere always in the ROI, whereas the off-centre sphere is gradually more and more outside the ROI with increasing truncation level. Since the central sphere is surrounded by a zero background and the sphere is entirely in the ROI, for any level of truncation, the data sufficiency condition is satisfied.

Experiment B - Interior Problem

In this experiment, the statistical properties of the estimator are investigated in case the ROI to be reconstructed is entirely inside the boundary of the object. This problem is typically defined as the interior problem and is not uniquely solvable.

The phantom, shown in Figure 6.1 - B is a uniform sphere positioned in the centre of the image space, with diameter $D = 24.6\ mm$. The activity in the sphere is set to $\lambda = 8kBq/cm^3$ and to $\lambda = 4kBq/cm^3$ in the background. The size and the position of the sphere has been chosen in order to have the sphere always in the ROI, whereas the activity in the background is gradually more and more outside the ROI with increasing truncation level.

6.4.1 Statistical Reconstruction and Statistical Calculation of the Uncertainty

In order to reconstruct the truncated projection data, an iterative PL reconstruction with an added quadratic regularisation penalty has been used. The noisy data sets are reconstructed to convergence, using an accelerated GPU implementation of the algorithm for PL estimation, implemented as part of the NiftyRec toolbox [Pedemonte et al., 2010]. It should be noted that, for ROI reconstruction, the matrix that samples the object space must be large enough to contain the whole support of the object even if only a region of interest is to be reconstructed.

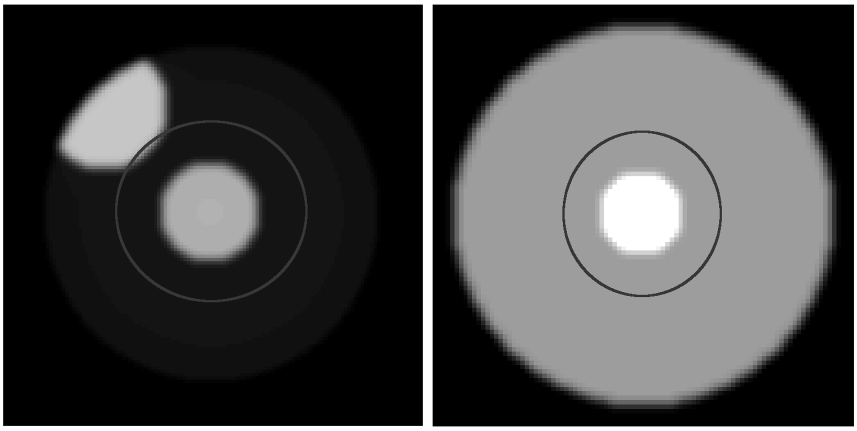

Figure 6.1: The interior problem in emission tomography. The red circle represents the ROI for the maximum level of truncation. A (left) - The image value in a sub-region within the ROI is known. An exact and stable reconstruction can be achieved. B (right) - The ROI to be reconstructed is entirely inside the boubdary of the object. For this measurement configuration, the ROI can not be reconstructed exactly (the solution is not unique)

10000 iterations are performed in order to maximise the cost function. A smoothing prior with a weight $\beta = 10^{-8}$ is included in the cost function. The value of the regularisation parameter is chosen after trial and error, as a minimum value that nearly guarantees convergence within 10000 iterations. In Figure 6.2 the objective function $\Psi = L - \beta R$, for the digital phantom described in 6.4 - A, is plotted as a function of the number of iterations. The different curves represent the objective function for different levels of truncation.

A series of independent noise realisations is computed using a pseudo-random Poisson noise generator. The mean and the variance are calculated as in equation (3.3) and (3.6) respectively, for both the experiments presented in 6.4 - A and 6.4 - B. Moreover the covariance is calculated for a voxel in the centre of imaging volume as in equation (3.8).

The number of noise instances is 10240. NiftyRec can process concurrently up to 1024 reconstructions in order to make efficient use of the GPU and 10 repetitions are chosen in order to obtain satisfactory images of variance.

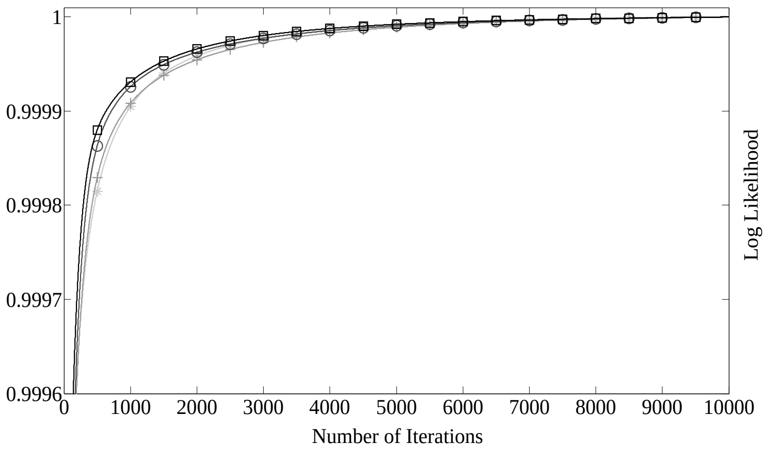

Figure 6.2: Objective functions Ψ curves as a function of number of iterations. The different curves represent the objective functions for different levels of truncation. Truncation is caused by a limited detector size. The ROI diameter varies from $u = 96$ to $u = 36$ (from dark grey $\square$ to light grey $*$). All the curves are scaled with respect to their maximum value.

6.4.2 Deterministic Method based on the Fisher Information

The main contribution of this chapter is given by the evaluation of noise properties of the system, in case of truncated projection data. The effect on the variance of the estimate, for region-of-interest reconstruction from truncated projection data, is evaluated using the deterministic derivation based on the Fisher Information matrix.

For the deterministic method, the subsampled version of the FIM has been calculated over three different grids of $g_1 = 27648$, $g_2 = 6912$ and $g_3 = 3072$ points equally distributed over the slice intersecting the point (or ROI) of interest and the two neighbouring slices.

The calculation of the covariance is then performed as in equation (4.47) for different levels of truncation. The variance images in Figures 6.7 (B-C-D) are obtained by re-ordering the diagonal of the covariance matrix Cov calculated as in (4.47). As described in section 4.8, a trilinear interpolation is then performed in order to facilitate the visual comparison between the variance images obtained with the different grid models. Variance images obtained with the circulant approximation method are also presented for comparison. Every pixel of the image in Figure 6.7 (E), is calculated according to equation (4.41).

The variance for a plane intersecting a point in the centre of the FOV and the CNR for a voxel positioned in the centre of the sphere are calculated for both the phantoms presented in 6.4 - A and 6.4 - B.

6.5 Results

In the following sections 6.5.2 and 6.5.3, results from ROI reconstruction for two different truncation cases (namely, the peripheral ROI and the interior ROI), are shown and discussed. The results are obtained with the presented subsampled approximation of the FIM and, for validation and comparison, with the circulant approximation of the FIM and with the reference statistical method. In section 6.5.1 results obtained with the singular value decomposition (SVD), to study the properties of the imaging system in case of truncation of the projection measurements, are discussed. In section 6.5.4, all validation points are gathered to get a global overview of the agreement between the new method and the reference method. In section 6.6 a visual representation of the Fisher Information matrices for the different approximation methods, is shown for comparison.

6.5.1 Singular Value Decomposition for ROI Acquisition

Figure 6.3 - A shows the singular value spectra of the system operator $\boldsymbol{H}$ with different levels of truncation. Truncation is caused by a limited detector size. The different curves show the SVD spectra for ROI diameter varying from $u = 96$ to $u = 16$ (from $236.16\ mm$ to $39.36\ mm$). For increasing level of truncation, as the ROI becomes smaller, an increasing number of singular vectors present a singular value that is nearly zero, indicating that the system matrix becomes more ill- conditioned.

Figure 6.3 - A shows the singular value spectra for the full crosstalk matrix with grid $g_1 = 27648$, whereas Figure 6.3 - B shows the singular value spectra for the subsampled crosstalk matrix with grid $g_2 = 6912$. Comparing the two figures, it is easy to notice that calculating the SVD for the subsampled crosstalk matrix is equivalent to performing a subsampling of the singular value spectrum. The singular values for every singular vector obtained from the subsampled crosstalk matrix are equivalent to the singular values obtained from the full crosstalk matrix for the respective singular vectors. Therefore, the points of inflection of the curves are equivalent with both methods. It should be noted that, for representation purposes, the vertical axis in the plots of Figure 6.3 present the same scale. Thus, the maximum values of

Table 6.1: Condition numbers for the interior problem

Svd Interior	u = 96	u = 76	u = 56	u = 36
g1	1.81 e+05	4.71 e+24	8.13 e+25	4.74 e+26
g2	4.29 e+03	2.15 e+23	6.89 e+23	1.02 e+25
g3	1.05 e+03	7.50 e+22	8.27 e+23	1.56 e+25
g4	4.23 e+02	2.37 e+22	1.69 e+23	2.44 e+23

Condition numbers of the full crosstalk matrix with grid $g_1 = 27648$ points and of the subsampled crosstalk matrices with grid $g_2 = 6912$ points, grid $g_3 = 3072$ points and grid $g_4 = 1728$ points. The condition number are presented for each grid and for different levels of truncation ranging from $u = 96$ to $u = 16$.

Figure 6.3 - A are not displayed.

The condition number provides a scalar measure to evaluate the system of linear equations. The condition number is defined as the ratio of the largest singular value and the smallest singular value of $\boldsymbol{H}'\boldsymbol{H}$. The condition numbers of the full crosstalk matrices and of the subsampled crosstalk matrices for different levels of truncation are presented in table 6.1. The absolute values of the condition numbers for the two methods are obviously different since subsampling is performed on the singular value spectra and the maximum singular values for the full crosstalk matrices are most likely to differ from the maximum singular values for the subsampled case.

6.5.2 A - A Subregion Within the ROI is Known

Figures 6.4 and 6.5 show images of the mean, the variance and the covariance calculated with the statistical reference method based on the reconstruction of 10240 noise realisations, for the non-truncated and the truncated cases respectively.

From these images we can see that ROI reconstruction from truncated projection data can lead to nearly unbiased reconstruction in a well-sampled ROI, under the condition in which the image value in a sub-region within the ROI is known.

Images of the mean for the non-truncated and truncated case are presented in Figures 6.4-A and 6.5-A. These images are congruent and present good accuracy throughout the circular ROI.

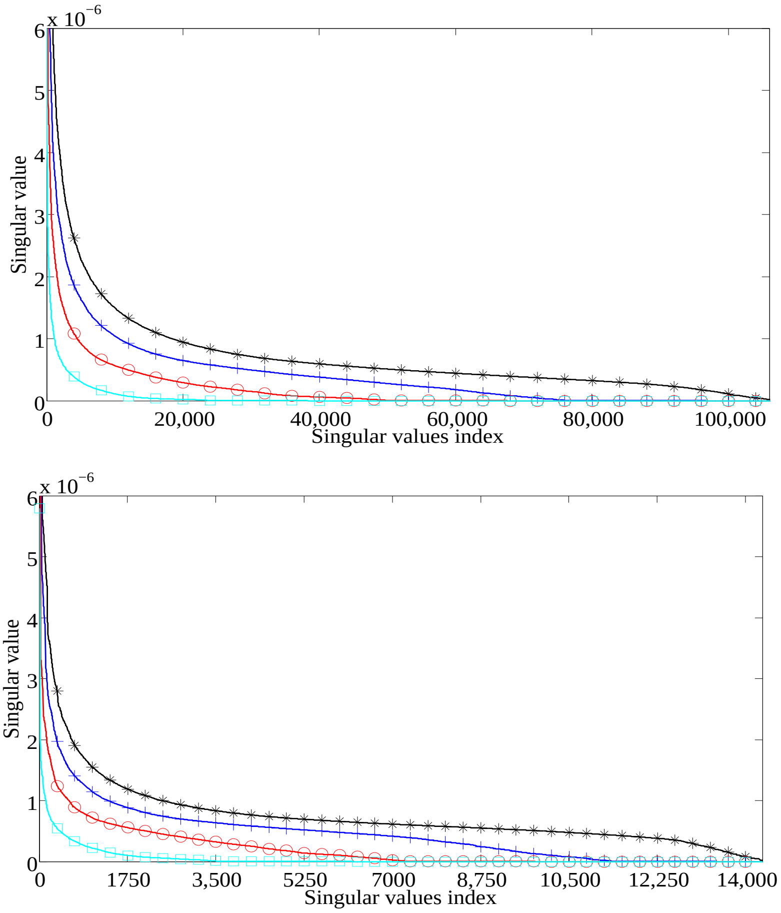

Figure 6.3: SVD spectra for the interior problem. A (top) - Singular value spectra for the full crosstalk matrix with grid $g_1 = 27648$. B (bottom) - Singular value spectra for the subsampled crosstalk matrix with grid $g_2 = 6912$. Truncation is caused by a limited detector size. The different curves show the SVD spectra for ROI diameter varying from $u = 96$ to $u = 36$ (from $236.16\ mm$ to $88.56\ mm$). Black ($*$) - ROI diameter $u = 96$, no truncation is performed. Blue ($+$) - ROI diameter $u = 76$. Red ($\circ$) - ROI diameter $u = 56$. Cyan ($\Box$) - ROI diameter $u = 36$.

The absolute value of the difference between the non-truncated and the truncated cases does not exceed the 8.2% of the maximum intensity and the average value of this absolute difference is the 2.3% of the maximum intensity.
Outside the ROI, where projections are truncated and data are available only from certain camera positions, the reconstruction algorithm overestimates the amount of activity in that region.

Figures 6.4-B and 6.5-B report the variance (standard deviation) and Figures 6.4-F and 6.5-F the covariance for a point in the centre of the FOV. In the non-truncated case the off-centre voxels have a lower variance than the central voxels of the phantom, whereas, for the truncated case, the variance is higher in the sphere outside the ROI. Even if ROI reconstruction from truncated projection data can lead to nearly unbiased reconstruction, in a well-sampled ROI, when the reconstruction algorithm is iterated to convergence; we noticed that a decrease in ROI size leads to an increase in variance, not only outside the ROI but also inside it. For this specific experiment we observe an overall increase in standard deviation of 6% in the centre of the sphere positioned inside the ROI. Moreover from the covariance profiles, we can observe a change in the statistical dependence between the estimate of the activity outside the ROI and the estimate of the activity inside the ROI, in case of truncation of the projection data. For the truncated case, in fact, the covariance does not have a localised content and therefore the uncertainty in a point in the centre of the FOV depends on the estimate of the activity outside the ROI.

6.5.3 B - Interior Problem

The mean and the variance images obtained with the reference statistical method, from truncated projection data with a ROI diameter $u = 36$, for a uniform sphere phantom and uniform non-zero background, are presented in Figure 6.6 - B and Figure 6.6 - C. The digital phantom used for this experiment is shown in Figure 6.6 - A. Figure 6.6 - D shows the horizontal profiles for the digital phantom and for the mean image. From these images we can see how, iterating the reconstruction algorithm to convergence (performing 10000 iterations), we obtain an estimate of the true unknown object $\boldsymbol{\lambda}$ with severe artifacts near the borders of the ROI. For this measurement configuration, where the ROI to be reconstructed is entirely inside the boundary of the object, the image inside this region can not be reconstructed exactly (since the solution is not unique) and the estimate is therefore biased.

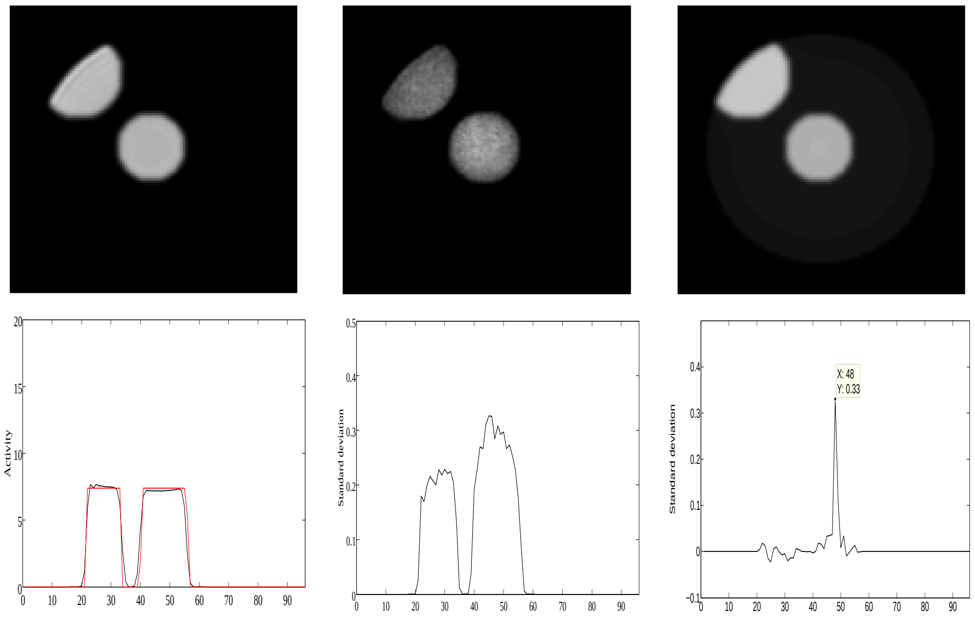

Figure 6.4: Non-truncated case. A (top left) - average image. B (top centre) - standard deviation image. C (top right) - activity phantom. D (bottom left) Black - diagonal profile of the average image. Red - diagonal profile of the phantom. E (bottom centre) - diagonal profile of the standard deviation image. F (bottom right) - diagonal profile of the covariance for a point at the centre of the FOV. The standard deviation in the point in the centre of the FOV equals 0.332.

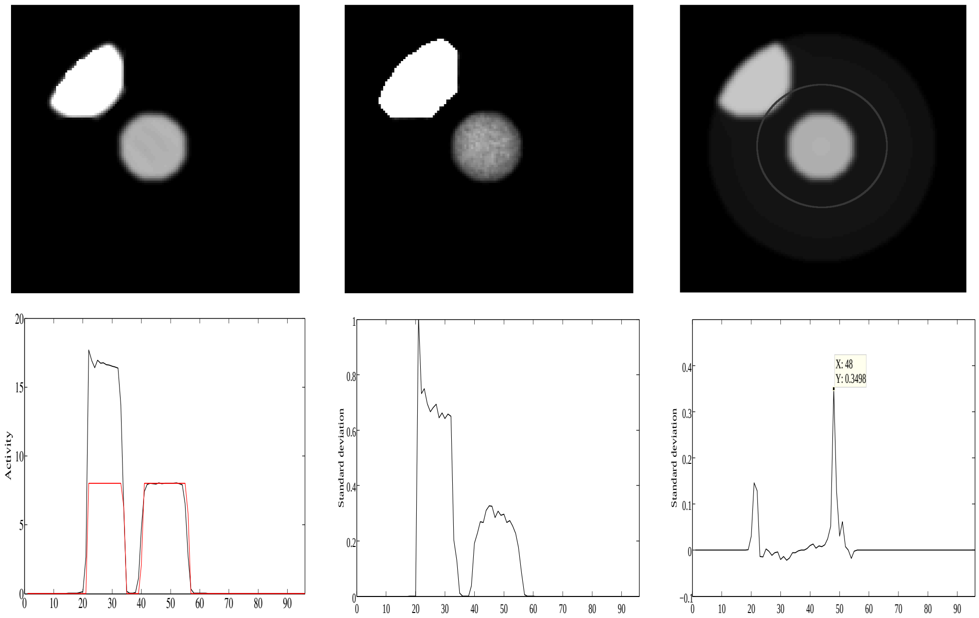

Figure 6.5: Truncated case. A (top left) - average image. B (top centre) - standard deviation image. C (top right) - activity phantom; the ROI is circled in red. D (bottom left) -Black - diagonal profile of the average image. Red - diagonal profile of the phantom. E (bottom centre) - diagonal profile of the standard deviation image. F (bottom right) - diagonal profile of the covariance for a point at the centre of the FOV. The standard deviation in the point in the centre of the FOV equals 0.349

.

It should be noted that, since there is no unique solution for the interior problem, the variance outside the ROI, theoretically, should present infinite value. However, the fact that we are including a smoothing prior to the cost function results in a variance outside the ROI that presents high but finite values. Even if the interior problem is not uniquely solvable, it is interesting to study the noise properties of the estimator for truncated emission data, using different approximations for the deterministic FIM-based method.

Fig. 6.7 shows variance images from truncated projection data with a ROI diameter $u = 36$. The variance image obtained from multiple noisy data sets and the variance images obtained from the full FIM and the subsampled FIM method with different grids are in good agreement. For the voxels outside the ROI, the variance increases considerably in respect to the non truncated case. Outside the ROI, in fact, we do not have full sampling, since we acquire data from that region only at certain angular positions of the camera.

The variance image obtained with the circulant approximation of the FIM method is displayed in Fig. 6.7 - E. From this image, we can notice an increase in variance in the voxels outside the ROI with respect to the non-truncated case. However the aforementioned effect is less accentuated with respect to the increase in variance estimated with full FIM method and with the subsampled FIM method in the same region. The horizontal profiles are shown in Fig. 6.7-F.

In Fig. 6.8 the CNR for a voxel in the centre of the sphere is plotted for different ROI diameters $u = 96, \cdots, 16$. The calculation of the CNR is obtained with the reference method (reconstruction of 1024 noisy data sets), with our novel approach for the approximation of the FIM (with grid of $g_1 = 27648$ (full FIM), $g_2 = 6912$, $g_3 = 3072$ points) and, for comparison, with the circulant approximation based on a single row of the FIM.

For this specific experiment we observe an overall decrease in CNR of 10% compared to the non-truncated case, using the reference method based on multiple noise realisations for the calculation of the variance. An important observation is that we see no effect due to truncation with the circulant approximation of the FIM (as stated in [Zhou et al., 2010]) whereas with the sumbsampled FIM, since we account for the interdependence between the voxels, we see a decreased CNR (increased variance) with increased level of truncation. This is an important feature of the method that we have introduced, because it enables the optimisation of systems for interior imaging, which is not possible with existing methods.

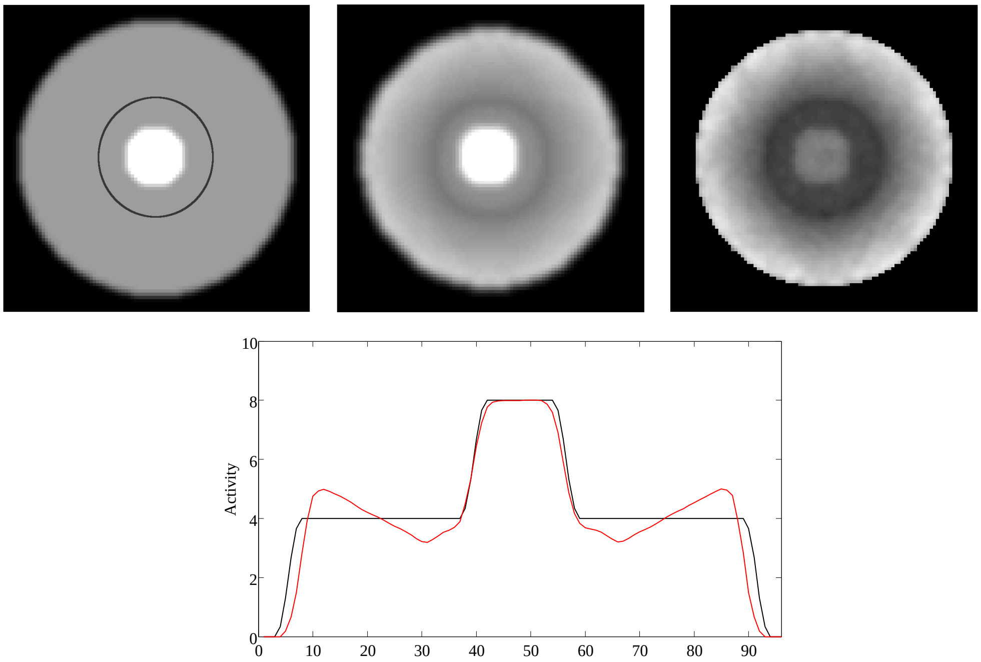

Figure 6.6: Mean and variance images from truncated projection data with a ROI diameter $u = 36$, for a uniform sphere phantom and uniform non-zero background. A (top left) - Digital phantom. B (top centre) - Mean image. C (top right) - Variance image. D (bottom) - Black line: phantom profile. Red line: mean profile.

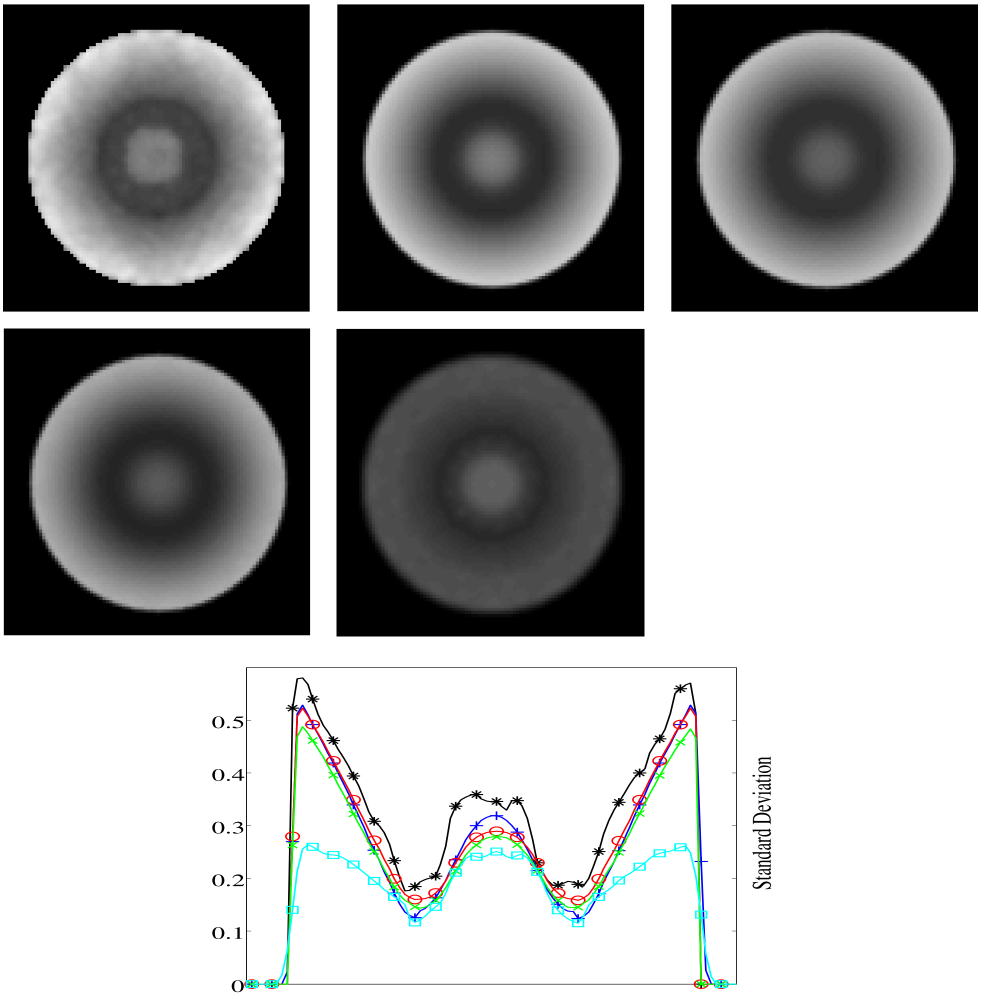

Figure 6.7: Interior tomography: Variance images for a uniform sphere phantom obtained with truncated projection data with ROI diameter $u = 36$. A (top left) - Reference method (variance image obtained from the reconstruction of 10240 noisy projection data sets), B (top central) - Full Fisher Information-based method with grid g_1, C (top right) - Subsampled Fisher Information based method with grid g_2, D (central left) - Subsampled Fisher Information based method with grid g_3, E (central)- Variance image obtained with the circulant approximation method. F (bottom)- Variance profiles at the centre of the heart: reference method (black line - $*$), grid g_1 (blue line - $+$), grid g_2 (red line - $\circ$), grid g_3 (green line - $\times$), circulant (cyan line - $\Box$).

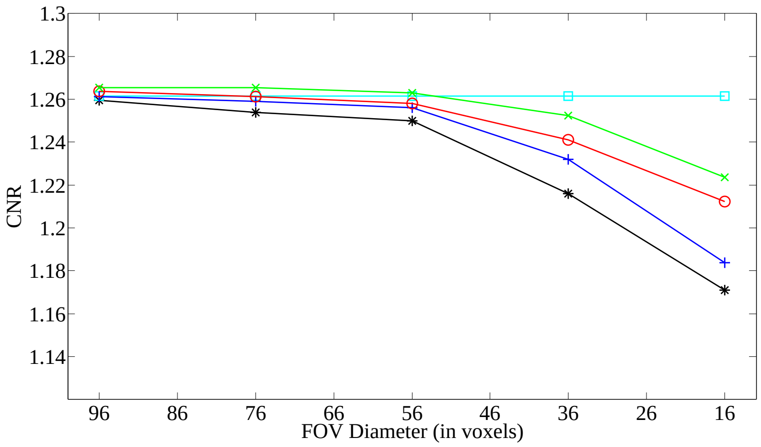

Figure 6.8: Interior tomography: CNR for different levels of truncation for a voxel in the center of the sphere. ROI diameter (from left to right) $u = 96, \cdots, 16$. Black line $(*)$ - CNR obtained with the reference method. Blue line $(+)$ - CNR obtained with full FIM method with grid g_1. Red line $(\circ)$ - CNR obtained with subsampled FIM method with grid g_2. Green line $(\times)$ - CNR obtained with subsampled FIM method with grid g_3. Cyan line $(\square)$ - CNR obtained from the circulant approximation method.

6.5.4 Validation

In section 6.5.3 we prove how our new approximation for the calculation of the FIM well predicts the variance of the estimate in case of truncation of the projection data. A validation for the subsampled Fisher Information-based variance calculation method, in case of data truncation, is presented in the following.

For the experiment presented in section 6.4, the variance obtained with the reference statistical method is plotted with respect to the variance predicted with the Fisher Information-based method, for grid $g_1 = 27648$ points, grid $g_2 = 6912$ points, grid $g_3 = 3072$ points, grid $g_4 = 1728$ points and grid $g_5 = 1106$ points. A least squares fitting is performed through the data. The regression coefficients, the intercepts of the line, the correlation coefficients and the standard error of the estimate for every experiment are presented in Table 6.2.

All validation points for the experiment in section 6.4 for those three grid models and for the circulant approximation method are shown in Figure 6.9. In these figures, the standard deviation calculated with the reference method with respect to the standard deviation predicted

Table 6.2: Validation of the Subsampled FIM

Inter	CC	SEE	int	RC
g1	0.9539	2.33 e-04	-0.0022	0.9529
g2	0.9417	4.90 e-04	-0.0047	0.9106
g3	0.9208	8.25 e-04	-0.0062	0.8687
g4	0.8942	0.0024	-0.0146	0.8174
g5	0.8296	0.0029	-0.0201	0.7955

Validation of the Subsampled FIM in comparison with the Reference Statistical Method for the experiment presented in section 6.4. CC: Correlation Coefficent. SEE: Standard Error of the Estimate. int:intercept of the least squares fit. RC: regression coefficient of the least squares fit.

with the subsampled FIM method and with the circulant approximation method are plotted. The solid line was fitted to minimise the least squares distance between these points. From these images it is possible to see how the subsampled FIM method well determines the variance of the estimate, although the level of approximation increases when the grid becomes more sparse. The correlation coefficient between the variance obtained with the reference statistical method and the variance obtained with the subsampled FIM with grid g_1, grid g_2 and grid g_3 ranges from 0.9 to 1. The Standard Error of the Estimate does not exceed 0.001.

Figure 6.9-D shows the least squares fitting between the reference method and the results obtained with the circular FIM. The fitted line is inclined towards the x-axis (slope 0.424, y-intercept 0.017), which means that the circulant method systematically underestimates the standard deviation. It is therefore clear from this figure that this approximation of the FIM does not account properly for truncation effects in the estimation of the variance.

6.6 Discussion

In this chapter we have used an iterative PL reconstruction algorithn with an added quadratic regularisation penalty to reconstruct two types of ROIs, namely, peripheral ROIs, and the interior ROIs.

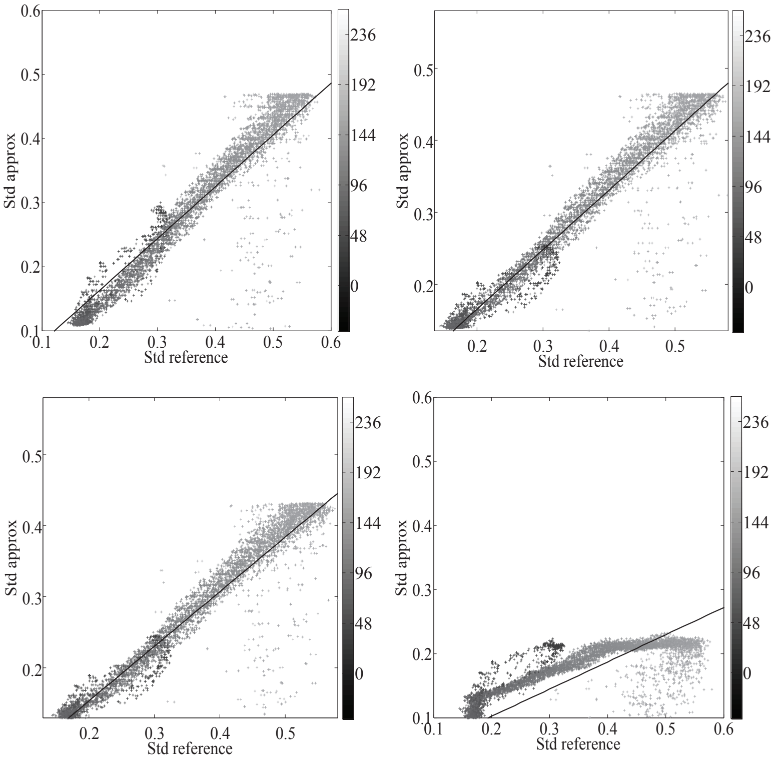

Figure 6.9: Validation of the Fisher information-based standard deviation calculation method for different grids, in case of truncated data. The standard deviation obtained with 10240 repeated simulations is plotted with respect to the standard deviation predicted with the Fisher information-based method. A (top left)- grid g_1, B (top right) - grid g_2, C (bottom left) - grid g_3 and D (bottom right) - circulant approximation. Spatial information is given by the colour of the dots. The light grey dots represent voxels which are more distant from the rotation axis (off-centre voxels); whereas the dark grey dots represent voxels which are closer to the rotation axis (central voxels). A colour bar with the respective distance from the rotation axis (in mm) is displayed for every plot.

We found that peripheral ROIs, which satisfy the analytical data sufficiency condition presented in section 6.4, can be exactly reconstructed using the PL reconstruction algorithm, while reconstructed images for the interior problem suffer from artifacts. These results are consistent with corresponding analytical data sufficiency conditions. The potential benefit of ROI statistical reconstruction with the PL algorithm over analytical ROI reconstruction method is that the PL reconstruction can handle both classes of problems in a unified framework.

For a general interior tomography case, it is extremely difficult to know the true value of the activity distribution from a certain region within the ROI. It has been suggested to perform a quick scout scan which is truncation-free to obtain an approximate estimate of the true activity distribution. However, the scout images are rather noisy due to the limited amount of data acquired and it may be difficult to determine the exact position of the object boundary from such images.

Given these considerations, it is important to study the statistical properties of the estimator (in terms of bias and variance) in case of interior ROI reconstruction from truncated projection data. In this chapter, simulation experiments have been presented to investigate the statistical properties of the estimator under the two different truncation cases. The peripheral ROI reconstruction and the interior ROI reconstruction have been investigated simulating different levels of truncation. In both cases the estimator $\hat{\lambda}$, suffers from increased variance, depending on the amount of data truncation. Even if peripheral ROI reconstruction can lead to nearly unbiased reconstruction, in a well-sampled ROI, we noticed that a decrease in ROI size leads to an increase in variance, not only outside the ROI but also inside it.

In section 6.5.3 of this chapter, a comparison between the results obtained using the subsampled FIM approximation and the circulant FIM approximation, in case of truncation, is presented. From these results, we can see how with the circulant approximation the variance inside the ROI appear to be independent of the present of truncation in the projection data, which is inconsistent with the results obtained with all the other methods for the calculation of the variance (and hence of the CNR).

Comparison of the variance (the diagonal of the covariance matrix) obtained from the circulant approximation, with the full FIM and with the reference statistical method, has therefore highlighted certain pitfalls of the circulant approximation in case of data truncation. In the following, the link between the shift-variance properties of the system response and the different approximations of the FIM, in case of data truncation, is discussed and illustrated with an example.

Figure 6.10 shows the full Fisher Information Matrix (Figures 6.10 A-B) the subsampled Fisher Information Matrix with grid g_2 (Figures 6.10 C-D) and the circulant Fisher Matrix (Figures 6.10 E-F) for the experiment described in section 6.4. Since the FIM is a very large matrix and therefore difficult to display, we show the FIMs for the 2-D case. Therefore the full FIM has $g_1 = 9216$ points and the subsampled FIM has $g_2 = 2304$ points.

As already described in section 4.5.2, all elements of the subsampled FIM at the grid points are calculated exactly, accounting for the acquisition geometry and for the object. We can see, in fact, how in the subsampled FIM we account for the system response and for the object dependency, whereas with the circulant approximation method we make the assumption that the FIM (and therefore the system response) is shift invariant.

However, what is ultimately of interest is the inverse of $(\boldsymbol{F} - \beta \cdot \boldsymbol{R})$ (the covariance matrix). Figure 6.11 shows the covariance matrix calculated from the full FIM (Figure 6.11 A), the covariance matrix calculated from the subsampled FIM with grid g_2 (Figure 6.11 C) and the covariance matrix calculated using the circulant approximation method (Figure 6.11 B) for the experiment described in section 6.4. The covariance matrix calculated using the circulant approximation method is spatially variant (non circulant) but clearly does not show the same structure as the covariance obtained from the full FIM in Figure 6.11 A. We can therefore deduce that the circulant FIM can not incorporate the effects of shift-variance, since it does not account for the effects of data truncation or missing data. Clearly, the inverse of $(\boldsymbol{F}_G - \beta \cdot \boldsymbol{R}_G)^{-1}$ will not be exact at the grid points because of the missing off-diagonal FIM entries between the grid points, however the matrices in 6.11 A and 6.11 C exhibit the same structure. This encompasses the capability of the method to incorporate non-stationary system models, effects of long-range correlations and data truncation.

In conclusion we can infer that no effect due to truncation is observed with the circulant approximation of the FIM, whereas the subsampling approach seems to be very accurate for evaluating the effects of data truncation or missing data, since it accounts for the interdependence between all the voxels, This is an important feature of the method that we have

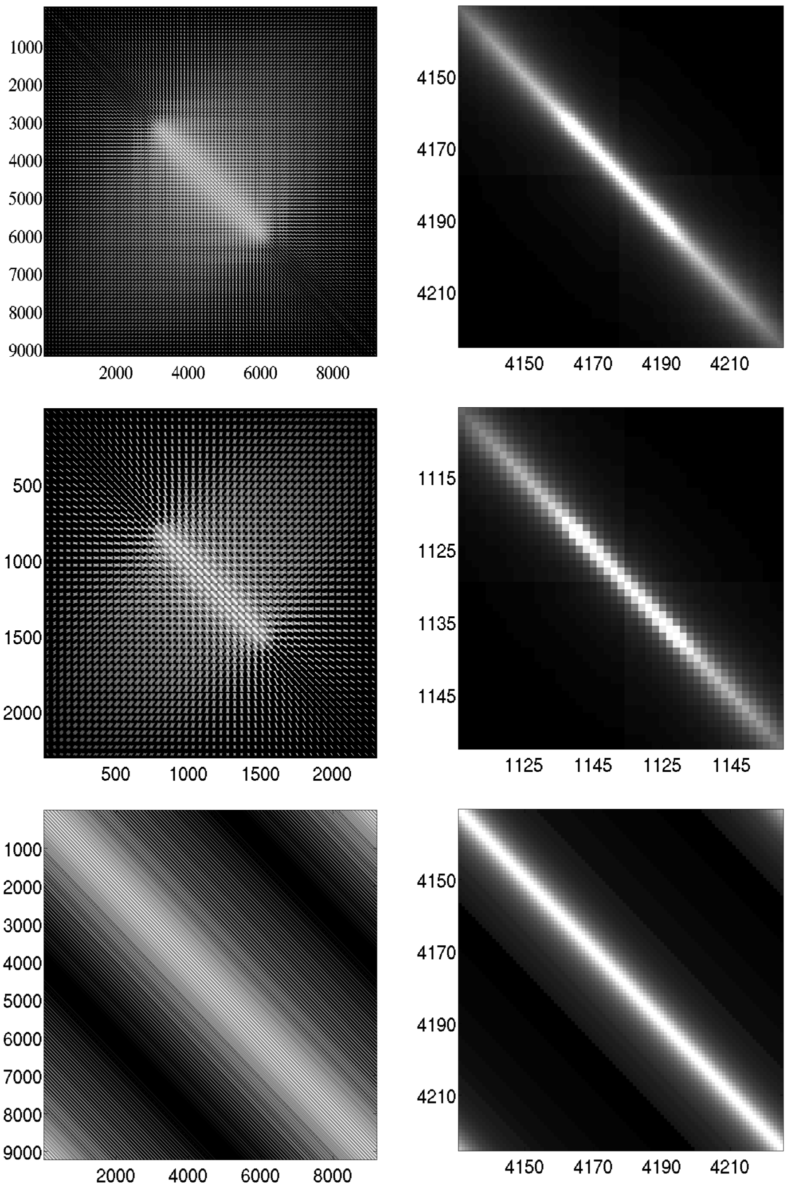

Figure 6.10: Fisher Information Matrix for the experiment in section 6.4. Comparison between full FIM, subsampled FIM and Circulant FIM. A (top left) - Full FIM, grid 9216 with points. B (top right) - A zoom-in part of A. C (central left) - subsampled FIM, grid with 2304 points. D (central right) - A zoom-in part of C. E (bottom left) - Circulant FIM for a voxel of interest i in the centre of the FOV only. F(bottom right) - A zoom-in part of E.

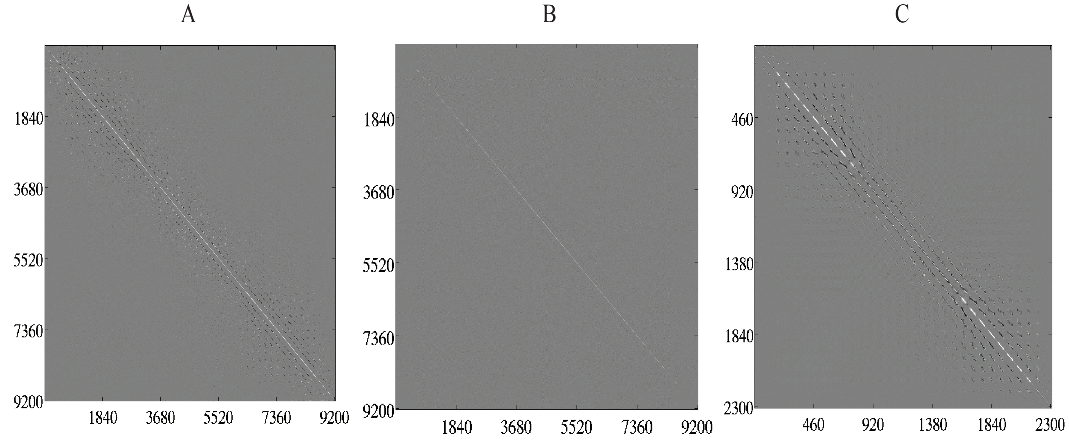

Figure 6.11: Inverse of $\boldsymbol{F} - \beta \cdot \boldsymbol{R}$ (Covariance matrix) for the experiment in section 6.4. Comparison between the Covariance matrix obtained from the full FIM, the subsampled FIM and obtained by row-by-row inversion of the circulant FIM, where each column of the matrix is evaluated separately using (4.41). A - Covariance matrix obtained from the Full FIM, grid $g_1 = 9216$ points. B - Covariance matrix obtained by row-by-row inversion of the circulant FIM. C - Covariance matrix obtained from the subsampled FIM, grid $g_2 = 2304$ points.

introduced, because it enables the optimisation of systems for interior imaging and of imaging systems that depart from the uniformly sampled circular trajectory of SPECT (as we will demonstrate in the next chapter).

Chapter 7

Changing Acquisition Trajectory: The D-SPECT System

7.1 Introduction

The D-SPECT system (Spectrum Dynamics, Caesarea, Israel), based on novel detector technology and a unique acquisition geometry, offers potential advantages in nuclear imaging compared to conventional gamma camera single photon emission computed tomography (SPECT) systems [Erlandsson et al., 2009] [Gambhir et al., 2009].

The design of a SPECT system has remained unchanged for many years. As described in section 2.1.2 and section 2.1.3, the primary components of this system are a collimator that limits the direction of the incoming photons, a scintillation crystal and a set of photomultiplier tubes that determine the location and the energy of the detected photons. The acquisition geometry usually involves the rotation of the camera along a circular trajectory at constant speed around the centre of the imaging volume; so that the detector collects photons during the same time interval for each angular position. Most commonly, dual detectors are used where both speed of rotation and detector location tend to be limited by the bulk of the overall system.

The disadvantages of scintillation detectors include bulkiness and relatively poor energy resolution. Solid-state detectors have long been used in spectroscopic applications due to their superior energy resolution, but have not been widely used in medical imaging for reasons of stability, practicality and cost. Solid-state detectors based on cadmium zinc telluride (CZT) can operate at room temperature, and recent technical advances have led to the development of pixelated CZT detector units appropriate for medical imaging applications [Wagenaar, 2004].

The absence of PM tubes allows for a compact and flexible design.

The design of the D-SPECT system is based on 9 arrays of CZT solid-state detectors units. This novel technology greatly reduces the bulkiness of the detectors and permits movements that would not be achievable with conventional gamma cameras.

This new design overcomes some limitations inherent to a SPECT system, allowing for a region-centric acquisition. By choosing to spend more time directing the detector heads towards a Region Of Interest (ROI), one can allocate more time to collect data from more informative regions at the expense of collecting less data from less informative regions (as for example outside the body contour). The trajectory of the gamma camera has a profound effect on the overall uncertainty of the measurement and on how the uncertainty is distributed throughout the imaging volume. Moving the camera along a circular trajectory at constant speed around the centre of the imaging volume produces, intuitively, optimum retrieval of the information when the object is (roughly) uniform and the PSF ideally not depth-dependent. Changing the detector angular movements allowing for a non uniform scanning pattern, just like for any other parameter of the acquisition system, the interdependence of the information changes and the D-SPECT system response may be highly shift-variant. Moreover, in the D-SPECT as in every SPECT system, there is an inherent space-variancy due to the intrinsic depth dependent response of the collimator and to the object dependency. Calculating the reconstructed image quality in a region-of-interest (ROI) is a complex problem that depends also on the presence of activity and on the system response outside the ROI.

It is therefore fundamental to compare different acquisition protocols and to investigate the influence of the presence of activity outside the ROI when optimising the design of this system.

The D-SPECT is an example of adaptive system, where the acquisition protocol (in terms of the camera trajectory and their position) can be modified depending on the data acquired during the scan, so as to image certain desired properties of the underlying object. In order to adapt the response of a system during acquisition, a set of different design parameters have to be compared in a reasonably short time. Thanks to a dedicated GPU implementation, the novel algorithm for the estimation of the uncertainty (presented in 4.5.2) drastically reduces the computational complexity and therefore is a good candidate method for adaptive optimisation problems.

The aim of this chapter is therefore to explore the use of the subsampled FIM approximation to compare a set of candidate scanning patterns and to optimise the acquisition protocol; emphasising how it enables us to explore the design of highly shift variant systems as a result of distance dependent resolution and adaptive data sampling.

The results obtained with the approximate subsampled FIM method are compared with the circulant approximation method (see section 4.5.1) and with the reference method based on reconstruction of multiple noise instances (see section 3).

7.2 System Description

Several new systems have recently been designed utilising semiconductors for detecting γ-photons [Madsen, 2007] [Patton et al., 2007]. Although presenting different scanning geometries, these systems, based on solid-state detectors, are all characterised by a compact design, and provide significantly higher count sensitivity, as well as spatial and energy resolution, than conventional SPECT systems.

7.2.1 CZT Detectors

Semiconductor detectors are solid-state devices that provide direct conversion of absorbed γ-ray energy into an electronic signal. Because there is no need for an intermediate high-gain amplification stage, these devices are compact and operate at low voltage. The absorbed energy from a γ-ray interaction liberates charge carriers (electrons and holes) within the charge-free depletion zone of the semiconductor. The induced charge on the terminals generates an electronic pulse with an amplitude proportional to the absorbed energy. Because there is no intermediate conversion stage, the precision of the signal is better than that of scintillators, resulting in correspondingly better energy resolution. Solid-state detectors based on cadmium zinc telluride (CZT) can operate at room temperature, and recent technical advances have led to the development of detector units appropriate for medical imaging applications [Wagenaar, 2004]. The intrinsic efficiency for CZT would be comparable to NaI(Tl) if the crystals were equally thick, however CZT crystals are typically thinner than scintillation crystals. CZT are available as pixelated detector arrays with a typical intrinsic spatial resolution of 2.46 mm and present a better energy resolution (than scintillator detectors) in the range of 2%- 5% for 140- keV γ-rays..

Table 7.1: pD-SPECT System Parameters

Detector block	1	2	3	4	5	6	7	8	9
Detector distance (mm)	228	203	205	229	248	256	254	249	262
FWHM (mm)	7.93	7.18	7.25	7.97	8.54	8.79	8.70	8.56	8.96

7.2.2 The D-SPECT System

In this section, the design of the novel D-SPECT system is described.

The commercial D-SPECT camera system is based on 9 collimated CZT detector columns arranged in a curved configuration in order to conform to the shape of the left side of the patient chest (Figure 7.1-A). Each detector column is then placed at a different distance from the centre of the imaging volume (Table 7.1). Each of the nine detector blocks is composed of 16×96 individual pixels with a size of $2.46\ mm$ in both dimensions, resulting in a total detector surface of $39.36\ mm \times 236.16\ mm$.

The tungsten parallel-hole collimators used by Spectrum Dynamics are shorter ($21.7\ mm$) and have larger square holes ($2.46\ mm$) than the standard LEHR parallel-hole collimators. This collimator design results in a linearly varying depth-dependent Gaussian response that has a slope of 0.136, which corresponds to $16.8\ mm$ FWHM at the centre of the field of view. The acceptance solid angle ($10.847 - 1023\ sr$) is therefore more than 8 times that of the standard LEHR lead parallel-hole SPECT collimator ($1.264 - 1023\ sr$). The square collimator holes are in registration with the crystal detector array, with 1 CZT pixel for each hole.

In this chapter we consider an adaptive SPECT system similar to the commercially available D-SPECT [Erlandsson et al., 2009]. In the following we refer to this adaptive SPECT system as the *pseudo* D-SPECT (pD-SPECT) . The design of the pD-SPECT differs from the commercially available D-SPECT in both the design of its collimators and the specifications of the acquisition protocol. Each detector block is equipped with a square PH collimator with septa thickness $q = 0.2\ mm$, hole diameter $v = 1.03\ mm$ and collimator length $l = 35\ mm$. The FWHMs, which depend on the distance of every detector from the centre of the FOV, are presented in Table 7.1. The square collimator holes are in registration with the pixelated detector array, with 4 collimator holes for each CZT pixel.

It can be noted that the collimator of the pD-SPECT system presents a response which is comparable to that of a LEHR collimator used in conventional SPECT. Therefore we do not account for higher sensitivity due to a broader collimator aperture but only for higher sensitivity due to a region-centric acquisition.

7.2.3 Acquisition Protocol

During data acquisition in the standard scanning mode, each of the nine individual detector blocks rotates in synchrony around its central axis in order to cover the whole field of view (FOV).

The adaptive dynamic sequence consists of two options:

Open-Sweep Acquisition: Each detector block rotates 110° in order to cover the whole FOV, performing 60 regular angular steps. In order to obtain a more complete tomographic sampling, the complete set of detectors is translated by 9° and the open-sweep acquisition is performed for a second time.

Region of Interest Acquisition: After a preliminary sweep mode scan, the operator defines a ROI contour. This ROI is then used to generate a specific scanning pattern, designed in such a way that each detector should spend more time acquiring data coming from the direction of the ROI than from regions outside. Data from the whole FOV are still acquired, however, so as to avoid truncation. The sequence of acquisition is therefore adapted in order to minimise the uncertainty in the ROI. The search of optimum scanning sequences is constrained by the following algorithm: each detector head covers the full angular span of $\varphi_3 - \varphi_0 = 110^\circ$, performing 60 angular steps $\delta\varphi$ (Figure 7.1 B):

$$\delta\varphi = \begin{cases} \dfrac{t((\varphi_3 - \varphi_2) + (\varphi_1 - \varphi_0))}{T(1-S)} & \text{if } \varphi_0 < \varphi \leq \varphi_1 \\ \dfrac{(\varphi_2 - \varphi_1)}{TS} & \text{if } \varphi_1 < \varphi \leq \varphi_2 \\ \dfrac{t((\varphi_3 - \varphi_2) + (\varphi_1 - \varphi_0))}{T(1-S)} & \text{if } \varphi_2 < \varphi \leq \varphi_3 \end{cases} \tag{7.1}$$

where T is the total scanning time, $t_a = T/60$ is the constant scanning time for every angular step, φ_1 and φ_2 are the angles subtended by the rays intersecting the centre of the detector and tangent to the ROI contour. The complete set of detectors is then translated by 9°; a new set of 60 angular steps is defined and the region centric acquisition is performed a second time.

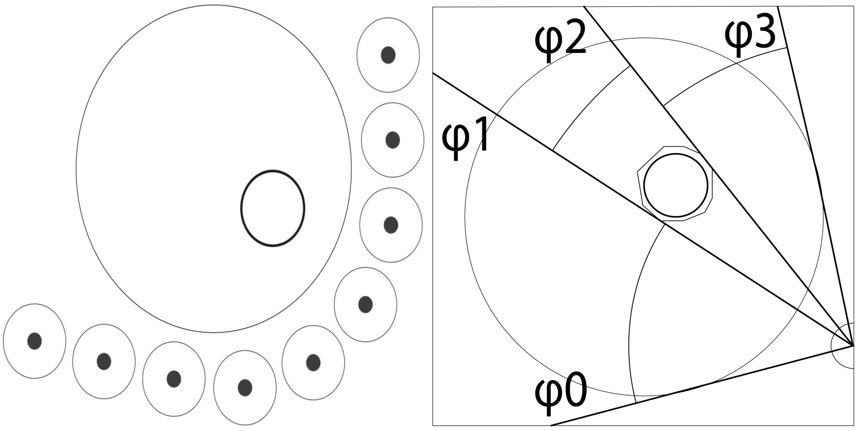

Figure 7.1: A (left) - Position of pD-SPECT detectors. B (right) - angular movement of a single pD-SPECT detector. Angular span of the FOV: $\varphi_3 - \varphi_0$. Angular span of the ROI: $\varphi_2 - \varphi_1$.

The only parameter defined in order to describe the scanning pattern is the time ratio S that each detector spends acquiring data from the ROI rather than from the surrounding region.

7.3 SVD for Different Acquisition Protocols

In chapter 6, the singular value decomposition (SVD) method has been used to study the properties of the imaging system in case of truncation of the projection measurements. Changing the camera trajectory has a profound effect on the overall system response and on its variability throughout the imaging volume. If with the Open-Sweep Acquisition ($S = 0.45$) the imaging volume is uniformly scanned, a ROI acquisition with $S = 1$ could lead to missing data in the projection domain, depending on the size and position of the ROI. It is therefore interesting to study the SVD of the system operator for different acquisition protocols, ranging from the open-sweep acquisition to the ROI-only acquisition.

The SVD of $\boldsymbol{H}'\boldsymbol{H}$ has been calculated as in equation (6.1), using the MATLAB's built-in function. The singular value spectra for the full and the subsampled crosstalk matrices are shown in Figure 7.3. The condition numbers for different acquisition protocols are presented in Table 7.2.

7.4 Statistical Calculation of the Uncertainty

In order to reconstruct the projection data, an iterative PL reconstruction, which converges to an image that maximises the Poisson likelihood of the data, with an added quadratic smooth-

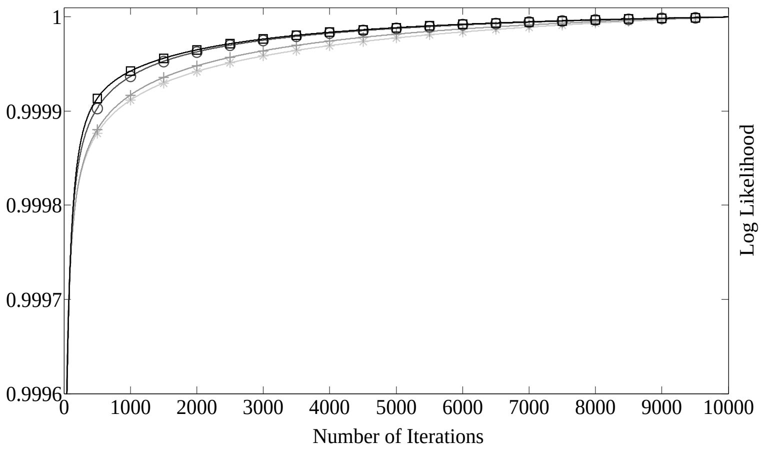

Figure 7.2: Objective function Ψ curves as a function of number of iterations. The different curves represent the objective function for different scanning patterns of a pD-SPECT system. The scanning pattern is defined by its time ratio S which varies ranging from $S = 0.45$, to $S = 0.9$ (from dark gray $\Box$ to light gray $*$). All the curves are scaled with respect to their maximum value.

ness penalty, has been used. The noisy data sets were reconstructed using an accelerated GPU implementation of the One Step Late algorithm for PL estimation, implemented as part of the NiftyRec toolbox [Pedemonte et al., 2010].

10000 iterations are performed in order to maximise the cost function. A smoothing prior with a weight $\beta = 10^{-10}$ was included in the cost function. The value of the regularisation parameter was chosen after trial and error, as a minimum value that nearly guarantees convergence within 10000 iterations. In Figure 7.2, the objective function $\Psi = L - \beta R$, for the digital phantom described in 7.5 - A, is plotted as a function of the number of iterations. The different curves represent the objective function for different acquisition protocols.

A series of independent noise realisations is computed using a pseudo-random Poisson noise generator. The mean and the variance are calculated as in equation (3.3) and (3.6) respectively. The number of noise instances is 10240. NiftyRec can process concurrently up to 1024 reconstructions in order to make efficient use of the GPU and 10 repetitions were chosen in order to obtain satisfactory images of variance.

7.4.1 Deterministic Method based on the Fisher Information

For the deterministic method, the subsampled version of the FIM has been calculated over three different grids of $g_1 = 27648$ (full FIM), $g_2 = 6912$ and $g_3 = 3072$ points equally distributed over the slice intersecting the point (or ROI) of interest and the two neighbouring slices.

The calculation of the covariance is then performed as in equation (4.47) for different acquisition protocols. The variance images in Figures 7.8 (B-C-D) and 7.9 (B-C-D), are obtained by re-ordering the diagonal of the covariance matrix Cov calculated as in (4.47). For the full FIM (g_1), the diagonal of Cov is simply reshaped to a three dimensional matrix, whereas for grids g_2 and g_3 every point of the diagonal of Cov_G is allocated to the respective points of the grid in the imaging volume. As described in section 4.8, a trilinear interpolation is then performed in order to facilitate the visual comparison between the variance images obtained with the different grid models.

Variance images obtained with the circulant approximation method are also presented for comparison. Each pixel of the images in Figures 7.8 (E) and 7.9 (E), is calculated according to equation (4.41).

7.5 Experiments

The main aim of this set of experiments is to compare different acquisition protocols for the D-SPECT system and to investigate the influence of the presence of activity outside the ROI in the optimisation.

7.5.1 A - NCAT Phantom

The first experiment is performed to compare different acquisition protocols with a realistic phantom.

The phantom used for this experiment is a heart phantom (NCAT) [Segars, 2001]. The activity within the phantom was $\lambda = 8kBq/cm^3$ and $\lambda = 4kBq/cm^3$ in the right and left ventricle myocardium respectively, $\lambda = 0.6kBq/cm^3$ in the left and right ventricle chamber, $\lambda = 0.6kBq/cm^3$ in the lungs, $\lambda = 8kBq/cm^3$ in the liver and $\lambda = 0.1kBq/cm^3$ in the background. The image volume dimensions are $96 \times 96 \times 12$ cubic voxels of $2.46\ mm$.

For this experiment, the scanning pattern varies ranging from $S = 0.45$, to $S = 0.9$ (where $S = 0.45$ is the time ratio for the open-sweep modality). The mean and the variance are calculated using the statistical method for the calculation of the uncertainty. The CNRs are calculated for the central point of the ROI at a constant target resolution $FWHM_t = 12\,mm$.

7.5.2 B - Contrast Phantom

The second experiment is performed to investigate the influence of the presence of activity outside the ROI in the optimisation of the acquisition protocols.

The software phantom is a uniform sphere positioned at a mean distance from the detectors of $205\,mm$. The sphere is then positioned in the half of the FOV closest to the detectors, where there is complete tomographic sampling. The activity in the sphere is set to $\lambda = 8kBq/cm^3$ and the background activity is set to three different levels $\lambda_{BK1} = 0.9kBq/cm^3$, $\lambda_{BK2} = 2.2kBq/cm^3$ and $\lambda_{BK3} = 4.4kBq/cm^3$. The image volume dimensions are $96 \times 96 \times 12$ cubic voxels of $2.46\,mm$.

We perform, for each of the three different backgrounds in the object, a set of experiments in which the scanning pattern varies ranging from $S = 0.45$, to $S = 0.9$.

For this experiment, the deterministic FIM based method for the calculation have been used for the calculation of the LLIR and the variance. Results obtained with the subsampled FIM approximation, the circulant approximation and the reference method are presented for comparison. The CNRs are calculated for the central point of the ROI at a constant target resolution $FWHM_t = 12\,mm$.

7.6 Results

In the following sections 7.6.2 and 7.6.3, results for different acquisition protocols for a pD-SPECT system are shown and discussed. The results are obtained with the presented subsampled approximation of the FIM and, for validation and comparison, with the circulant approximation of the FIM and with the reference statistical method. In section 7.6.1, results obtained with the singular value decomposition (SVD) are discussed to study the properties of the imaging system. In section 7.7, all validation points are gathered to get a global overview of the agreement between the new method and the reference method.

7.6.1 SVD for Different Acquisition Protocols

Figure 7.3 - A shows the singular value spectra of the system operator $\boldsymbol{H}$, for different acquisition protocols of a pD-SPECT system. The ROI is the one defined for the experiment in section

7.5.1. The acquisition protocol is determined by the time ratio that ranges from $S = 0.45$, the open-sweep acquisition, to $S = 1$, for a ROI-only acquisition.

Performing an open-sweep acquisition, the imaging volume is uniformly scanned and only few singular vectors present a singular value below that is nearly zero. For increasing time ratio S, as we allocate more time directing the detector heads towards the ROI, an increasing number of singular vectors present a singular value that is nearly zero, indicating that the system matrix becomes increasingly more ill-conditioned. It should be noted, however, that only for $S = 1$ the singular value spectrum presents a significant number of singular values close to the zero value, since a ROI-only acquisition leads to missing data in the projection domain.

Figure 7.3 - A shows the singular value spectra for the full crosstalk matrix with grid $g_1 = 27648$, whereas Figure 7.3 - B shows the singular value spectra for the subsampled crosstalk matrix with grid $g_2 = 6912$. As previously discussed in section 6.5.1, comparing these figures, it is possible to infer that calculating the SVD for the subsampled crosstalk matrix is equivalent to performing a subsampling of the singular value spectrum. Therefore, the points of inflection of the curves are equivalent with the two methods. It should be noted that, for representation purposes, the vertical axis in the plots of Figure 6.3 present the same scale. Thus, the maximum values of Figure 6.3 - A are not displayed.

The condition numbers of the full crosstalk matrices and of the subsampled crosstalk matrices, for different acquisition protocols, are presented in table 6.1. The absolute values of the condition numbers for the two methods are obviously different since subsampling is performed on the singular value spectra and the maximum singular values for the full crosstalk matrices are most likely to differ from the maximum singular values for the subsampled case.

Thus the condition number alone can not be employed as a metric for comparison between different systems when subsampled crosstalk matrices (or subsampled Fisher matrices) are used.

7.6.2 A - NCAT Phantom

Figure 7.4 and 7.6 show the mean image obtained from the reconstruction of 10240 noisy projection data sets, for two different slices of an NCAT phantom. In Figure 7.4 the slice includes the left and right ventricle myocardium, whereas in Figure 7.6 the slice includes the apex of the heart and part of the liver. Figure 7.4 - A and 7.6 - A show the two corresponding slices of the digital phantom. Figure 7.4 - B - C - D and 7.6 - B - C - D show the mean images

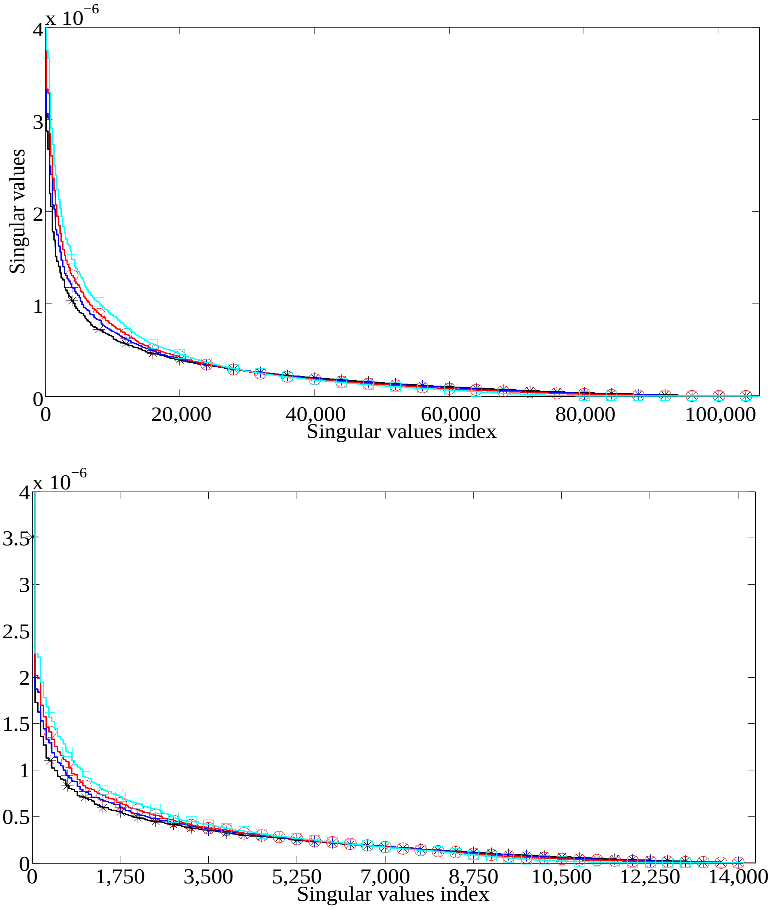

Figure 7.3: for the D-SPECT for different acquisition protocols. A (top) - Singular value spectra for the full crosstalk matrix with grid $g_1 = 27648$. B (bottom) - Singular value spectra for the subsampled crosstalk matrix with grid $g_2 = 6912$. The different curves show the SVD spectra for time ratios varying from $S = 0.45$ to $S = 1$. Black ($*$) - Time ratio $S = 0.45$, open-sweep acquisition. Blue ($+$) - Time ratio $S = 0.65$. Red ($\circ$) - Time ratio $S = 0.85$. Cyan ($\Box$) - Time ratio $S = 1$.

Table 7.2: Condition numbers for the pD-SPECT system with different acquisition protocols.

Svd pD-SPECT	S = 0.45	S = 0.65	S = 0.85	S = 1
g1	4.43 e+03	7.28 e+03	9.33 e+03	1.76 e+04
g2	220.94	400.94	694.12	5.98 e+03
g3	65.33	116.75	215.9	3.04 e+03
g4	35.26	59.62	109.13	925.95

Condition numbers of the full crosstalk matrix with grid $g_1 = 27648$ points and of the subsampled crosstalk matrices with grid $g_2 = 6912$ points, grid $g_3 = 3072$ points and grid $g_4 = 1728$ points. The condition number are presented for each grid and for different time ratios, ranging from $S = 0.45$ to $S = 1$.

for time ratio $S = 0.45$, $S = 0.85$ and $S = 1$. The horizontal profiles for the digital phantom and for the mean images are given in Figure 7.4 - E and 7.6 - E. From these images we can see how the structure inside the ROI can be accurately resolved for both time ratios $S = 0.45$ and $S = 0.75$. On the contrary, for time ratio $S = 1$, in both slices, the amount of activity in the right ventricle is underestimated. From Figure 7.6 - D we can also see how for time ratio $S = 1$ the reconstructed image outside the ROI (the liver) is affected by severe artifacts. From these results we can therefore deduce that for time ratios $S = 0.45$ and $S = 0.75$ the activity distribution can be accurately reconstructed, whereas for time ratio $S = 1$ the reconstructed image is affected by artifacts.

The calculated variance images for the two slices of the NCAT phantom are shown in Figure 7.5 and Figure 7.7. Figure 7.5 - A - B - C - D and 7.7 - A - B - C - D show the variance images for time ratios $S = 0.45$, $S = 0.65$, $S = 0.85$ and $S = 1$. The horizontal profiles for the variance images are given in Figure 7.5 - E and 7.7 - E.

From these images we can see how, for an NCAT phantom, changing the detector angular movements and allocating more time to collect data from a ROI, the variance inside the ROI decreases with increasing time ratio S. We can however notice how with time ratio $S = 1$ the variance in the ROI is higher than the variance obtained with time ratio $S = 0.65$ and $S = 0.85$. Moreover from Figure 7.7 we can see how, with increasing time ratio S, the variance increases in the liver whereas the variance inside the ROI decreases. This effect on the variance is due to the fact that with increasing time ratio S more time is spent on the ROI at the expense of acquiring less information on the surrounding region. Though the information that is ultimately acquired about the ROI also depends on the information that is acquired in the surrounding region, imaging an NCAT phantom, the net effect of increasing S is to increase the overall information about the activity in the ROI.

The optimum time ratio obtained with the proposed approach is a complex function of many factors, including the location of the ROI with respect to the orbit of the camera and the relative location and strength of the background features. Sampling angles corresponding to larger probabilities for detecting the gamma rays originated from the ROI are generally more important, whereas sampling angles that lead to greater sensitivity to strong background features are generally less favourable. For an NCAT phantom where the amount of activity in the background is low with respect to the activity in the ROI an acquisition protocol that spend more time acquiring data from the ROI may be preferable. For time ratio $S = 1$, there is an increase level of variance due to missing data in the projection domain. A ROI-only acquisition is therefore not recommended even when imaging a cardiac phantom.

7.6.3 B - Contrast Phantom

Figures 7.8 and 7.9 show the calculated variance for a slice intersecting the centre of the sphere for time ratio $S = 0.45$ and 0.85 respectively and background set at $\lambda = 2.2kBq/cm^3$. Figures 7.8-A and 7.9-A show the results obtained from the reference statistical method based on the reconstruction of multiple noise realisations. The results obtained with the deterministic method based on a subsampled version of the FIM are shown in Figures 7.8-B and 7.9-B for a grid characterised by $g_1 = 27648$ points and in Figures 7.8-C and 7.9-C for a grid of $g_2 = 3072$ points. Moreover the results obtained with the circulant approximation of the FIM method are shown in Figures 7.8-D and 7.9-D. The profiles of the image taken from a diagonal line intersecting the centre of the sphere are shown in the Figures 7.8-E and 7.9-E.

From the images obtained with the reference method and with the full and subsampled FIM, we see how with increasing time ratio S, the variance increases in the region outside the ROI whereas the variance in the uniform sphere decreases. This intuitive effect on the uncertainty in the measurements is due to the fact that, with an open-sweep acquisition, the entire

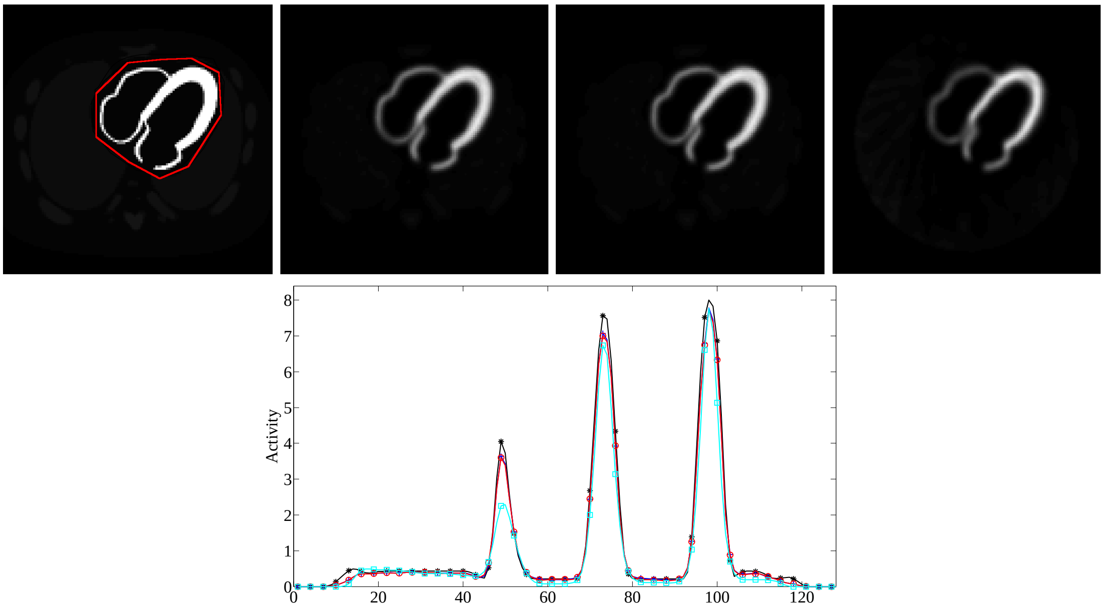

Figure 7.4: Mean images obtained from the pD-SPECT system with varying scanning pattern time ratio. The mean images are obtained from the reconstruction of 10240 noisy projection data sets. A (top left) - A slice of the digital NCAT phantom including the left and right ventricle myocardium. The ROI is highlighted by a red line. B (top centre-left) - Mean image for time ratio $S = 0.45$. C (top centre-right) - Mean image for time ratio $S = 0.75$. B (top right) - Mean image for time ratio $S = 1$. E (bottom) - Mean image profiles over a diagonal intersecting the centre of the ROI: digital phantom (black line - $*$), time ratio $S = 0.45$ (blue line - $+$), time ratio $S = 0.85$ (red line - $\circ$), time ratio $S = 1$ (cyan line - $\Box$).

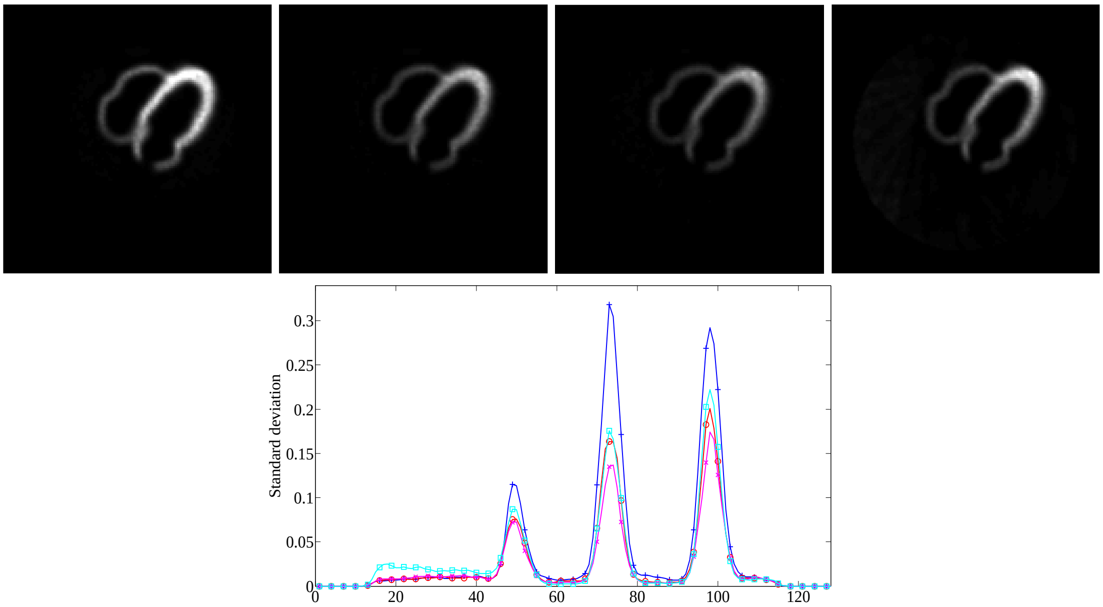

Figure 7.5: Variance images obtained from the pD-SPECT system with varying scanning pattern time ratio. The variance images are obtained from the reconstruction of 10240 noisy projection data sets. The slice of the digital NCAT phantom includes the left and right ventricle myocardium. A (top left) - Variance image for time ratio $S = 0.45$. B (top centre-left) - Variance image for time ratio $S = 0.65$. C (top centre-right) - Variance image for time ratio $S = 0.85$. B (top right) - Variance image for time ratio $S = 1$. E (bottom) - Variance image profiles over a diagonal intersecting the centre of the ROI: time ratio $S = 0.45$ (blue line - $+$), time ratio $S = 0.65$ (red line - $\circ$), time ratio $S = 0.85$ (magenta line - $\times$), time ratio $S = 1$ (cyan line - $\square$).

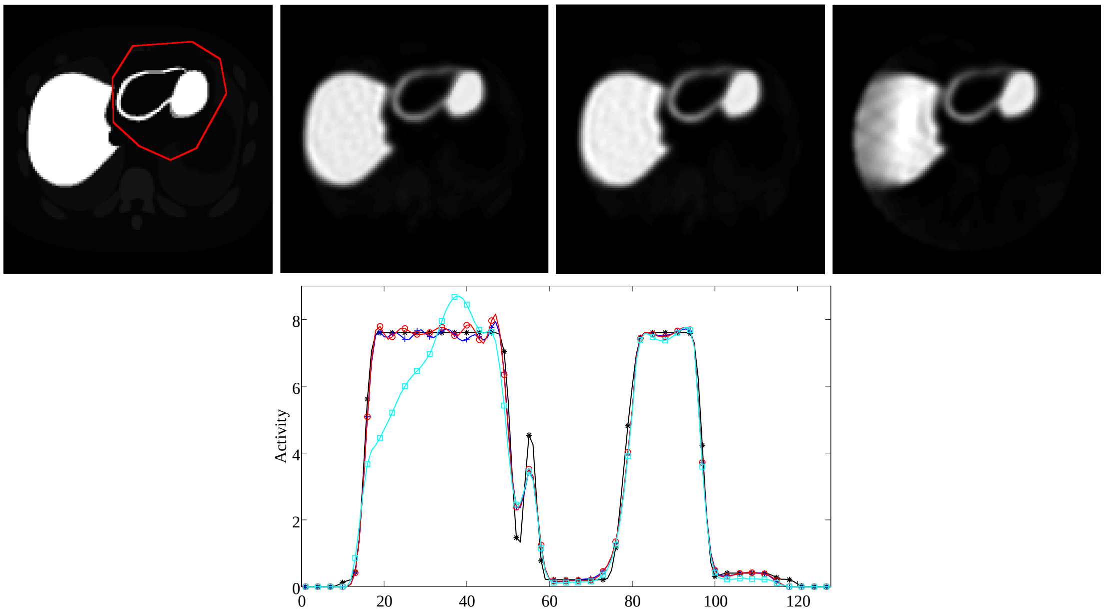

Figure 7.6: Mean images obtained from the pD-SPECT system with varying scanning pattern time ratio. The mean images are obtained from the reconstruction of 10240 noisy projection data sets. A (top left) - A slice of the digital NCAT phantom including the the apex of the heart and part of the liver. The ROI is highlighted by a red line. B (top centre-left) - Mean image for time ratio $S = 0.45$. C (top centre-right) - Mean image for time ratio $S = 0.75$. B (top right) - Mean image for time ratio $S = 1$. E (bottom) - Mean image profiles over a diagonal intersecting the centre of the ROI: digital phantom (black line - $*$), time ratio $S = 0.45$ (blue line - $+$), time ratio $S = 0.85$ (red line - $\circ$), time ratio $S = 1$ (cyan line - $\Box$).

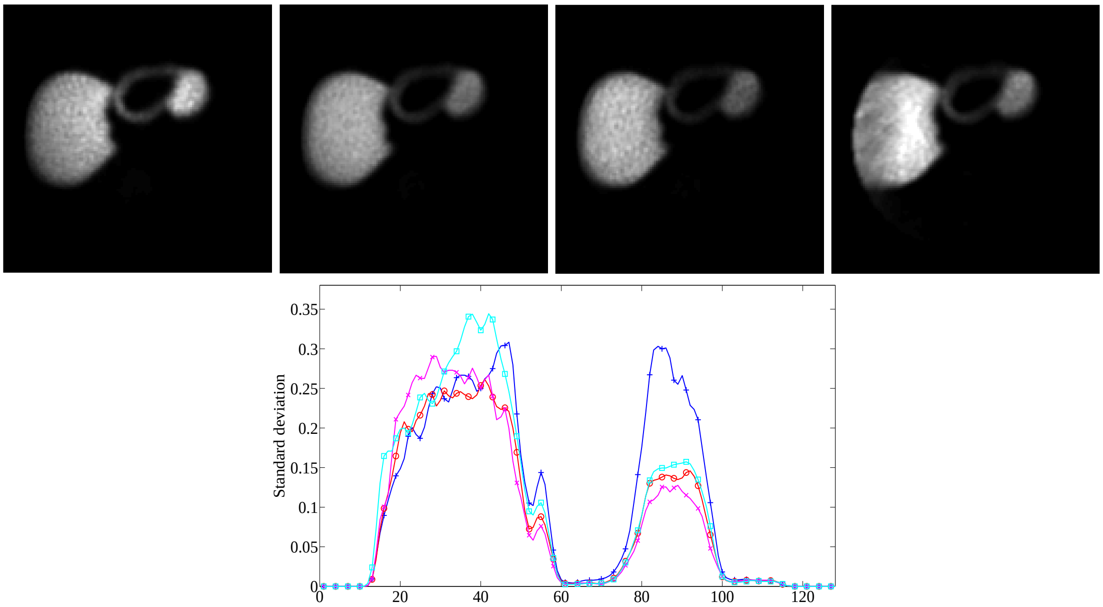

Figure 7.7: Variance images obtained from the pD-SPECT system with varying scanning pattern time ratio. The variance images are obtained from the reconstruction of 10240 noisy projection data sets. The slice of the digital NCAT phantom includes the apex of the heart and part of the liver. A (top left) - Variance image for time ratio $S = 0.45$. B (top centre-left) - Variance image for time ratio $S = 0.65$. C (top centre-right) - Variance image for time ratio $S = 0.85$. B (top right) - Variance image for time ratio $S = 1$. E (bottom) - Variance image profiles over a diagonal intersecting the centre of the ROI: time ratio $S = 0.45$ (blue line - $+$), time ratio $S = 0.65$ (red line - $\circ$), time ratio $S = 0.85$ (magenta line - $\times$), time ratio $S = 1$ (cyan line - $\Box$).

FOV is scanned uniformly, whereas with increasing time ratio S more time is spent on the ROI at the expense of acquiring less information on the surrounding region. Once again, for this set of experiments, using the circulant approximation of the FIM method, the aforementioned effect of increasing variance outside the ROI with increasing time ratio S is less accentuated with respect to the increase in variance obtained from the full FIM method and also with respect to the increase in variance obtained from the subsampled FIM method, since the circulant method does not account for effects of long distant correlations.

The plots in Figure 7.10 show the variation of CNR in the central voxel of the sphere for different acquisition protocols whose time ratio varies ranging from $S = 0.45$ to $S = 0.9$. Three experiments were performed for different values of the activity in the background $\lambda_{BK1} = 0.9kBq/cm^3$ (Figure 7.10-A), $\lambda_{BK2} = 2.2kBq/cm^3$ (Figure 7.10-B) and $\lambda_{BK2} = 4.4kBq/cm^3$ (Figure 7.10-C). From these plots it can be seen that the optimal scanning pattern is sensitive to the level of activity in the background. If the activity in the background is high with respect to the activity in the ROI, an acquisition that more uniformly scans the whole FOV may be preferable. This effect is captured by the subsampled Fisher Information based method with grid g_1 and grid g_2, whereas it is not captured by the circulant approximation of the FIM method. The circulant approximation method only accounts for the increased sensitivity in the ROI with increased time ratio S; whereas it does not account for the effects of long distant correlations due to a non-uniform scanning pattern.

In order to reinforce the observation that there is an influence of the acquisition parameters on the covariance; in Figure 7.11, we show images of covariance of a point in the centre of the sphere, for two different scanning patterns whose time ratios are $S = 0.45$ and 0.9 (from left to right). The covariances have been calculated using the reference method (which implies the reconstruction of multiple noise realisations). From these images we can deduce that with increasing time ratio S, there is a change in the dependency between the presence of activity outside the ROI and the uncertainty of the estimation inside the ROI.

7.7 Validation

In section 7.6.3, we prove how our new approximation for the calculation of the FIM well predicts the variance of the estimate for different acquisition protocols of a pD-SPECT system.

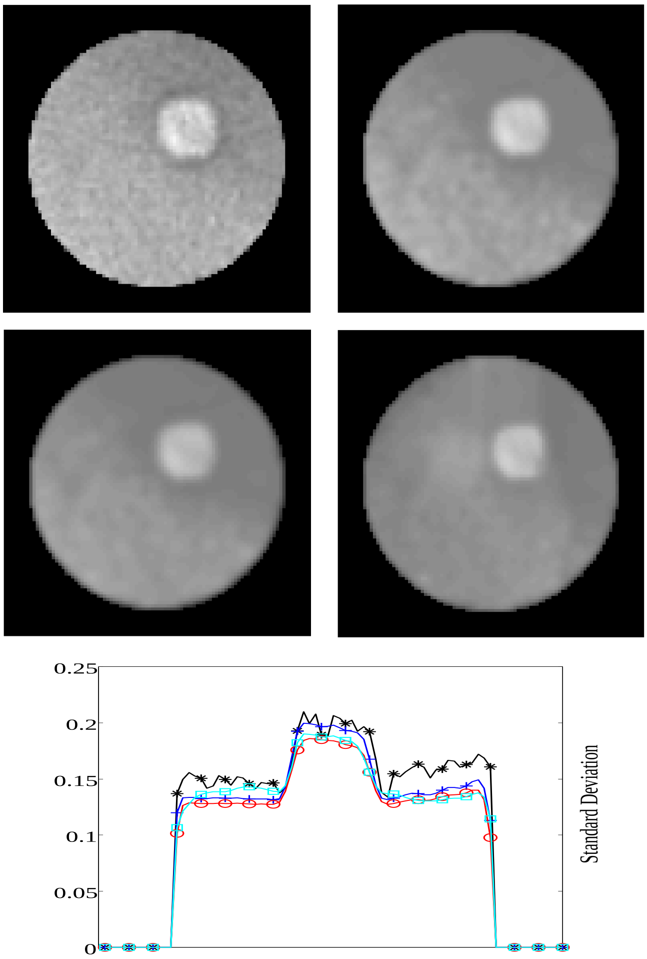

Figure 7.8: Variance images of a uniform sphere ($\lambda = 8kBq/cm^3$) and uniform background ($\lambda = 2.2kBq/cm^3$) obtained from the pD-SPECT system with scanning pattern time ratio $S = 0.45$. A (top left) - reference method (variance image obtained from the reconstruction of 10240 noisy projection data sets). B (top right) - Fisher Information-based method with grid g_1 (full FIM). C (central left) - Fisher Information based method with grid g_2. D (central right) - Variance image obtained with the circulant approximation method. E (bottom) - Image profiles over a diagonal intersecting the centre of the sphere: reference method (black line - $*$), grid g_1 (blue line - $+$), grid g_2 (red line - $\circ$), circulant approximation method (cyan line - $\square$)

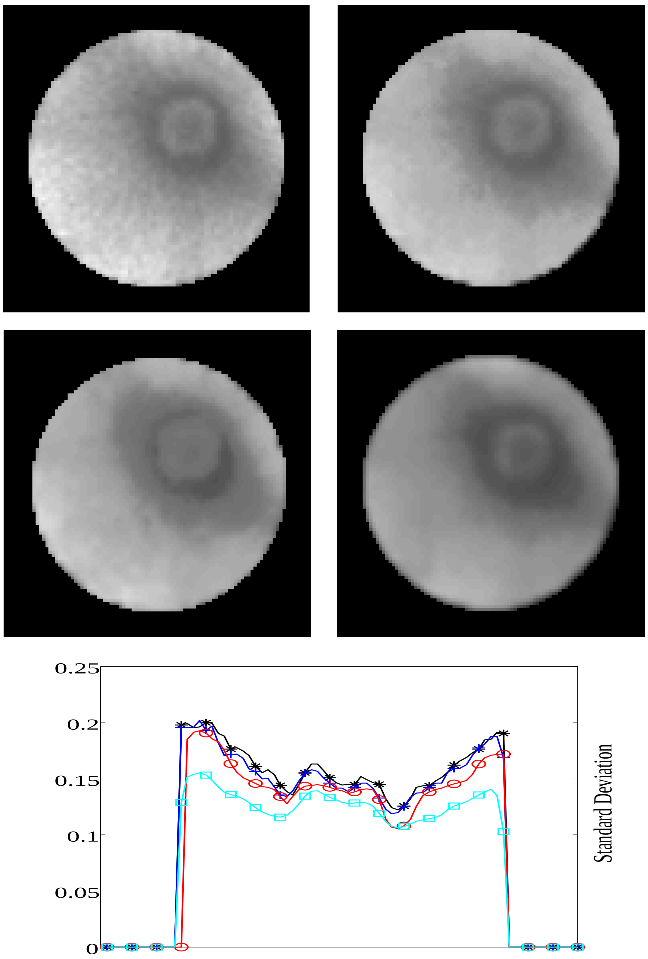

Figure 7.9: Variance images of a uniform sphere ($\lambda = 8kBq/cm^3$) and uniform background ($\lambda = 2.2kBq/cm^3$) obtained from the pD-SPECT system with scanning pattern time ratio $S = 0.9$. A (top left)- reference method (variance image obtained from the reconstruction of 10240 noisy projection data sets). B (top right) - Fisher Information-based method with grid g_1 (full FIM). C (central left) - Fisher Information based method with grid g_2. D (central right) - Variance image obtained with the circulant approximation method. E (bottom) - Image profiles over a diagonal intersecting the centre of the sphere: reference method (black line - $*$), grid g_1 (blue line - $+$), grid g_2 (red line - $\circ$), circulant approximation method (cyan line - $\square$)

Table 7.3: Validation of the Subsampled FIM for the pD-SPECT system.

Dspect	CC	SEE	int	RC
g1	0.95928	2.43 e-04	-0.00464	0.94834
g2	0.92837	3.47 e-04	-0.00646	0.89897
g3	0.89902	4.90 e-04	-0.01069	0.81060
g4	0.82983	0.0026	-0.01864	0.79574
g5	0.79345	0.0032	-0.02287	0.76643

Validation of the Subsampled FIM in comparison with the Reference Statistical Method for the experiment presented in section 7.6.3. CC: Correlation Coefficent. SEE: Standard Error of the Estimate. int: intercept of the least squares fit. RC: regression coefficient of the least squares fit.

A validation for the subsampled Fisher Information-based variance calculation method is presented in the following.
For the experiment presented in section 7.6.3, the variance obtained with the reference statistical method is plotted with respect to the variance predicted with the Fisher Information-based method, for grid $g_1 = 27648$ points, grid $g_2 = 6912$ points, grid $g_3 = 3072$ points, grid $g_4 = 1728$ points and grid $g_5 = 1106$ points. A least squares fitting is performed through the data. The regression coefficients, the intercepts of the line, the correlation coefficients and the standard error of the estimate for every experiment are presented in Table 7.3.

Figure 7.12 and Figure 7.13 show all validation points for the experiment in section 7.6.3, with time ratio $S = 0.45$ and $S = 0.9$ respectively. In these figures, the standard deviation calculated with the reference method with respect to the standard deviation predicted with the subsampled FIM method and with the circulant approximation method are plotted. The solid line was fitted to minimise the least squares distance between these points.

From these images we can observe how the subsampled FIM method well determines the variance of the estimate, although the level of approximation increases when the grid becomes more sparse. The correlation coefficient between the variance obtained with the reference statistical method and the variance obtained with the subsampled FIM with grid g_1 and grid g_2

ranges from 0.9 to 1. The Standard Error of the Estimate does not exceed 0.001.

Figure 7.12-D shows the least squares fitting between the reference method and the results obtained with the circular FIM for time ratio $S = 0.45$. From this figure, we can see how the circulant approximation method well determines the variance for the open-sweep acquisition and how the results obtained with the least squares fitting are comparable with the results obtained with the full FIM. Figure 7.12-D shows the least squares fitting between the reference method and the results obtained with the circular FIM for time ratio $S = 0.9$. In this case, the fitted line is inclined towards the x-axis (slope 0.77, y-intercept 0.044), which means that the circulant method systematically underestimates the standard deviation. It is therefore clear from this figure that this approximation of the FIM does not account properly for the effects of long distant correlations due to a non-uniform scanning pattern.

7.8 Discussion

The D-SPECT camera provides a novel photon collection method and scanning geometry that permits the independent movement of multiple detector columns in order to achieve a region-centric acquisition. The D-SPECT system has been developed to improve the trade-off between spatial resolution and sensitivity inherent in conventional SPECT systems. A direct comparison of the performances of a D-SPECT system with respect to a conventional SPECT system has been performed by Erlandsson et al. [2009]. In this study, the D-SPECT system presents a significant improvement in sensitivity with respect to the conventional SPECT system. The count improvement has been demonstrated in terms of absolute measurement in both planar and tomographic modes and it is due to the combined wide-angle collimator and region centric acquisition.

A higher spatial resolution for the D-SPECT system with respect to the conventional SPECT has been demonstrated in cardiac phantom studies and was confirmed in cardiac patient studies, with a better myocardial edge definition noted in both cases [Gambhir et al., 2009]. However, in this study, the D-SPECT image of the activity was reconstructed accounting for resolution modelling, while the conventional SPECT was reconstructed not accounting for that. Moreover, in this specific experiment, the D-SPECT camera could come closer to the heart than the conventional SPECT system, because of its smaller dimensions. This explain why, for the specific experiment presented in [Gambhir et al., 2009], a collimator with poorer resolution could lead to a reconstructed images with superior resolution.

Reconstructed spatial resolution is estimator and object-dependent and a comparative study for a set of different phantoms has not been performed yet. The system is designed specifically for cardiac SPECT use and the advantages in resolution and sensitivity of the D-SPECT camera reported from cardiac imaging may not necessary hold as other organs are imaged.

We considered an adaptive SPECT system similar to the D-SPECT as an example of a system in which the camera trajectory can be modified in response to the characteristics of the underlying activity distribution. The design of the pseudo D-SPECT differs from the commercially available D-SPECT in the design of its collimators. The collimator used for the simulations performed in this chapter has a response which is comparable to that of a LEHR collimator used in conventional SPECT. Therefore we do not account for higher sensitivity due to a broader collimator aperture but only for higher sensitivity due to a region-centric acquisition. The main aim of this chapter is, in fact, to compare different acquisition protocols for a pD-SPECT system and to investigate the influence of the presence of activity outside the ROI in the optimisation. This study therefore plays a role in studying the performance of the D-SPECT system, not only for a single phantom but for a class of objects.

The optimum time ratio for the acquisition protocols is a complex function of many factors, including the location of the ROI with respect to the position of the cameras and the relative location and strength of the background features. While the effects of these factors are coupled to each other and therefore difficult to quantify, one may draw a general conclusion from these results: the optimal scanning pattern is sensitive to the level of activity in the background. With increasing time ratio S, the uncertainty in the estimation in the ROI is increasingly more dependent on the presence of activity outside the ROI. If the activity in the background is high with respect to the activity in the ROI , an acquisition that more uniformly scans the whole FOV may be preferable. Therefore calculating the reconstructed image quality in a region of interest (ROI) is a complex problem that depends also on the presence of activity and on the system response outside the ROI. This observation emphasises the need for a fast method to compare different acquisition protocols. The statistical method based on the reconstruction of multiple noise instances is in fact extremely time consuming.

In order to be able to compare a set of candidate scanning patterns we use a deterministic method, based on an approximated expression of the Fisher Information Matrix (FIM). In section 4.5.2, a new formulation that relies on a subsampled version of the FIM has been presented. This formulation reduces the computational complexity in inverting the FIM but nevertheless accounts for the global interdependence between the variables. The results obtained with the approximate subsampled FIM method are compared with the circulant approximation method (see section 4.5.1) and with the reference method based on reconstruction of multiple noise instances. The circulant approximation of the FIM has been used in the past to evaluate adaptive angular sampling in SPECT imaging [Meng and Li, 2011]. However, the incapability of the method to incorporate non-stationary system models and effects of long-range correlations (e.g., evaluation of the effects of data truncation or missing data) has lead to counterintuitive results. In this chapter, the subsampled FIM method has been shown to enable the exploration of design spaces previously precluded by the use of the circulant approximation, such as the evaluation of effects of changing camera trajectory and the optimisation of adaptive data sampling.

The recent development of adaptive SPECT systems has introduced a class of optimisation problems where the parameters of the imaging system may be modified in order to image certain desired properties of the underlying object and in order to adapt, during acquisition, in response to the projection data. The D-SPECT is an example of such a system, where the acquisition protocol (in terms of the trajectory of the cameras) can be modified depending on the data acquired during the scan. In order to adapt the response of the system during acquisition, a set of different design parameters have to be compared in real time. Thanks to the novel approximation of the FIM and thanks to an efficient GPU implementation, our novel algorithm for the estimation of the uncertainty, drastically reduces the computational complexity and therefore is a good candidate method for such optimization problems.

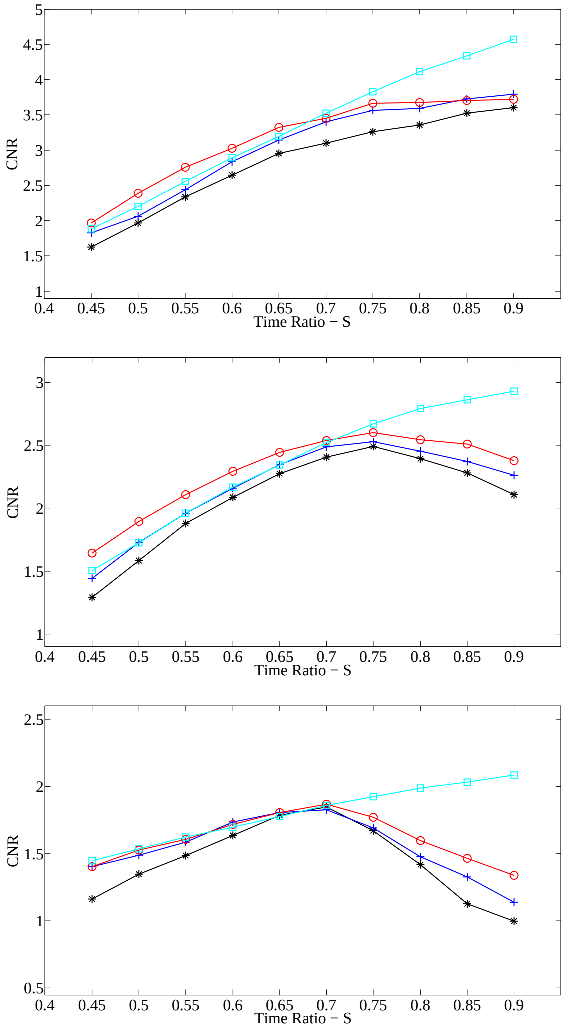

Figure 7.10: CNRs for different scanning patterns of the pD-SPECT system, obtained from the reference method (black line - $*$), from the Subsampled Fisher Information based method with grid g_1 (blue line - $+$), from the Subsampled Fisher Information based method with grid g_2 (red line - $\circ$) and from the circulant approximation method (cyan line - $\square$). The time ratio ranges from $S = 0.45$ to $S = 0.85$. The optimal time ratios are calculated for different level of background $\lambda_{BK1} = 0.9kBq/cm^3$ (A - top figure), $\lambda_{BK2} = 2.2kBq/cm^3$ (B - central figure) and $\lambda_{BK2} = 4.4kBq/cm^3$ (C - bottom figure). The target resolution $P_{target} = 12$ remains the same for all the experiments.

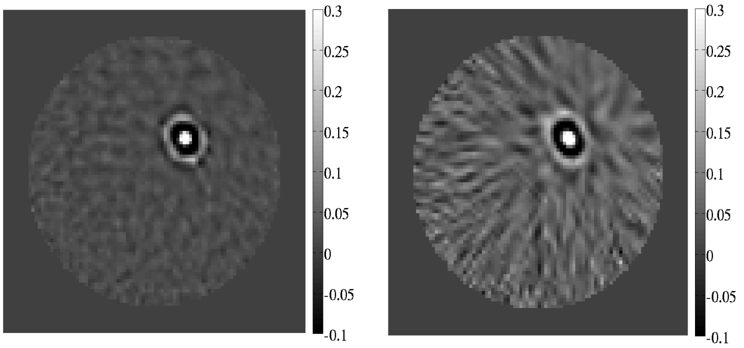

Figure 7.11: Images of covariance for a point in the centre of the uniform sphere for different acquisition protocols of a pD-SPECT system. The covariances are calculated with the reference statistical method. D-SPECT system with different acquisition protocols: time ratio $S = 0.45$ (A - left figure) and time ratio $S = 0.9$ (B - right figure).

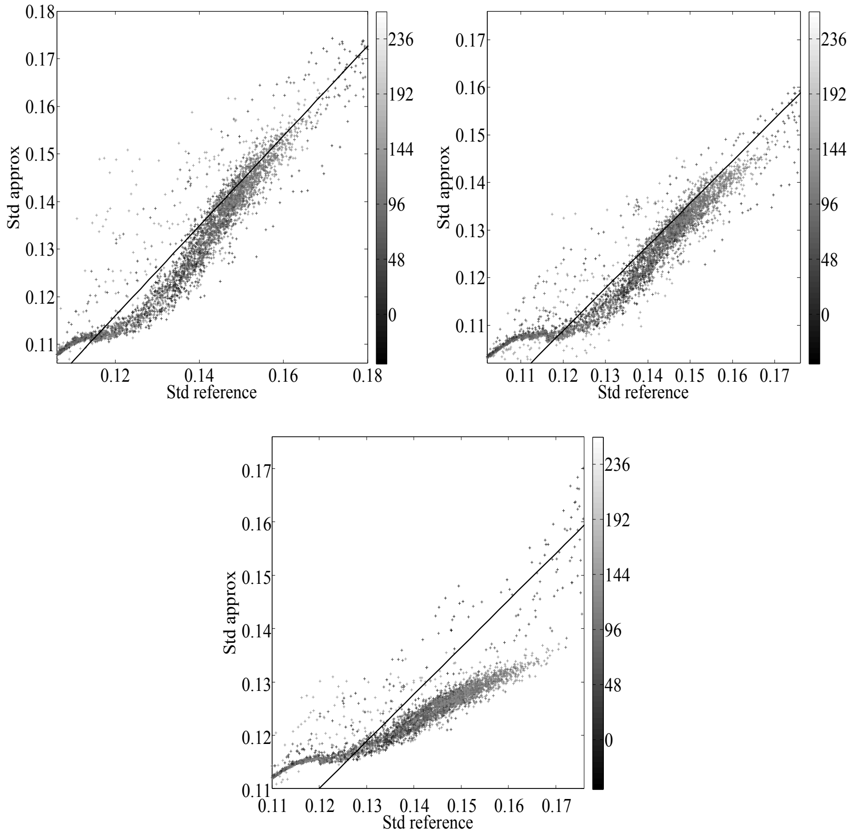

Figure 7.12: Validation of the Fisher information-based standard deviation calculation method for a pD-SPECT acquisition protocol with time ratio $S = 0.45$. The standard deviation obtained with 10240 repeated simulations is plotted with respect to the standard deviation predicted with the Fisher information-based method. A (top left) - grid g_1, B (top right) - grid g_2 and C (bottom)- circulant approximation. Spatial information is given by the colour of the dots. The light grey dots represent voxels which are more distant from the rotation axis (off-centre voxels); whereas the dark grey dots represent voxels which are closer to the rotation axis (central voxels). A colour bar with the respective distance from the rotation axis (in mm) is displayed for every plot.

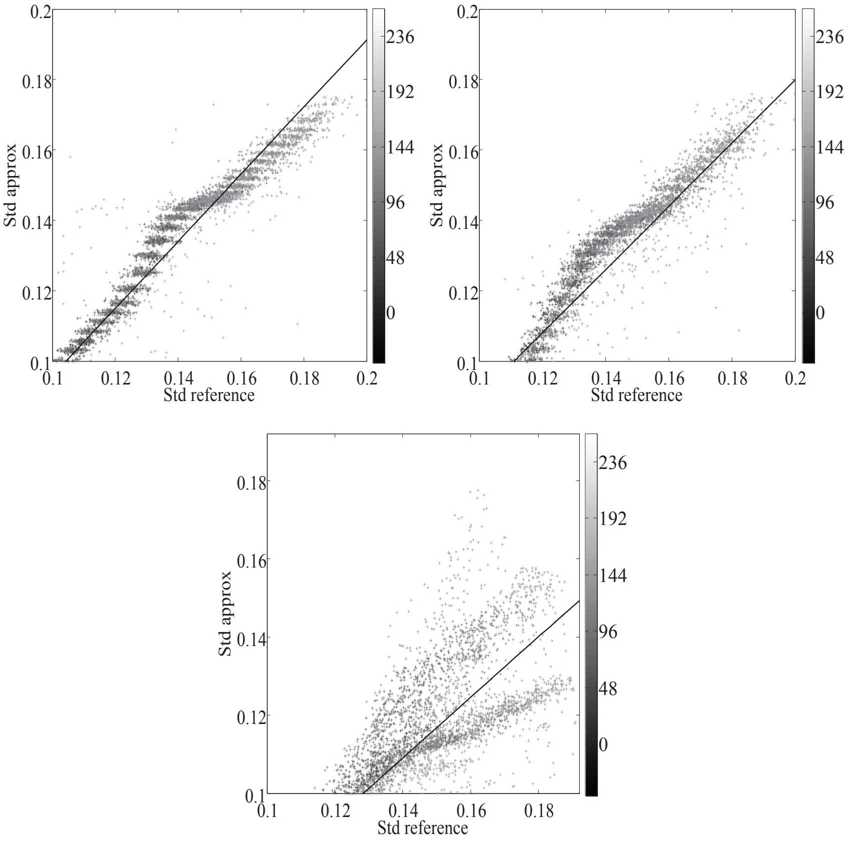

Figure 7.13: Validation of the Fisher information-based standard deviation calculation method for a pD-SPECT acquisition protocol with time ratio $S = 0.9$. The standard deviation obtained with 10240 repeated simulations is plotted with respect to the standard deviation predicted with the Fisher information-based method. A (top left) - grid g_1, B (top right) - grid g_2 and C (bottom)- circulant approximation. Spatial information is given by the colour of the dots. The light grey dots represent voxels which are more distant from the rotation axis (off-centre voxels); whereas the dark grey dots represent voxels which are closer to the rotation axis (central voxels). A colour bar with the respective distance from the rotation axis (in mm) is displayed for every plot.

Chapter 8

Conclusions and Future Work

8.1 Main Contributions and General Discussion

An overview on the main contributions of this PhD work is proposed in this section.

The most important contributions of this thesis are the derivation, validation and application of a novel algorithm for the characterisation of the system design in Emission Tomography. The main aim of this new approach is to introduce a less computationally expensive approximation of the FIM, that still takes into account the global interdependence between the variables. We have applied the method for the estimation of the optimal parameters of a SPECT system, in comparison with both the circulant approximation and the reference statistical method based on the reconstruction of multiple noise instances.

We have pointed out the shortcomings of the circulant approximation for a range of optimisation problems where the system response is markedly shift-variant. Such optimisation problems include the choice of a collimator and the tuning of its parameters and the choice of the scanning parameters of a D-SPECT system. Moreover the sub-sampled FIM method enables the exploration of design spaces previously inaccessible by the circulant approximation, such as the evaluation of effects of data truncation in interior tomographic imaging.

8.1.1 Estimation of the Uncertainty

In chapter 2, a cost function for the PL Estimator $\hat{\boldsymbol{\lambda}}$ has been defined. The absence of a closed analytical formulation that expresses $\hat{\boldsymbol{\lambda}}$ explicitly in terms of $\boldsymbol{\nu}$ makes it difficult to study the properties (e.g. mean and covariance) of the PL estimator $\hat{\boldsymbol{\lambda}}$ defined in section 2.5.1. Thus, in order to compare system designs, one has to compute time-consuming simulations of thousands of reconstructions, as described in chapter 3. The computational complexity of such simulations hinders the on-line optimisation of the parameters of adaptive imaging systems. Alternatively, an approximate estimate of the covariance may be obtained via the FIM, as described in chapter 4. In order to tackle the problem of the computational load in calculating and inverting the FIM, it has been proposed to approximate it with a circulant matrix (see section 4.5.1). The use of the circulant approximation has been explored for the purpose of measuring the image quality in [Fessler and Rogers, 1996] [Qi and Leahy, 2000] [Stayman and Fessler, 2000] and for system design optimisation in [Nuyts, 2009] [Zhou et al., 2010] [Vunckx et al., 2008a]. The computational complexity of the reference method, involving reconstruction to convergence using thousands of noise realisations, has precluded a systematic evaluation of the effect and the limitations of the circulant approximation. For a systematic characterisation of the effect of the circulant approximation on the estimates of the covariance matrix, one would have to consider not only a single phantom, but a class of objects. The problem is further complicated by the choice of the regularisation parameter β.

In this thesis we have described a criterion for the choice of β and a purpose- made GPU accelerated reconstruction software that processes multiple reconstructions in parallel (see chapter 3), enabling the estimation of the reference variance in a reasonably short time (see table 4.1). The circulant FIM is generally considered to yield a good approximation of the covariance matrix for nearly shift-invariant systems, however 3-D imaging systems are inherently shift variant, presenting a block-circulant FIM even in the case of an ideal uniform object in the FOV. Comparison of the variance (the diagonal of the covariance matrix) obtained from the circulant approximation, with the full FIM and with the reference statistical method, has highlighted certain pitfalls of the circulant approximation.

The first contribution of this paper consists in having highlighted these effects. Comparison of the variance (the diagonal of the covariance matrix) obtained from the circulant approximation, with the full FIM and with the reference statistical method, are presented in chapter 5, chapter 6 and chapter 7.

The second and main contribution of this thesis is the introduction of a new approximation which relies on a subsampled version of the FIM and that addresses the shortcomings of the circulant approximation (see section 4.5.2). A comparison between the subsampled FIM method and the circulant approximation of the FIM has also been performed in chapter 5, chapter 6 and chapter 7. The experiments show good results in comparison with the circulant approximation, for the specific choice of the metric of image quality based on the CNR. The results obtained with the subsampled Fisher information matrix, in fact, outperform the results obtained with the circulant approximation, when compared with the reference statistical method. Furthermore, the approximation based on the subsampled FIM shows certain effects that are not visible with the circulant approximation, enabling the exploration of design spaces that were previously forbidden. In particular, the subsampled FIM approach is well-suited for situations where the voxel variance is significantly influenced by activity outside of its neighbourhood, such as when there is high activity in regions of the phantom that are distant from the voxel of interest.

Regarding the relationship between the subsampling and reliability of the estimate of the covariance, we would like to point out that the algorithm described in the paper has been designed to reduce the computational complexity of the estimation of the covariance matrix. It has indeed been designed to enable a fast calculation of the image quality, in order to modify the parameters of an adaptive system during acquisition. The way the proposed algorithm has been designed permits the degree of approximation in the estimation to be defined by the user. Hence a trade-off between computational complexity and reliability of the estimation of the covariance matrix arises.

The sub-sampled FIM formulation and the software tool described in this paper may be employed for the optimisation of a range of design parameters of emission imaging systems. However, it is not possible to define an absolute criterion for the choice of the sub-sampling scheme. In chapter 4, we suggest that a test must be performed for every system under investigation. Moreover we present a criterion for the selection of the grid size under which the subsampled FIM method can provide a reliable estimate of the image quality for varying conditions.

A validation for the subsampled Fisher Information-based variance calculation method was performed for every experiment presented in chapter 5, chapter 6 and chapter 7. For validation purposes, the variance obtained with the reference statistical method was plotted with respect to the variance predicted with the Fisher Information-based method, for grid $g_1 = 27648$ points, grid $g_2 = 6912$ points, grid $g_3 = 3072$ points, grid $g_4 = 1728$ points and grid $g_5 = 1106$ points. A least squares fitting was performed through the data. The regression coefficients, the intercepts of the line, the correlation coefficients and the standard error of the estimate for every experiment were presented in table 5.2, table 6.2 and table 7.1.

The method can an be applied to a variety of systems and design parameters in emission computed tomography. However, evaluation of the trade-off between computational complexity and accuracy of the estimates for the optimum parameters is an open problem and needs to be evaluated case by case.

The code is available on-line (URL: http://niftyrec.scienceontheweb.net) and is open source, in order to foster further development and the evaluation of the algorithm for varying imaging conditions and sub-sampling schemes.

8.1.2 Experimental Findings

/media/LACIE/thesis In chapter 5, chapter 6 and chapter 7 several simulation studies were described, with the purpose of illustrating how the sub-sampled FIM method enables the exploration of design spaces previously forbidden by the circulant approximation. In these chapters, we show how our new algorithm applies to a range of optimisation problems where the system response is markedly shift-variant. Such problems include choice of a parallel hole collimator for SPECT and the tuning of its parameters, as well as the evaluation of the reconstructed image quality in the case of missing projection data and for different acquisition protocols for the D-SPECT system.

Even if the main purpose of the experiments presented in these chapters was to present some possible applications for the novel algorithm and to illustrate the reliability of our approximation, some general conclusions may be drawn from these results.

From the experiment presented in chapter 5, it has been found that the optimal collimator aperture is proportional to the target resolution imposed in the reconstruction. Similar results were presented in an other study [Zhou et al., 2010], where it has been shown that for parallel hole collimators, the FWHM of the aperture that yields the minimal variance equals the desired spatial resolution divided by $\sqrt{3}$ (for volume imaging).

In chapter 6, simulation experiments have been performed to investigate the statistical properties of the estimator under two different truncation cases. The peripheral ROI reconstruction and the interior ROI reconstruction have been investigated simulating different levels of truncation. It has been found that for both truncation cases, a decrease in ROI size leads to an increase in variance, not only outside the ROI but also inside it. To our knowledge, this characterisation of the noise properties of the estimator for truncated acquisition has not been previously reported.

In chapter 7, simulation experiments have been performed to compare different acquisition protocols for a D-SPECT systems and to investigate the influence of the presence of activity outside the ROI in the optimisation. It has been found that the optimal scanning pattern is sensitive to the level of activity in the background. With increasing time ratio S, the uncertainty in the estimation in the ROI is increasingly dependent on the presence of activity outside the ROI.

8.1.3 Summary of the Contributions

A summary of the most important contributions of this PhD work is presented in the following:

- Formulation of the problem concerning the selection of the optimum design of a tomographic imaging system in the information theoretic framework (chapter 3 and chapter 4).

- Development of a purpose-made GPU accelerated reconstruction software that processes multiple reconstructions in parallel, enabling the estimation of the reference variance in a reasonably short time (chapter 3). This reconstruction software has been developed by Stefano Pedemonte [Pedemonte et al., 2010].

- Formulation of a deterministic method, based on a subsampled Fisher Information Matrix, for the efficient estimation of the uncertainty in emission computed tomography (4).

- Investigation of the noise properties of the estimator, in case of interior ROI reconstruction from truncated projection data (chapter 6).

- Optimisation of the acquisition protocols for a D-SPECT system, investigating the influence of the presence of activity outside the ROI in the evaluation of the image quality (chapter 7).

8.2 Suggestions for Future Work

In this section, suggestions are made for improvement of the proposed algorithm and for its future application to the optimisation of novel collimator designs and a novel adaptive system.

8.2.1 Improvement of FIM-based Method

The Fisher information-based method is an efficient and reliable deterministic method to study the properties (e.g. mean and covariance) of the PL estimator $\hat{\boldsymbol{\lambda}}$. However, this method presents some limitations due to the approximations made during its derivation. One of the major limitations is that the calculation of the uncertainty based on the Fisher Information is restricted to problems where $\boldsymbol{\lambda}$ is a continuous parameter in $\mathbb{R}^N$, thus, strictly speaking, its results are not comparable with methods where non-negativity constraints are imposed on $\hat{\boldsymbol{\lambda}}$. However, for cost functions that include an appropriate penalty function for regularisation, non-negativity constraints are active relatively infrequently. Negative values can be avoided by keeping the weight of the prior beta low enough [Bruyant, 2002]. An approach to tackle this problem has been presented by Li et al. [2004], with assumptions that still affect the approximation accuracy. How to overcome this limitation is thus still an open question.

The sub-sampled FIM trades off computational complexity and accuracy of the estimation, enabling the adaptation of the accuracy of the estimation based on the available computational resources. When sufficient resources are available, the GPU-accelerated software (described in section 4.5.2) can compute the full covariance matrix exactly on a grid g_1. One important advantage of the scalable sub-sampled FIM approximation is that the algorithm provides an estimate of the full covariance matrix, though sub-sampled, accounting for the global interdependence between the variables of the tomogram. This enables the use of global metrics for system design optimisation. In other fields of imaging, where the lesser dimensionality of the parameter space enables the storage and inversion of the full FIM, a wide range of global optimality criteria has been explored, such as D-optimality [Delzell et al., 2012] and I-optimality [Khodja et al., 2012]. In the future, a global figure of merit that can account for the off-diagonal entries of the FIM will be investigated.

The experiments presented in this paper account for a uniform attenuation map. If the attenuation map or model for randoms and scatter are available, they can be included in the calculation of the Fisher Information Matrix to study their effect on image quality.

The proposed approximation for the prediction of the covariance matrix can also be useful in the context of task specific system optimisation strategies. In these studies the image quality assessment is based on the performance of numerical observers in classification, such as in the detection of a specific class of tumours. The ideal observer requires the computation of the inverse covariance matrix. The circulant approximation of the Fisher information matrix has been applied in the domain of numerical observers before [Yendiki and Fessler, 2006]. It would therefore be interesting to compare the circulant FIM approximation with the subsampled FIM approximation when using this particular figure of merit for image quality evaluation.

It might also be interesting to compare the subsampled FIM approach with the "small ROI method" proposed in [Cloquet et al., 2010], which avoids the computation of the entire Fisher information matrix by evaluating the Fisher elements only for voxels within a neighbourhood of a point of interest. This formulates invertible sub-matrices of the Fisher information matrix, which are then used to estimate the best achievable variance.

In addition, a comparison between the subsampled FIM approach with the Non Uniform Object-Space Pixelation (NUOP) approach is being performed in collaboration with the University of Ghent. In the NUOP method, suggested by Meng and Li [2009], the image-space is divided in non-uniformly sized "voxels". In particular, the voxel size increases as we go to regions further away from the region of interest. In this way we keep approximately the same image quality in the ROIs as when the image is divided into voxels with the smallest size considered, but the system modelling becomes much more efficient. In other words, this means that we can reduce the size of the image-space from N to N_{nuop}, with $N_{nuop} < N$, thereby reducing the size of the FIM to $N_{nuop} \times N_{nuop}$, while still obtaining a good evaluation of image quality in the ROIs. To apply the NUOP, we start with a uniform image-space, of size N, whose voxels have the smallest size that we want to consider. Afterwards the image-space is divided in regions according to how many voxels we want to group together to form larger voxels (the rebinning strategy). Neighbouring voxels within the same region and with the same activity value are grouped together to form a cube (when possible), and this larger voxel is given a new index $m \in \{1, \ldots, N_{nuop}\}$ in the non-uniform image-space. For a more detailed explanation of the algorithm see [Meng and Li, 2009].

This method can significantly increase the efficiency of the calculation of the variance at voxel i, since it only requires the inversion of a $N_{nuop} \times N_{nuop}$ matrix. Note that in this formulation only voxels with the same activity value can be grouped together. As such, the method is specific to each phantom, and to get a significant speed up in the case of a realistic phantom we need to first approximate the phantom to have reasonably large regions with uniform activity values and only then apply the approximation, which can be a downside of the approach.

In essence, the difference between the two methods is that in the subsampled approach the approximation is made at the level of the FIM, by only keeping some of the matrix elements, whereas in the NUOP it is made at the level of the image-space (in fact, it was originally presented as a method to speed up reconstruction). The subsampled algorithm also does not require patches of nearly uniform activity in the image, although the more variability in the FIM the less accurate the approximation will be.

8.2.2 Design of Novel Collimators

In chapter 5 the use of the subsampled FIM approximation has been explored for the optimisation of parallel hole collimators in SPECT; emphasising how it enables us to explore the design of highly shift variant systems as a result of distance dependent resolution. In appendix A a method for the investigation of new collimator design is introduced and the implementation of a ray-tracing algorithm is described. This ray-tracing algorithm can be used to model the geometric response and the septal penetration of novel collimator geometries.

Traditionally collimators are fabricated using folded sheets of foil, or are cast using moulding methods. The recent development of novel production techniques [Abe et al., 2009] has introduced the possibility to design collimators, for clinical SPECT, with septa geometries that differ from the conventional parallel holes.

In the future, the efficient method for the estimation of the image quality presented in this thesis and the ray-tracing algorithm presented in appendix A will be used to investigate the design of novel collimator geometries, optimised for a range of different phantoms.

8.2.3 Interior Problem

In chapter 6 the deterministic FIM-based method has been used to investigate the noise properties of the estimator in case of interior ROI reconstruction from truncated projection data.

Simulation experiments have been presented to investigate the statistical properties of the estimator under two different truncation cases: the peripheral ROI reconstruction and the interior ROI reconstruction. It has been shown that even if peripheral ROI reconstruction can lead to nearly unbiased reconstruction, the interior ROI reconstruction results in images that suffer from artifacts.

The presence of artifacts is expected to be reflected in the covariance matrix. However, there is no straight- forward way to estimate or quantify the artifacts directly from covariance matrix. It would therefore be interesting to develop a methodology to obtain quantitative evaluation of the artifacts with an efficient deterministic method.

8.2.4 D-SPECT System optimization

In chapter 7, the novel design of a D-SPECT system has been presented. The D-SPECT camera provides a novel photon collection method and scanning geometry that permits the independent movement of multiple detector columns in order to achieve a region-centric acquisition. The D-SPECT system has been developed to improve the trade-off between spatial resolution and sensitivity inherent in conventional SPECT systems.

However, the system is designed specifically for cardiac SPECT use. The advantages in resolution and sensitivity of the D-SPECT camera reported from cardiac imaging may not necessarily hold as other organs are imaged.

In chapter 7, the deterministic method, that relies on a subsampled version of the FIM, has been used to compare different acquisition protocols for a D-SPECT system and to investigate the influence of the presence of activity outside the ROI in the optimisation. This study therefore plays a role in studying the performance of the D-SPECT system, not only for a single phantom but for a class of objects.

This efficient method will be used in the future to investigate novel D-SPECT collimator geometries for a range of objects. Moreover the subsampled FIM-based method will be used to design the scan pattern of each single detector individually and to adapt the system design, in real time, during acquisition.

8.2.5 Adaptive Systems

The recent development of adaptive SPECT systems has introduced a class of optimisation problems where the parameters of the imaging system may be modified in order to image certain desired properties of the underlying object and in order to adapt, during acquisition, in response to the projection data. The D-SPECT is an example of such a system, where the acquisition protocol (in terms of the trajectory of the cameras) can be modified depending on the data acquired during the scan.

However, a set of different design parameters has to be compared in real time in order to adapt the response of the system during acquisition. Thanks to the novel approximation of the FIM and thanks to an efficient GPU implementation, our novel algorithm for the estimation of the uncertainty drastically reduces the computational complexity and is therefore a good candidate method for such optimisation problems.

However, the deterministic FIM-based method poses a problem when used for the optimisation of adaptive systems in real time. The first assumption we make in (4.30), in order to calculate the FIM, is that the reconstruction is locally linear, meaning that the mean of the noisy reconstruction can be well estimated by the reconstruction of noiseless data. This in turns means that we need to know the activity distribution in advance to be able calculate the FIM. Fessler and Rogers [1996] argued that even for real noisy measurements we can predict the variance simply by replacing $\overline{\nu}$ with ν in (4.30). However this approximation may be problematic and eventually cause convergence problems when the scanning parameters are updated iteratively. The optimisation of adaptive systems is thus still an open problem and it needs further investigation.

8.3 Publications Arising from Thesis Work

8.3.1 International Journals

- **N. Fuin**, S. Pedemonte, S. Arridge, S. Ourselin, B. Hutton. Efficient Determination of the Uncertainty for the Optimization of SPECT System Design : A Subsampled Fisher Information Matrix. IEEE Transactions on Medical Imaging (Accepted pending minor revision).

8.3.2 International Conference Proceedings

- **N. Fuin**, S. Pedemonte, S. Arridge, S. Ourselin, B. Hutton. Use of the Fisher Information Matrix to Optimize the Acquisition Protocol for a D-SPECT System. IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2012, pp. 2137 - 2142.

- **N. Fuin**, S. Pedemonte, S. Arridge, S. Ourselin, B. Hutton. Subsampled Fisher Information Matrix for Efficient Estimation of the Uncertainty in Emission Tomography. IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2012, pp. 2251 - 2256.

- **N. Fuin**, A. Bousse, S. Pedemonte, S. Arridge, S. Ourselin, B. Hutton. Collimator Design in SPECT, an Optimisation Tool. IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2010, pp. 2061 - 2065 .

- A. Bousse, **N. Fuin**, K. Erlandsson, S. Pedemonte D. Kazantsev S. Ourselin, S. Arridge B. Hutton. Point Spread Function Optimization in SPECT. IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2010, pp. 3149 - 3154.

Appendix A

Methodology for the Optimisation of Novel Collimator Design

In chapter 5 the use of the subsampled FIM approximation has been explored for the optimisation of parallel hole collimators in SPECT; emphasising how it enables us to explore the design of highly shift variant systems as a result of distance dependent resolution. Thanks to an efficient GPU implementation, the novel algorithm for the estimation of the uncertainty, drastically reduces the computational complexity in calculating the FIM and its inverse, making the algorithm a good candidate for fast collimator design optimisation.

In this appendix a method for the investigation of new collimator design is introduced. Although no results have been obtained yet, this method allows us to accelerate the optimisation process, permitting the investigation of completely novel collimator geometries whose properties are not described analytically. In section A.2, we describe the implementation of a ray-tracing algorithm that can be used to model the geometric response and the septal penetration of novel collimator geometries. In section A.3, we describe the use of a genetic Algorithm (GA) to determine the optimal collimator design . A GA was selected, among other optimisation techniques (i.e. Powell's method or Simulated Annealing), as it gives us a greater chance of finding an optimal solution when the optimisation space presents multiple local minima or is a curved flat valley.

Traditionally collimators are fabricated using folded sheets of foil, or are cast using moulding methods. The recent development of novel production techniques [Abe et al., 2009] has introduced the possibility to design collimators, for clinical SPECT, with septa geometries that differ from the conventional parallel holes.

In the future, the efficient method for collimator design presented in this appendix will be used to investigate the design of novel and unforeseen collimator geometries, optimised for a range of different phantoms.

A.1 Design Parameters and Cost Function

In chapter 5, we discussed the optimisation of the geometrical parameters of parallel hole (PH) collimators that affect geometric resolution and detection sensitivity. The PH collimator usually consists of a two-dimensional array of parallel holes and variation between collimators is confined to the geometric dimensions that describe the holes.

In order to design a novel collimator, we allow for a wider choice of the collimator's geometry parameters. First, we propose a single layer collimator geometry that does not necessarily imply parallel holes, but could involve a polygonal shape of the septa. The parameters that define the optimisation space are: the length of the holes through the collimator l; the thickness of the septa on the side of the collimator that faces the scintillation crystal q_1; the thickness of the septa on the external side of the collimator q_2; the shorter diameter of the holes v; and, finally, the material used to make the collimator. It can be noted that the septa can have any polygonal shape. Second, a multilayer geometry has been proposed. The additional parameters that define the optimisation space in this case are: the number of collimator layers n_l and the distances between those layers o. The constraints imposed to the dimensions are due to: weight constraints imposed by the camera gantry, the minimal septa thickness imposed by the limitations of fabrication and the visibility of collimator hole pattern in the images. Some constraints, with regard to geometrical symmetry, can be also imposed in order to maintain spatial invariance of the PSF along planes parallel to the detector surface. Changing the aforementioned design parameters to find new and possibly unforeseen collimator designs, one would probably lead to a change in other parameters, including e.g. the (local) convergence of the collimator. That would imply a projector/backprojector that can deal with all these potential collimators. While for parallel-hole SPECT the implementation of a projector/backprojector is relatively simple, it can be more complex for alternative geometries. In order to address this problem, an approach, based on angular rebinning, to compute the projector/backprojector operator for any SPECT system has been proposed in [Bousse et al., 2013].

A cost function for design optimisation, based on the trade-off between bias and variance of the estimator, was introduced in section 4.6.1. In order to compare different systems parameterisation, a target bias gradient is defined. Consequently, for every system parameterisation, a specific post-smooth filter is designed and the variance of the estimator can be compared at equal bias gradient. Thus, the cost function is given by the variance of the estimator at a fixed target bias gradient.

A flowchart of the optimisation method presented in this appendix is shown in Figure A.1. Given the set of design parameters and the cost function described above; a Genetic Algorithm (GA) was selected to stochastically guide the algorithm through the solution space to the optimal design of the collimator. Note that in order to obtain the fitness score for each set of parameters, we have to perform the whole ray-tracing process described in the next section.

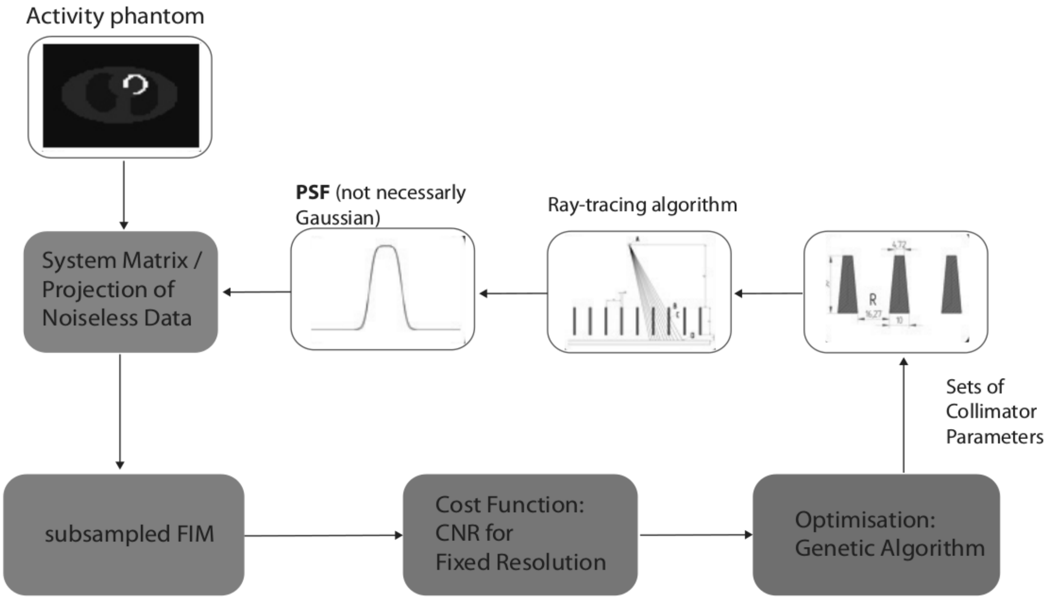
Activity phantom
System Matrix / Projection of Noiseless Data
PSF (not necessarly Gaussian)
Ray-tracing algorithm
4.72
R
9.27
10
Sets of Collimator Parameters
subsampled FIM
Cost Function: CNR for Fixed Resolution
Optimisation: Genetic Algorithm

Figure A.1: Flow chart of the method for optimisation of novel collimator designs.

A.2 Ray-Tracing Algorithm

As already described in chapter 2, the uncertainty about the origin of the detected photons is modelled by a Point Spread Function (PSF). The PSF of a gamma camera describes the photon count density distribution at the detector surface when a point source is imaged.
In orderer to obtain the PSF for novel collimator geometries, we introduce a ray-tracing algorithm that models the detection efficiency, the geometric response and the septal penetration in the collimator, since no analytical treatment of this effect appears to exist in the literature (except for standard collimators).

The model that we use consists of a point source above the collimator and a simple detector below the collimator which detects all photons that strike it. The probability that such a photon will actually reach the image plane is determined by the distance travelled through the collimator septum and the attenuation coefficient of the collimator material. The septal scatter on the contrary is not taken into account.
The fundamental idea of the ray-tracer program is shown in Figure A.2. A photon emitted by the source 'A', enters the collimator at point 'B', travels through part of a collimator septum, exits the collimator at 'C', and finally is imaged by the scintillation camera at point 'D' [Han et al., 1999]. By tracing rays through the collimator and finding the distance travelled in the septum, the flux of γ-photons that reach the detector 'D' is determined. Afterwards, in order to account for the intrinsic response of the detector, the obtained collimator PSF is convolved with the intrinsic gamma-camera's PSF. The intrinsic PSF is usually well approximated by a radially symmetric Gaussian function, with $FWHM = 3.6\ mm$ as per current Anger cameras; although a different function could be determined for alternative detectors.
The PSF of the collimator, for a single source-to-detector distance, is found by repeating this process for more than 4 millions rays, discretised by angle of emission at the point source. Moreover for a complete characterisation of the system response the PSF has to be calculated for several point source-to-detector distances, hence the need for a fast algorithms where ray-tracing process is performed in parallel.
The algorithms have been implemented in the CUDA programming language for parallel execution on Graphics Processing Units (GPU). The ray-tracing algorithm implemented in NiftyRec [Pedemonte et al., 2010] is based on the efficient ray-box intersection algorithm described on the

Siggraph Education web-site (http://www.siggraph.org/education/ materials/HyperGraph/raytrace/rtinter3.htm). Each GPU thread casts one ray from the point source to one of the detector pixels, computes the coordinates of the intersection with the volume of the septa and integrates the attenuation coefficient along the ray, re-sampling with tri-linear interpolation the attenuation coefficient at regular intervals.

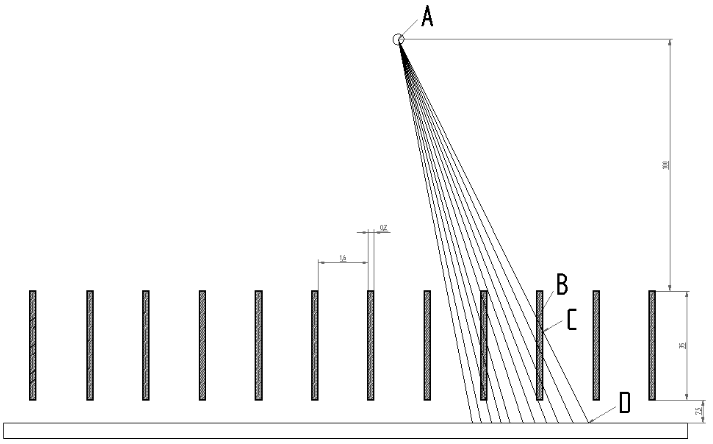

Figure A.2: Raytracer: rays are traced from the source to the detector.

The results obtained using the ray-tracing algorithm described above are in agreement with measurements of FWHM from a LEHR collimator (GE Healthcare), within an accuracy of 2%.

A.3 Optimisation Using Genetic Algorithm

Upon review of the optimisation techniques available in the literature, a Genetic Algorithm (GA) was selected to stochastically guide the algorithm through the solution space to the optimal design of the collimator.

Hollande [1975] showed that a GA combines both exploration (random search) and exploitation (hill-climbing) at the same time, in an optimal way. Combinations of those two strategies (Simulated Annealing) can be quite effective, but it is difficult to know where the best balance lies. Another potentially useful point is that genetic algorithm is intrinsically parallel. Most other algorithms are serial (Gradient Descent and Powell algorithm has been used) and can only explore the solution space to a problem in one direction at time. GA, in the contrary, can explore the solution space in multiple directions at once, giving a greater chance of finding an optimal solution when the optimisation space presents multiple local minima or is a curved

flat valley. The GA also enables the creation of virtual entities without requiring an understanding of the procedures or parameters used to generate them, which is useful when we are solving a global optimisation problem [Beasley et al., 1993].
The creation of the initial population of our GA does not involve any specific initialisation. The only constraints imposed to the dimensions are due to limitations of fabrication or are related to the visibility of the collimator hole pattern in the images.

A flow chart of the algorithm is shown in Figure A.3. The GA uses populations of genotypes consisting of strings of binary digits. Those populations are usually read-out for producing an offspring, which is then evaluated according to a fitness criteria before being selectively reproduced. At each step a ray-tracing process is performed to obtain a PSF, which is then used to evaluate the cost function (variance of the estimator at a fixed target resolution) for each set of parameters (genotypes). Giving more probability of selection to high-scored genotypes, these are selected for a subsequent genetic manipulation process consisting of two steps. In the first step, the crossover operation, recombining the dimensions (genes) of each two selected genotypes (chromosomes), is executed. Various types of crossover operators are found in the literature. For this study, the single point crossover operation was selected for use. During the second step, the dimension at one or more randomly selected positions of the chromosomes are altered, this is known as mutation. The mutation process helps overcome trapping at local minima. The offspring produced by the genetic manipulation process are the next population to be evaluated. The cycle of evolution is then repeated until convergence.

Because of a high rate of correlation among the parameters, a high number of individuals in the population is required [Goldberg, 1989] (400 population individuals, 20 elite individuals). Evolution was tested with different values of crossover fraction and mutation rate, in order to find a set of parameters that avoid genetic drift and provide a better convergence. Furthermore, in order to avoid premature convergence or slow finishing, the parent selection technique of choice is the fitness ranking where the individuals are sorted in order of raw fitness, and then reproductive fitness values are assigned according to rank [Baker, 1985].

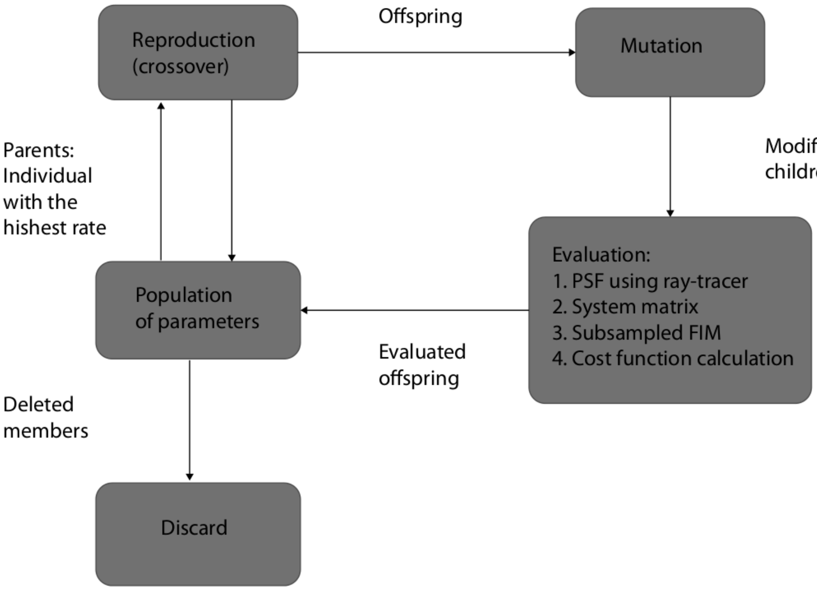
Reproduction (crossover)
Offspring
Mutation
Parents: Individual with the hishest rate
Modified children
Population of parameters
Evaluation:
1. PSF using ray-tracer
2. System matrix
3. Subsampled FIM
4. Cost function calculation
Evaluated offspring
Deleted members
Discard

Figure A.3: Flow chart of the Genetic Algorithm.

A.4 Conclusion

In the last few decades, the optimisation of the collimator has essentially proceeded by trial and error: certain designs that have proven good image quality for certain imaging conditions, such as a certain energy range, have been designed and standardised [Keller, 1994]. Only in the last few years the problem has been treated systematically, driven by the need to compare and optimise collimator designs prospectively, by computer simulation.

The development of a new method for the estimation of the uncertainty allows us to accelerate the optimisation process, permitting the fast calculation of a cost function. The ray-tracing algorithm models the detection sensitivity and the collimator response in software simulations, permitting the investigation of completely novel collimator geometries whose properties are not described analytically. Moreover, the use of a genetic algorithm allows us to optimise multiple collimator parameters at the same time without making assumptions about which one takes priority.

All the different algorithms that compose this method exploit the use of parallel coding and have been implemented in the CUDA programming language for execution on GPU, thereby considerably reducing the computational complexity. In the future, this efficient method will be used to investigate the design of novel collimator geometries, optimised for a range of different phantoms.

References

Abe, F., Osakada, K., Shiomi, M., Matsumoto, M., and Shiomi, M. (2009). The Manufacturing of Hard Tools from Metallic Powders by Selective Laser Melting. *J. Mater. Proc. Tech.*, 50(4):635–643.

Ahn, S. and Fessler, J. (2003). Globally convergent image reconstruction for emission tomography using relaxed ordered subsets algorithms. *Trans. Med. Imag.*, pages 101–111.

Alpert, N., Chesler, D., Correia, J., Ackerman, R., Chang, J., Finklestein, S., Davis, S., Brownell, G., and Taveras, J. (1982). Estimation of the Local Statistical Noise in Emission Computed Tomography. *IEEE Trans. Med. Imag.*, 1(2):142–146.

Anger, H. (1964). Scintillation Camera with Multichannel Collimators. *J. Nucl. Med.*, 5(7):515–531.

Baker, J. (1985). Adaptive selection methods for genetic algorithms. *Proc. of the 1st Int. Conf. on Gen. Alg.*, pages 101–111.

Barrett, H., Abbey, C., and Clarkson, E. (1998). Objective Assessment of Image Quality. III. ROC Metrics, Ideal Observers, and Likelihood Generating Functions. *J. Opt. Soc. Amer. A Opt. Image Sci.*, 1(15):1520–1535.

Barrett, H., Denny, J., Wagner, R., and Myers, K. (1995). Objective Assessment of Image Quality. II. Fisher Information, Fourier Crosstalk, and Figures Of Merit for Task Performance. *J.Opt. Soc. Am.*, 29(A):834 – 852.

Barrett, H., Myers, J., Devaney, N., and Dainty, C. (2006). Objective Assessment of Image Quality. IV. Application to Adaptive Optics. *J. Opt. Soc. Amer. A Opt. Image Sci.*, 1(23):3080–3105.

Barrett, H. and Myers, K. (2004). *Foundations of Image Science*. Princeton : Princeton University Press.

Barrett, H., Wilson, D., and Tsui, B. (1994). Noise Properties of the EM Algorithm: I. Theory. *Phys. Med. Biol.*, 39(5):833–846.

Beasley, D., Bull, D. R., and Martin, R. R. (1993). An overview of genetic algorithms: Part 1, fundamentals. *Univ. Comp.*, 15:58–69.

Beck, R. and Redtung, L. (1985). Collimator Design Using Ray-Tracing Techniques. *IEEE Trans. Nucl. Sci.*, 32(1):865–869.

Beekman, F., Slijpen, E., and Niessen, W. (1998). Selection of Task-Dependent Diffusion Filters for the Post-Processing of SPECT Images. *Phys. Med. Biol.*, 43(6):1713–1730.

Bousse, A., Erlandsson, K., Pedemonte, S., Ourselin, S., Arridge, S., and Hutton, B. (2013). Angular Rebinning for Geometry Independent SPECT Reconstruction. *Proc. Fully 3-D Imag. Reconstr. Conf.*

Brill, A. and Beck, R. (2004). *Evolution of Clinical Emission Tomography*. Academic Press, San Diego.

Bruyant, P. P. (2002). Analytic and Iterative Reconstruction Algorithms in SPECT. *J. Nucl. Med.*, 43(10):1343–1358.

Butler, J., Lingren, C., Friesenhahn, S., Doty, F., Ashburn, W., Conwell, R., Augustine, F., Apotovsky, B., Collins, B., Zhao, S., and Isaacson, C. (2001). CdZnTe Solid-State Gamma Camera. *IEEE Trans. Nucl. Sci.*, 48(4):950–959.

Carson, R., Yan, Y., Chodkowski, B., Yap, T., and Daube-Witherspoon, M. (1994). Precision and Accuracy of Regional Radioactivity Quantitation Using the Maximum Likelihood EM Reconstruction Algorithm. *IEEE Trans. Med. Imag.*, 13(3):526–537.

Chesler, D., Riederer, S., and Pelc, N. (1977). Noise Due to Photon Statistics in Computed X-Ray Tomography. *J. Comp. Assisted Tomo.*, 1(1):64–74.

Clackdoyle, R. and Noo, F. (2004). A Large Class of Inversion Formulae for the 2D Radon Transform of Functions of Compact Support. *Inv. Prob.*, 20(4):1281–1291.

Clackdoyle, R., Noo, F., Guo, J., and Roberts, J. (2004). A Quantitative Reconstruction from Truncated Projections in Classical Tomography. *IEEE Trans. Nucl. Sci.*, 51(5):2570–2578.

Clarkson, E., Palit, R., and Kupinski, M. (2010). SVD for Imaging Systems with Discrete Rotational Symmetry. *Opt. Exp.*, 18(24):25306–25320.

Cloquet, C., Goldman, S., and Defrise, M. (2010). Cramer-Rao Bound for Gated PET. *Proc. IEEE Nucl. Sci. Symp. Med. Imag. Conf.*, pages 2267–2272.

Cot, A., Sempau, J., Pareto, D., Bullich, S., Pavia, J., and Calvino, F. (2002). Evaluation of the Geometric, Scatter, and Septal Penetration Components in Fan-Beam Collimators Using Monte Carlo Simulation. *IEEE Trans. Nucl. Sci.*, 49(1):12–16.

Cramér, H. (1946). *Mathematical Methods of Statistics*. Wiley, New York.

Defrise, M., Noo, F., Clackdoyle, R., and Kudo, H. (2006). Truncated Hilbert Transform and Image Reconstruction from Limited Tomographic Data. *Inv. Prob.*, 1(22):1037–1053.

Delzell, D., Gunst, R., Schucany, W., Carmack, P., Lin, Q., Spence, J., and Haley, R. (2012). Key Properties of D-Optimal Designs for Event-Related Functional MRI Experiments With Application to Nonlinear Models. *Stat. Med.*, 31(29):3907–3920.

Erlandsson, K., Kacperski, K., Gramberg, D. V., and Hutton, B. (2009). Performance Evaluation of D-SPECT: a Novel SPECT System for Nuclear Cardiology. *Phys. Med. Bio.*, 1(54):2635–2649.

Fessler, J. (1996). Mean and Variance of Implicitly Defined Biased Estimators (Such as Penalized Maximum Likelihood): Applications to Tomography. *IEEE Trans. Imag. Proc.*, 5(3):493–506.

Fessler, J. (1998). Spatial Resolution and Noise Tradeoffs in Pinhole Imaging System Design: A Density Estimation Approach. *Opt. Exp.*, 2(6):237–253.

Fessler, J. (2012). Image Reconstruction Toolbox. *URL: http://web.eecs.umich.edu/ fessler/irt/irt/.*

Fessler, J. and Rogers, W. (1996). Spatial Resolution Properties of Penalized-Likelihood Image Reconstruction: Space-Invariant Tomographs. *IEEE Trans. Imag. Proc.*, 9(5):1346–1358.

Gambhir, S., Berman, D., Ziffer, J., Nagler, M., Sandler, M., Patton, J., Hutton, B., Sharir, T., Haim, S., and Haim, S. (2009). A Novel High-Sensitivity Rapid-Acquisition Single-Photon Cardiac Imaging Camera. *J. Nucl. Med.*, 50(4):635–643.

Geman, S. and McClure, D. (1985). Bayesian Image Analysis: an Application to Single Photon Emission Tomography. *Proc. of Stat. Comp. Sect. of Amer. Stat. Assoc.*, pages 12–18.

Gifford, H., King, M., Pretorius, P., and Wells, R. (2005). A Comparison of Human and Model Observers in Multislice LROC Studies. *IEEE Trans. Med. Imag.*, 1(24):160–169.

Gilland, K., Tsui, B., Qi, Y., and Gullberg, G. (2006). Comparison of Channelized Hotelling and Human Observers in Determining Optimum OS-EM Reconstruction Parameters for Myocardial SPECT. *IEEE Trans. Nucl. Sci.*, 53(3):1200–1204.

Goldberg, D. (1989). Sizing populations for serial and parallel genetic algorithms. *Proc. of the 3rd Int. Conf. on Gen. Alg.*, 22(5):613–626.

Green, P. (1990). Bayesian Reconstructions From Emission Tomography Data Using a Modified EM Algorithm. *IEEE Trans. Med. Imag.*, 9(1):84–93.

Han, G., Liang, Z., and You, J. (1999). A Fast Ray-Tracing Technique for TCT and ECT Studies. *Proc. IEEE Nucl. Sci. Symp.*, 3:1515–1518.

Hebert, T. and Leahy, R. (1989). A Generalized EM Algorithm for 3-D Bayesian Reconstruction from Poisson Data Using Gibbs Priors. *IEEE Trans. Med. Imag.*, 8(2):194–202.

Hebert, T. J. (1990). Statistical Stopping Criteria for Iterative Maximum Likelihood Reconstruction of Emission Images. *Phys. Med. Biol.*, 35(9):1221–1232.

Helmer, R., Heath, R., Putnam, M., and Gipson, D. (1967). Photopeak Analysis Program for Photon Energy and Intensity Determinations: Ge (Li) and NaI (Tl) Spectrometers. *Nucl. Inst. and Meth.*, 57:46–57.

Henkin, R., Boles, M., Dillehay, G., Halama, J., karesh, S., Wagner, R., and Zimmer, A. (1996). *Nuclear Medicine*. Mosby: St. Louis, MO.

Hero, A., Fessler, J., and Usman, M. (1996). Exploring Estimation Bias-Variance Tradeoffs Using the Uniform CR Bound. *IEEE Trans. Sign. Proc.*, 44(8):2026 – 2040.

Hollande, J. (1975). *Adaptation in Natural and Artificial Systems*. MIT Press.

Hudson, H. and Larkin, R. (1994). Accelerated Image Reconstruction Using Ordered Subsets of Projection Data. *IEEE Trans. Med. Imag.*, 13(4):601–609.

Jacobson, M., Levkovitz, R., Ben-Tal, A., Thielemans, K., Spinks, T., Pagani, D., Bettinardi, V., Gilardi, M., Zverovich, A., and Mitra, G. (2000). Enhanced 3D PET OSEM Reconstruction Using Inter-Update Metz Filtering. *Phys. Med. Biol.*, 45(8):2417–2439.

Jan, S., Santin, G., Strul, D., Staelens, S., Assi, K., Autret, D., Avner, S., Barbier, R., Bardis, M., Bloomfield, P., Brasse, D., Breton, V., Bruyndonckx, P., Buvat, I., Chatziioannou, A., Choi, Y., Chung, Y., Comtat, C., Donnarieix, D., Ferrer, L., Visvikis, D., de Walle, R. V., Wieers, E., and Morel, C. (2004). GATE: a Simulation Toolkit for PET and SPECT. *Phys. Med. Biol.*, 49(19):4543–4561.

Jaszczak, R., Floyd, C., Manglos, S., Greer, K., and Coleman, R. (1986). Cone Beam Collimation for Single Photon Emission Computed Tomography: Analysis, Simulation, and Image Reconstruction Using Filtered Backprojection. *Med. Phys.*, 13(5):484–493.

Jorgensen, A. and Zeng, G. (2008). SVD-Based Evaluation of Multiplexing in Multipinhole SPECT Systems. *Int. J. Biom. Imag.*, (769195).

Keller, E. (1994). Optimum Dimensions of Parallel-Hole, Multi-Aperture Collimators for Gamma-Ray Cameras. *J. Nucl. Med.*, 9:233–235.

Khodja, M., Prange, M., and Djikpesse, H. (2012). A Heuristic Bayesian Design Criterion for Imaging Resolution Enhancement. *IEEE Stat. Sig. Proc. Workshop (SSP)*, pages 9–12.

Khurd, P. and Gindi, G. (2005). Decision Strategies that Maximize the Area Under the LROC Curve. *IEEE Trans. Med. Imag.*, 1(24):1626–1636.

Lange, K. (1990). Convergence of EM Image Reconstruction Algorithms with Gibbs Smoothing. *IEEE Trans. Med. Imag.*, 9(4):439–46.

Li, Q., Asma, E., Qi, J., Bading, J., and Leahy, R. (2004). Accurate Estimation of the Fisher Information Matrix for the PET Image Reconstruction Problem. *IEEE Trans. Med. Imag.*, 23(9):1057–1062.

Maass, P. (1992). The Interior Radon Transform. *SIAM J. Appl. Math.*, 52(3):710–724.

Macovski, A. (1983). *Medical Imaging Systems*. Prentice-Hall, New Jersey.

Madsen, R. (2007). Recent Advances in SPECT Imaging. *J. Nucl. Med.*, 48(2):661–673.

Matej, S. and Lewitt, R. (1996). Practical Considerations for 3-D Image Reconstruction Using Spherically Symmetric Volume Elements. *IEEE Trans. Med. Imag.*, 15(1):68–78.

Meng, L. and Clinthorne, N. (2004). A Modified Uniform Cramer-Rao Bound for Multiple Pinhole Aperture Design. *IEEE Trans. Med. Imag.*, 23(7):896–902.

Meng, L. and Li, N. (2000). A Vector Uniform Cramer-Rao Bound for SPECt System Design. *IEEE Trans. Nucl. Sci.*, 56(1):81–90.

Meng, L. and Li, N. (2009). Non-Uniform Object-Space Pixelation (NUOP) for Penalized Maximum-Likelihood Image Reconstruction for a Single Photon Emission Microscope System. *IEEE Trans. Nucl. Sci.*, 56(5):2777–2788.

Meng, L. and Li, N. (2011). Adaptive Angular Sampling for SPECT Imaging. *IEEE Trans. Nucl. Sci.*, 58(5):2205–2219.

Moore, S., Kouris, K., and Cullum, I. (1992). Collimator Design for Single Photon Emission Tomography. *Eur. J. Nucl. Med.*, 19(2):138–150.

Muehllehner, G. (1973). Septal Penetration in Scintillation Camera Collimators. *Phys. Med. Biol.*, 18(6):855–862.

Natterer, F. (1986). *The Mathematics of Computerized Tomography*. New York: Wiley.

Newiger, H. and Jordan, K. (1985). Optimization of Collimators for Imaging Positron Emitters by a Gamma Camera. *Eur. J. Nucl. Med.*, 11:230–234.

Noo, F., Clackdoyle, R., and Pack, J. (2004). A Two-Step Hilbert Transform Method for 2D Image Reconstruction. *Phys. Med. Biol.*, 49(17):3903–3923.

Nuyts, J. (2009). Small Animal Imaging with Multi-Pinhole SPECT. *Methods*, 48(2):83–91.

Pan, X., Zou, Y., and Xia, D. (2005). Image Reconstruction in Peripheral and Central Regions-Of-Interest and Data Redundancy. *Med. Phys.*, 32(3):673–684.

Patton, J., Slomka, P., Germano, G., and Berman, D. (2007). Recent Technologic Advances in Nuclear Cardiology. *J. Nucl. Cardiol.*, 14(4):501–513.

Pedemonte, S., Bousse, A., Erlandsson, K., Modat, M., Arridge, S., Hutton, B., and Ourselin, S. (2010). GPU Accelerated Rotation-Based Emission Tomography Reconstruction. *Proc. IEEE Nucl. Sci. Symp. Med. Imag. Conf.*, pages 2657–2661.

Pierro, A. R. D. (1995). A modied Expectation Maximization Algorithm for Penalized Likelihood Estimation in Emission Tomography. *IEEE Trans. Med. Imag.*, 14(1):132–137.

Qi, J. (2003). A unified noise analysis for iterative image estimation. *Phys. Med. Biol.*, 48(21):3505–3520.

Qi, J. and Leahy, R. (2000). Resolution and Noise Properties of MAP Reconstruction for Fully 3-D PET. *IEEE Trans. Imag. Proc.*, 1(5):1346–1358.

Qi, J. and Leahy, R. M. (1999). A Theoretical Study of the Contrast Recovery and Variance of MAP Reconstructions from PET Data. *IEEE Trans. Med. Imag.*, 18(4):293–305.

Rockmore, A. and Macovski, A. (1976). A Maximum Likelihood Approach to Emission Image Reconstruction from Projections. *IEEE Trans. Nucl. Sci.*, 23:1428–1432.

Segars, W. (2001). *Development of a New Dynamic NURBS-Based Cardiactorso (NCAT) Phantom*. PhD thesis, The University of North Carolina.

Shepp, L. and Vardi, Y. (1982). Maximum Likelihood Reconstruction for Emission Tomography. *IEEE Trans. Med. Imag.*, 13(4):601–609.

Sidky, E. and Pan, X. (2005). Recovering a Compactly Supported Function from Knowledge of its Hilbert Transform on a Finite Interval. *IEEE Sign. Proc. Lett.*, 12(2):97–100.

Silverman, B. W., Jones, M. C., Wilson, J. D., and Nychka, D. W. (1990). A Smoothed EM Approach to Indirect Estimation Problems, with Particular Reference to Stereology and Emission Tomography. *J. Royal Stat. Soc. Ser.*, 52(2):271–324.

Slijpen, E. and Beekman, F. (1999). Comparison of Post-Filtering and Filtering Between iIterations for SPECT Reconstruction. *IEEE Trans. Nucl. Sci.*, 46(6):2233–2238.

Soares, E., Byrne, C., and Glick, S. (2000). Noise characterization of block-iterative reconstruction algorithms. I. Theory. *IEEE Trans. Med. Imag.*, 19(4):261–270.

Soares, E., Glick, S., and Hoppin, J. (2005). Noise characterization of block-iterative reconstruction algorithms: II. Monte Carlo simulations. *IEEE Trans. Med. Imag.*, 24(1):112–121.

Staelens, S., de Wit, T., and F.Beekman (2007). Fast Hybrid SPECT Simulation Including Efficient Septal Penetration Modelling (SP-PSF). *Phys. Med. Biol.*, 52(11):3027–3043.

Stamos, J., Rogers, W., Clinthorne, N., and Koral, K. (2010). Object-Dependent Performance Comparison of two Iterative Reconstruction Algorithms. *IEEE Trans. Nucl. Sci.*, 35(1):611–614.

Stayman, J. and Fessler, J. (2000). Regularization for Uniform Spatial Resolution Properties in Penalized-Likelihood Image Reconstruction. *IEEE Trans. Med. Imag.*, 19(6):601–616.

Stayman, J. and Fessler, J. (2004a). Compensation for Nonuniform Resolution Using Penalized-Likelihood Reconstruction in Space-Variant Imaging Systems. *IEEE Trans. Med. Imag.*, 23(3):269–285.

Stayman, J. and Fessler, J. (2004b). Efcient Calculation of Resolution and Covariance for Penalized-Likelihood Reconstruction in Fully 3-D SPECT. *IEEE Trans. Med. Imag.*, 23(12):1543–1556.

Takahashi, T. and Watanabe, S. (2001). Recent Progress in CdTe and CdZnTe Detectors. *IEEE Trans. Nucl. Sci.*, 48(4):950–959.

Toossi, M. B., Islamian, J. P., Momennezhad, M., Ljungberg, M., and Naseri, S. (2010). SIMIND Monte Carlo Simulation of a Single Photon Emission CT. *Phys. Med. Biol.*, 35(1):42–47.

Veklerov, E. and Llacer, J. (1987). Stopping Rule for the MLE Algorithm based on Statistical Hypothesis Testing. *IEEE Trans. Med. Imag.*, 6(4):313–319.

Vetterling, W., Teukolsky, S., Vetterling, W., and Flannery, B. (2002). *Numerical Recipes in C - Second Edition*. Cambridge University Press.

Vunckx, K., Beque, D., Defrise, M., and Nuyts, J. (2008a). Single and Multipinhole Collimator Design Evaluation Method for Small Animal SPECT. *IEEE Trans. Med. Imag.*, 27(1):36–46.

Vunckx, K., Suetens, P., and Nuyts, J. (2008b). Effect of Overlapping Projections on Reconstruction Image Quality In Multipinhole SPECT. *IEEE Trans. Med. Imag.*, 27(7):972–983.

Wagenaar, D. (2004). *CdTe and CdZnTe Semiconductor Detectors for Nuclear Medicine Imaging Emission Tomography: The Fundamentals of SPECT and PET*. San Diego, CA: Elsevier.

Wilson, D., Tsui, B., and Barrett, H. (1994). Noise Properties of the EM Algorithm: II. Monte Carlo Simulations. *Phys. Med. Biol.*, 39(5):847–872.

Yendiki, A. and Fessler, J. (2006). Analysis of Observer Performance in Known-Location Tasks for Tomographic Image Reconstruction. *IEEE Trans. Med. Imag.*, 1(25):28–41.

Yendiki, A. and Fessler, J. A. (2004). A Comparison of Rotation- and Blob-Based System Models for 3D SPECT with Depth-Dependent Detector Response. *Phys. Med. Biol.*, 49(11):2157–2168.

Yu, D. and Fessler., J. (2000). Mean and Variance of Singles Photon Counting with Deadtime. *Phys. Med. Biol.*, 45(7):2043–2056.

Zeng, G. and Gullberg, G. (1992). Frequency Domain Implementation of the Three-Dimensional Geometric Point Response Correction in SPECT Imaging. *IEEE Trans. Nucl. Sci.*, 39(5):1444–1454.

Zeng, G. and Gullberg, G. (1997). An SVD Study of Truncated Transmission Data in SPECT. *IEEE Trans. Nucl. Sci.*, 44(1):107–111.

Zeng, G. and Gullberg, G. (2000). Unmatched Projector/Backprojector Pairs in an Iterative Reconstruction Algorithm. *IEEE Trans. Med. Imag.*, 19(5):548–555.

Zeng, G. and Gullberg, G. (2012). Null-Space Function Estimation for the Interior Problem. *Phys. Med. Biol.*, 57:1873–1887.

Zhang, B. and Zeng, G. (2007). Two-Dimensional Iterative Region-Of-Interest (ROI) Reconstruction from Truncated Projection Data. *Med. Phys.*, 34(3):935–944.

Zhang-O'Connor, Y. and Fessler, J. (2007). Fast Predictions of Variance Images for Fan-Beam Transmission Tomography with Quadratic Regularization. *IEEE Trans. Med. Imag.*, 26(3):335–246.

Zheng, G. and Gullberg, G. (2002). A Channelized-Hotelling-Trace Collimator Design Method Based on Reconstruction Rather Than Projects. *IEEE Trans. Nucl. Sci.*, 49(5):2155–2158.

Zhou, L., Defrise, M., Vunckx, K., and Nuyts, J. (2010). Comparison Between Parallel Hole and Rotating Slat Collimation: Analytical Noise Propagation Models. *IEEE Trans. Med. Imag.*, 29(12):2028–2052.

Zhou, L. and Gindi, G. (2008). Collimator Optimization in SPECT Based on an LROC Ideal Observer. *Proc. IEEE Nucl. Sci. Symp. Med. Imag. Conf.*, pages 5540–5547.

Zou, Y., Pan, X., and Sidky, E. (2005). Image Reconstruction in Regions-Of-Interest from Truncated Projections in a Reduced Fan-Beam Scan. *Phys. Med. Biol.*, 50(1):13–27.