QUANTIFICATION AND OPTIMIZATION TECHNIQUES FOR DYNAMIC BRAIN IMAGING IN HIGH RESOLUTION POSITRON EMISSION TOMOGRAPHY

by

JU-CHIEH (KEVIN) CHENG

B.Sc., The University of British Columbia, 2001
M.Sc., The University of British Columbia, 2004

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

(Physics)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)

December 2008

© Ju-Chieh (Kevin) Cheng, 2008
ABSTRACT

Modern positron emission tomography (PET) scanners have a large number of detector crystals which corresponds to an even larger number of lines of response (LOR) addressing the coincidence events. This increasing complexity makes the image reconstruction much more challenging especially in dynamic imaging where the acquired number of counts per frame or per LOR is much less as compared to that in the static case. The low statistical quality of the data thus degrades the quantitative accuracy of the images. Moreover, the increasing number of LORs and dynamic frames requires a longer computational time and larger data storage for the image reconstruction task.

A dual reconstruction scheme, a novel scatter calibration, and a practical scatter and randoms approximation methods were developed in this work. These methods have been validated using phantom, non-human primate, and human studies and have been demonstrated to improve the quantification accuracy of the images, to accelerate the image formation task, and to reduce the data storage requirement for dynamic brain imaging in high resolution PET.

In conclusion, these studies contribute to increasing the accuracy and to decreasing the computational burden for dynamic high resolution quantitative PET imaging. The proposed methods are modular and can be applied to any PET scanners with the exception of the dual reconstruction scheme which requires list-mode acquisition capability. These methods are particularly beneficial for high resolution scanners which have a large number of LORs, such as the high resolution research tomograph (HRRT).
Table of Contents

Abstract ............................................................................................................................. ii
Table of Contents ............................................................................................................... iii
List of Tables .................................................................................................................... vii
List of Figures ................................................................................................................... viii
Glossary .............................................................................................................................. xiii

Acknowledgements ......................................................................................................... xv

Chapter 1 .......................................................................................................................... 1
Introduction to Positron Emission Tomography (PET) ..................................................... 1
  1.1 Basic Principles of PET Imaging ........................................................................... 2
  1.2 Motivation of This Work ......................................................................................... 3
  1.3 Thesis Overview ...................................................................................................... 5
  1.4 Radioactivity and PET Tracers .............................................................................. 7
    1.4.1 Radioactivity and Radiation Dose in PET ...................................................... 7
    1.4.2 PET Tracer Production and Limitations ...................................................... 8
  1.5 Applications of PET .............................................................................................. 9
    1.5.1 Oncology ....................................................................................................... 10
    1.5.2 Cardiology .................................................................................................... 11
    1.5.3 Neurology and Psychiatry .......................................................................... 11
    1.5.4 Pharmacology .............................................................................................. 12
  1.6 Coincidence Detections and Data Acquisitions .................................................... 12
    1.6.1 Types of Detected events ............................................................................. 12
    1.6.2 Energy Window and Data Collection ......................................................... 15
    1.6.3 Sinograms .................................................................................................... 16
  1.7 PET Scanner Design ............................................................................................... 19
    1.7.1 Scintillation Detectors .................................................................................. 19
    1.7.2 Photo-multiplier Tubes (PMT) .................................................................... 21
    1.7.3 Depth of Interaction (DOI) .......................................................................... 24
    1.7.4 The High Resolution Research Tomograph (HRRT) .................................... 25
    1.7.5 Time of Flight (TOF) PET ........................................................................... 27
    1.7.6 Multi-modality Scanners with PET Capability ........................................... 28
  1.8 Quantitative Corrections for the Physical Interactions in PET ............................. 29
    1.8.1 Decay Correction ......................................................................................... 29
    1.8.2 Dead Time Correction .................................................................................. 29
    1.8.3 Detector Normalization Correction .............................................................. 30
    1.8.4 Attenuation Correction ................................................................................ 31
    1.8.5 Randoms Correction .................................................................................... 34
    1.8.6 Motion Correction ....................................................................................... 36
5.1 Chapter Overview ................................................................. 96
5.2 Summary Table for the Proposed Methods in Chapters 5 and 6 .......... 96
5.3 Effect of High RF and/or with Low Number of Counts in the Scatter Tail-Fitting Process .................................................... 97
5.4 Phantom Experiment .......................................................... 99
  5.4.1 Method ................................................................. 99
  5.4.2 Results .................................................................. 100
5.5 A Novel Scatter Calibration Technique ........................................ 106
  5.5.1 A Global-level Scatter Calibration (GSC) ............................... 106
    5.5.1.1 Overall Concept ................................................. 106
    5.5.1.2 Experimental phantom results for GSC .................... 109
  5.5.2 A Segment-level Scatter Calibration (SSC) ............................ 110
    5.5.2.1 Description of Technique ...................................... 110
    5.5.2.2 Experimental phantom results for SSC .................... 111
5.5.3 Applying the Scatter Calibration in the Plane Level ................... 112
  5.5.3.1 Non-human Primate Experiment .................................. 114
    5.5.3.1.1 Methods ....................................................... 114
    5.5.3.1.2 Results ....................................................... 114
  5.5.4 Validation of Constant Axial Scatter Distribution: Early and Later Stages of the Scan ......................................................... 119
    5.5.4.1 Methods .......................................................... 119
    5.5.4.2 Results .......................................................... 122
5.5.5 Experimental Phantom Results for the SPSC Method ................... 126
  5.5.5.1 Methods .......................................................... 126
  5.5.5.2 Results .......................................................... 127
5.5.6 SPSC Method Leading to a Practical Scatter Approximation ........... 129
5.5.7 Applying SSC to Frames with a Very Short Duration ................. 136
5.6 Discussions and Conclusion .................................................. 138

Chapter 6 ................................................................................. 139
A Practical Scatter and Randoms Approximation Technique and Experimental Validations ................................................................. 139
  6.1 Chapter Overview .......................................................... 140
  6.2 Introduction .................................................................. 140
  6.3 A Practical Scatter and Randoms Approximation Technique .......... 141
  6.4 Experiments ................................................................ 144
    6.4.1 Methods .............................................................. 144
    6.4.2 Results .............................................................. 147
  6.5 Combining the Practical Scatter/Randoms Approximation with the Dual Reconstruction Scheme ................................. 153
  6.6 Conclusion ................................................................ 155

Chapter 7 ................................................................................. 156
Conclusion ............................................................................... 156
  7.1 Summary ................................................................ 157
7.2 Concluding Remarks ........................................................................................................ 158

References ............................................................................................................................. 161
List of Tables

TABLE 1.1: Applications of various PET tracers ......................................................... 10
TABLE 1.2: Characteristics of various scintillators ....................................................... 21
TABLE 1.3: Positron range and energy for typical PET radioisotopes ......................... 39
TABLE 4.1: Differences between the 3D-OP and OP-LMEM algorithms; note that the last row shows the pure list-mode approach without any data compression which has not been considered here .................................................. 82
TABLE 4.2: BP value comparison between 3D-OP and OP-LMEM ................................ 88
TABLE 5.1: Summary of the proposed methods in Chapters 5 and 6 ................................. 97
TABLE 5.2: The number of prompts required to produce a relatively accurate global SF for various RFs ........................................................................................................... 102
TABLE 5.3: The standard deviation over the mean of the axial profile ratio comparison ............................................................................................................................. 128
TABLE 5.4: The standard deviation over the mean of the axial profile ratio comparison ............................................................................................................................. 129
TABLE 6.1: Binding potential (BP) comparisons between the conventional and the approximate scatter and randoms correction method ........................................... 152
List of Figures

FIGURE 1.1: Positive beta decay and positron annihilation .............................................3
FIGURE 1.2: The three types of measured coincidence events; prompts = true +
scatters + randoms........................................................................................................13
FIGURE 1.3: The basic concept of the Compton interaction ..............................................14
FIGURE 1.4: Typical range of the scatter angle for the annihilation photons .................15
FIGURE 1.5: The list-mode data acquisition ......................................................................16
FIGURE 1.6: Illustration of the parameters characterizing the address of each LOR in a
sinogram ........................................................................................................................17
FIGURE 1.7: An example of a source distribution (the grey-scale is inversed in this
case) and the corresponding set of sinograms ............................................................18
FIGURE 1.8: (a) An example of the direct planes (segment zero) and planes in two
other oblique segments across the axial direction of the scanner, and (b) the axial
rebinning for span 3: 1-to-1, 2-to-1; for span 9: 5-to-1, 4-to-1. The colour of the
LOR is assigned to match with the colour of the text in order to demonstrate the
rebinning scheme more clearly.....................................................................................19
FIGURE 1.9: (a) A detector block for the HRRT and (b) a block coupled with 4 PMTs
........................................................................................................................................22
FIGURE 1.10: (a) The conventional design for detector blocks coupled with PMTs and
(b) the quadrant sharing design using circular PMTs ....................................................23
FIGURE 1.11: The red arrows show the actual annihilation photon paths, and the black
arrows show the assigned LOR..................................................................................24
FIGURE 1.12: The HRRT scanner located at the UBC hospital, Vancouver, BC,
Canada .......................................................................................................................25
FIGURE 1.13: The cross section ring geometry of the HRRT with the dimensions in
centimetres....................................................................................................................26
FIGURE 1.14: Photon attenuation path along a LOR ......................................................32
FIGURE 1.15: Polaris™ with the trackers shown on the right .........................................37
FIGURE 1.16: LOR shift due to the photon non-collinearity; the ideal LOR is
represented by solid green, and the assigned LORs are represented by dashed red
for different size of detector rings (the angle of deviation is exaggerated in this
case)............................................................................................................................40
FIGURE 2.1: Illustration of the back-project process with (A) being the original source
distribution, (B) with 1, (C) with 3, (D) with 4, (E) with 16, (F) with 32, and (G)
with 64 angular views of back-projection as adapted from [42]..............................45
FIGURE 2.2: An example of the finite voxel parameterization of the image estimate ..47
FIGURE 2.3: The flow chart of the iterative reconstruction algorithms .......................49
FIGURE 2.4: The reconstruction procedures for the histogram-mode reconstruction
(note that the grey-scale of the image is inverted)....................................................51
FIGURE 2.5: An example of the spatial angular subset grouping scheme used in the
histogram-mode OSEM algorithm ...........................................................................52
FIGURE 2.6: The reconstruction procedures for the list-mode reconstruction ............53
FIGURE 2.7: The temporal subset formation scheme in the list-mode reconstruction.
The data of a 'division' are a small part (short duration) of the scan, and the data of
a 'portion' are a small part of the division. The portions with the same colour are
grouped to form the temporal subsets. .......................................................... 55
FIGURE 2.8: Dynamic imaging in PET ............................................................ 57
FIGURE 2.9: An example of the transaxial profile analysis for a phantom image with
and without the scatter correction reconstructed with 3D-OP and OP-LMEM
algorithms .......................................................... 58
FIGURE 2.10: An example of the axial profile analysis ........................................ 59
FIGURE 3.1: A contrast phantom reconstructed with (right) and without (left) the
scatter correction; the single scatter simulation (SSS) was used as the scatter
correction in this case .......................................................... 63
FIGURE 3.2: Ideal projection data from the unscattered true events (left) and realistic
projection data (red) which contains both unscattered and scattered true events
(right); the magnitude of the estimated scatter projection distribution (pink) can be
corrected by fitting the tail of the estimated scatter distribution to the tail of the
measured true close to the boundary of the object ................................................ 65
FIGURE 3.3: The illustration of the SSS calculation for a given LOR ....................... 69
FIGURE 3.4: The scatter scaling factors (y-axis) across the direct planes (x-axis); the
raw scaling factors are depicted in red, and the ones after applying the axial
smoothness constraint (boxcar filter) are in green ........................................... 72
FIGURE 3.5: Prompts and RF in the dynamic frames of typical human raclopride
studies; each color represents a different scan .................................................. 73
FIGURE 3.6: Scatter scaling factors across the direct planes for a frame with 2M
prompts and 94% RF; the raw scaling factors are depicted in red, and the ones after
applying the axial smoothness constraint (boxcar filter) are in green .......... 74
FIGURE 3.7: The temporal SF values with the corresponding motion data with and
without motion compensation ........................................................................ 77
FIGURE 4.1: The transaxial (line) profile across the summed plane (a) without the
scatter correction and (b) with the scatter correction .................................... 83
FIGURE 4.2: (a) The transaxial or radial profile on a summed plane and (b) the axial
profile comparison between 3D-OP and OP-LMEM image reconstruction on a
representative frame ...................................................................................... 84
FIGURE 4.3: The transaxial profile comparison on a single plane for a frame with
~15M true counts .......................................................................................... 85
FIGURE 4.4: (a) Contrast vs number of iteration comparison, (b) ROI noise vs number
of iteration comparison; there is virtually complete overlap between the results
obtained with the two algorithms, (c) Voxel noise vs number of iteration
comparison, and (d) Voxel noise vs number of iteration comparison after applying
a 2mm 3D Gaussian filter ............................................................................ 86
FIGURE 4.5: (a) TAC comparison between 3D-OP and OP-LMEM for the phantom
study and (b) TAC comparison for the non-human primate study; the TACs for the
right caudate (rt-cau) and the left cortex (lf-cortex) regions are shown here .... 87
FIGURE 4.6: The reconstruction time cost v.s the number of counts using the same
number of processors for both 3D-OP and OP-LMEM reconstructions .......... 88
FIGURE 4.7: The flow chart for the dual reconstruction scheme ......................... 90
FIGURE 4.8: The result of the dual reconstruction scheme from the phantom study... 92
FIGURE 4.9: The result of the dual reconstruction scheme from the non-human primate
study ................................................................................................................. 93
FIGURE 5.1: Although the net true counts are 2, there are -2 counts in the true sinogram in this case; the positivity-constraint sets the negative values in the plane to some pre-set positive threshold during the scatter scaling process (overestimation). ................................................................. 98

FIGURE 5.2: The global SF as a function of RF for various numbers of counts using the conventional scatter tail-scaling with and without the positivity-constraint ...... 100

FIGURE 5.3: The SF estimated from each segment (i.e. segment SF) of the sinogram for the frame with 90% RF and 2M prompts with and without the positivity-constraint in the scatter tail-scaling process .......................................... 102

FIGURE 5.4: The segment SF obtained with the p-constraint as a function of the number of true counts per plane within each segment for the frame with 90% RF and 2M prompts ................................................................. 103

FIGURE 5.5: The segment SF for the frame with 75% RF and 2M prompts with and without the positivity-constraint in the scatter tail-scaling process; the reference was obtained from 300M prompts with similar RF ........................................ 104

FIGURE 5.6: The z-r view of the segment zero scatter sinogram which shows all 207 planes (i.e. the vertical axis is along the tomograph axis, the horizontal axis is the radial distance from the tomograph’s centre (see Section 1.6.3), and segment zero refers to the direct planes or sinograms with zero oblique angle) (a) for a frame with 300M counts and (b) for a (biased) frame with 2M counts .......... 105

FIGURE 5.7: The estimated global SF values as a function of time for a typical dynamic brain study; the estimated scatter was obtained without the positivity-constraint ................................................................. 107

FIGURE 5.8: Segment SFs obtained from the frame with 75% RF and 2M prompts using the conventional, GSC methods, and that obtained from the reference frame is also shown ................................................................. 109

FIGURE 5.9: The z-r view of the segment zero scatter sinogram obtained from the frame with 75% RF and 2M prompts using (a) the conventional with p-constraint and (b) the GSC method ................................................................. 110

FIGURE 5.10: The estimated segment SF values for 5 representative segments as a function of time for a typical dynamic brain study ................................................................. 111

FIGURE 5.11: Segment SF obtained from the frame with 75% RF and 2M prompts using the SSC, GSC methods, and that obtained from the reference frame is also shown ................................................................. 112

FIGURE 5.12: The plane SF ratio between the different (temporal) frames ............. 115

FIGURE 5.13: The plane SF ratio between the calibrated and original target frame.... 116

FIGURE 5.14: Segment SF comparison between the original and the calibrated frames ................................................................. 117

FIGURE 5.15: Segment SF comparison between SPSC and GPSC for a frame with 94% RF and 2.6M prompts ................................................................. 118

FIGURE 5.16: The global true rate vs time plot or the global TAC; the frames within the red circle are grouped to form the reference frame; the SSC and SPSC methods are applied to the scatter estimate in the target frame, and the differences in the emission image between the gold-standard and the images reconstructed with the calibrated scatter estimates are compared ........................................ 120

FIGURE 5.17: The global true rate vs time plot or the global TAC; the frames within the red circles are grouped to form the reference frames, and the difference in the
axial distribution is investigated between the gold-standard and the images
reconstructed with the calibrated scatter estimates ........................................ 122

FIGURE 5.18: The first row shows the gold-standard image for the first frame in the
early stage of the scan; the second row shows the ratio of the SSC image to the
gold-standard; the third row shows the ratio of the SPSC image to the gold-
standard; the transaxial, coronal, and sagittal views are shown from left to right
respectively ........................................................................................................ 123

FIGURE 5.19: The percent difference in the axial profile between the images
reconstructed with the SPSC calibrated scatter estimates and the gold-standard
using the reference frames defined in Figure 5.17; the 3300s reference frame
covers the entire later stage of the scan ......................................................... 124

FIGURE 5.20: The first row shows the gold-standard image for the first frame in the
later stage of the scan; the second row shows the ratio of the SPSC image using the
entire later stage of the tracer as the reference to the gold-standard; the transaxial,
coronal, and sagittal views are shown from left to right respectively; the axial
boundaries (plane 80 to 200) as shown in Figure 5.19 are depicted by the two line
profiles in the coronal and sagittal views in the second row ...................... 125

FIGURE 5.21: The z-r view of the segment zero scatter sinogram which shows all 207
planes (a) for the reference with 400M prompts and 40% RF, (b) for the frame with
3M prompts and 40% RF, (c) after calibrating with SSC and (d) SPSC ........... 127

FIGURE 5.22: The voxel noise comparison between the phantom images reconstructed
with the conventional, SSC, and SPSC scatter estimates ............................. 128

FIGURE 5.23: The voxel noise comparison between the phantom images reconstructed
with the SPSC and the approximate scatter estimates .................................. 130

FIGURE 5.24: The z-r view of the segment zero scatter sinogram for the first dynamic
frame of the human study (a) using (prompts – delayed coincidences) with the
positivity-constraint, (b) using (prompts – delayed coincidences) without the
constraint, (c) using (prompts – smoothed randoms) with the positivity-constraint,
(b) using (prompts – smoothed randoms) without the constraint, (e) calibrated with
SSC, and (f) the reference used for the calibration ........................................ 132

FIGURE 5.25: The profiles across the planes defined in Figure 5.24 (a), (b), (c), (d), and
(e) together with the corresponding global scatter fraction (SF) .................... 133

FIGURE 5.26: Comparison of segment SF as a function of the number of true counts
per plane within each segment using the conventional scatter estimate between the
phantom and human study .............................................................................. 134

FIGURE 5.27: The segment SF comparison between all the scatter estimates described
in Figure 5.25 .................................................................................................... 135

FIGURE 5.28: An example which demonstrates the effect of the biased scatter on the
(true) emission image (middle 3), and the same image after applying the calibration
(bottom 3). The image reconstructed without any scatter correction is also shown
(top 3). The transaxial, coronal, and sagittal views are shown from left to right
respectively ..................................................................................................... 136

FIGURE 5.29: A representative TAC comparison between the conventional, the SSC
methods, and without any scatter correction (nosc) ..................................... 137

FIGURE 6.1: A global TAC for a $^{11}$C-dihydrophtenbazine dynamic non-human
primate study. The solid line connects the measured data points and aids in
determining larger changes in the true coincidence rate .............................. 143
FIGURE 6.2: ROIs for the selected hot and cold spots in the striatum plane placed (a) in the emission image and (b) in the scatter image which is obtained by reconstruction the scatter events using the Filtered Back Projection .................................................. 147

FIGURE 6.3: The radial profile comparison in the scatter images of the non-human primate study between the conventional (FBP-S-conventional) and the approximate method with three estimates obtained using the global TAC as guidance (FBP-S-approx) in the plane with the greatest observed difference .... 148

FIGURE 6.4: The scatter TAC (without decay and dead time corrections) comparison for the right caudate region of the non-human primate study between the conventional (FBP-S-conventional) and the approximate method with one (FBP-S-approx1e) and three estimates obtained using the global TAC as guidance (FBP-S-approx) ........................................................................................................ 149

FIGURE 6.5: (a) The ratio of the scatter TAC comparisons for the right-caudate region of the non-human primate study between the conventional and the approximate method with one (FBP-S-approx1e) and three estimates obtained using the global TAC as guidance (FBP-S-approx), (b) The ratio of the scatter TAC comparisons for the selected cold spots as shown in Figure 6.2 which contain low radioactivity but correspond to higher density regions, (c) The ratio of the TAC comparisons for the right-caudate region of the non-human primate study between the conventional and the approximate method with one (OP-approx1e) and three estimates obtained using the global TAC as guidance (OP-approx), and (d) The ratio of the TAC comparisons for the selected cold spots as shown in Figure 6.2 with three estimates ................................................................................................................................. 150

FIGURE 6.6: The ratio of the voxel TAC comparisons for the representative hot and cold voxels with three estimates ................................................................................................................................. 152

FIGURE 6.7: The TAC comparison for a cortex ROI between conventional 3D-OP and the approximated dual reconstruction scheme (left) and the ratio of the TAC between the conventional 3D-OP and the approximated dual reconstruction scheme comparison for 4 representative ROIs (right) ................................................................................................................................. 153

FIGURE 6.8: The BP value comparison (left) and the reconstruction time cost comparison (right; the number of counts ranges from 140M to 13M, and the frames were sorted by number of counts in a descending order) ...................................................................................... 154

FIGURE 7.1: An example of the scatter calibration scheme in the practical scatter approximation ........................................................................................................................................... 159
# Glossary

Abbreviation and Acronyms

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3D-OP</td>
<td>3D OP-OSEM</td>
<td>reconstruction algorithm</td>
</tr>
<tr>
<td>AD</td>
<td>Alzheimer’s Disease</td>
<td>neurodegenerative disease</td>
</tr>
<tr>
<td>BP</td>
<td>Binding Potential</td>
<td>biological parameter</td>
</tr>
<tr>
<td>CT</td>
<td>(X-ray)Computed Tomography</td>
<td>imaging technique</td>
</tr>
<tr>
<td>DEW</td>
<td>Dual Energy Window</td>
<td>scatter correction method</td>
</tr>
<tr>
<td>DOI</td>
<td>Depth of Interaction</td>
<td>imaging term</td>
</tr>
<tr>
<td>EM</td>
<td>Expectation Maximization</td>
<td>reconstruction algorithm</td>
</tr>
<tr>
<td>ETM</td>
<td>Estimation of Trues Method</td>
<td>scatter correction method</td>
</tr>
<tr>
<td>FBP</td>
<td>Filtered Back-Projection</td>
<td>reconstruction algorithm</td>
</tr>
<tr>
<td>FDG</td>
<td>$^{18}$F-Fluoro-Deoxy-Glucose</td>
<td>radio-tracer</td>
</tr>
<tr>
<td>FOV</td>
<td>Field of View</td>
<td>imaging term</td>
</tr>
<tr>
<td>FWHM</td>
<td>Full Width at Half Maximum</td>
<td>analysis term</td>
</tr>
<tr>
<td>GPSC</td>
<td>Global-ratio Plane-level Scatter Calibration</td>
<td>scatter calibration</td>
</tr>
<tr>
<td>GSC</td>
<td>Global-level Scatter Calibration</td>
<td>scatter calibration</td>
</tr>
<tr>
<td>HRRT</td>
<td>High Resolution Research Tomograph</td>
<td>scanner</td>
</tr>
<tr>
<td>LOR</td>
<td>Line of Response</td>
<td>imaging term</td>
</tr>
<tr>
<td>LSO</td>
<td>Lutetium Oxyorthosilicate</td>
<td>scintillator</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
<td>Type</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>MRI</td>
<td>Magnetic Resonance Imaging</td>
<td>imaging technique</td>
</tr>
<tr>
<td>OP-EM</td>
<td>Ordinary Poisson EM</td>
<td>reconstruction algorithm</td>
</tr>
<tr>
<td>OP-LMEM</td>
<td>Subsetized List-Mode OP-EM</td>
<td>reconstruction algorithm</td>
</tr>
<tr>
<td>OSEM</td>
<td>Ordered Subset EM</td>
<td>reconstruction algorithm</td>
</tr>
<tr>
<td>PD</td>
<td>Parkinson’s Disease</td>
<td>neurodegenerative disease</td>
</tr>
<tr>
<td>PET</td>
<td>Positron Emission Tomography</td>
<td>imaging technique</td>
</tr>
<tr>
<td>PIB</td>
<td>Pittsburgh Compound-B</td>
<td>radio-tracer</td>
</tr>
<tr>
<td>PMT</td>
<td>Photo-Multiplier Tube</td>
<td>hardware</td>
</tr>
<tr>
<td>PSF</td>
<td>Point Spread Function</td>
<td>imaging term</td>
</tr>
<tr>
<td>QC</td>
<td>Quality Control</td>
<td>imaging term</td>
</tr>
<tr>
<td>RF</td>
<td>Random Fraction</td>
<td>imaging term</td>
</tr>
<tr>
<td>ROI</td>
<td>Region of Interest</td>
<td>imaging term</td>
</tr>
<tr>
<td>SF</td>
<td>Scatter Fraction</td>
<td>imaging term</td>
</tr>
<tr>
<td>SPSC</td>
<td>Segmented Plane-level Scatter Calibration</td>
<td>scatter calibration</td>
</tr>
<tr>
<td>SSC</td>
<td>Segment-level Scatter Calibration</td>
<td>scatter calibration</td>
</tr>
<tr>
<td>SSS</td>
<td>Single Scatter Simulation</td>
<td>scatter correction method</td>
</tr>
<tr>
<td>TAC</td>
<td>Time Activity Curve</td>
<td>image analysis term</td>
</tr>
<tr>
<td>TOF</td>
<td>Time of Flight</td>
<td>imaging term</td>
</tr>
</tbody>
</table>
Acknowledgements

First of all, I would like to thank my supervisor, Dr. Vesna Sossi, for her continuous support throughout the course of this study. It has been great that she guided me to think outside the box. I would also like to thank all my committee members: Dr. Thomas Ruth, Dr. Andrzej Kotlicki, and Dr. David Measday for their valuable time and advice. Many thanks to my past and present colleagues, Arman Ramim, Marie-Laure Camborde, Stephan Blinder, Siobhan McCormick, Katie Dinelle, Sarah Lidstone, Eric Vandervoort, Geoff Topping, and Joon-Young Suk for their help and being good friends. Furthermore, I would like to thank Carolyn English and Caroline Williams for their assistance during the scan sessions and to thank Ryan Thomson for his help with the computer and network issues. In addition, I would like to thank Dr. Doris Doudet for providing me with the monkey data. Finally, I would like to thank my family members and friends for taking care of me.
Chapter 1

Introduction to Positron Emission Tomography (PET)
1.1 Basic Principles of PET Imaging

Positron Emission Tomography (PET) is a functional imaging modality used in various fields such as oncology, cardiology, pharmacology, psychiatry, and neurology. In particular in brain research, it is used to investigate in-vivo metabolic processes, neurotransmitter, and receptor activity. This ability makes PET a unique tool to investigate disease mechanisms and progression in addition to being extremely useful in drug development.

PET was first developed in the early 1970s soon after x-ray computed tomography (CT) and at about the same time as the magnetic resonance imaging (MRI). In PET, a radioactively labeled compound (radio-pharmaceutical or radio-tracer labeled with positron-emitting radioisotope) is typically injected into a living subject. The tracer then selectively binds to sites of interest and/or undergoes specific metabolic processes. The obtained PET images which show the static 3-D or dynamic 4-D (with the 4th dimension being the temporal or the time domain) distribution of the tracer can then be combined with appropriate mathematical models to investigate and quantify physiologic or metabolic processes, and different tracers can be used to study the function of different targets.

The primary function of PET scanners is to capture the annihilation photons resulting from positron or positive beta decay of the radio-tracer. In positive beta decay, a bound proton decays to a neutron with an emission of positron and an electron neutrino. The positron then travels a short distance (called ‘positron range’ which depends on the energy of the emitted positron), losing its kinetic energy due to Coulomb interactions with the surrounding electrons and eventually annihilates with an electron. The result of the annihilation produces a pair of approximately anti-parallel photons each of which has energy of 511 keV corresponding to the rest mass energy of an electron. The schematic of positive beta decay and positron annihilation is illustrated in Figure 1.1. The pairs of photons are then measured as coincident events with the PET scanner which typically consists of rings of discrete scintillation crystals coupled with photomultiplier tubes (PMT). The measurement of the coincident events originating in the object is termed the ‘emission scan’. Each coincident event has a line of response.
(LOR) associated with the pair of anti-parallel photons or the pair of scintillation crystals which detect the event. A PET image is then obtained by localizing or ‘reconstructing’ the events. The PET image reconstruction results in the spatial distribution of annihilation points which, within the limits of the positron range, corresponds to the radio tracer distribution.

![Diagram of Positive beta decay and positron annihilation](image)

**FIGURE 1.1: Positive beta decay and positron annihilation**

The accuracy of PET imaging depends on the quantitative corrections applied to the data, which account for the physical interactions of the annihilation photons in the subject and also in the detectors. In order to gain accurate information on the tracer distribution as well as the biological processes, it is very important that the images obtained from PET scanners after applying the corrections are quantitatively accurate as the intensity within the regions of interest (ROI) in the image correctly corresponds to the activity concentration within the ROI.

### 1.2 Motivation of This Work

In typical dynamic PET imaging in which the data from the PET scan are divided into a number of ‘dynamic frames’ or individual frames with short durations, the image
reconstruction time as well as data storage is dependent on the number of frames in the study (e.g. frame-dependent or frame-based corrections as will be discussed later). For example of the High Resolution Research Tomograph (HRRT) used in this work, the frame-based correction method (both scatter and randoms; see Chapter 6) takes about 30 minutes per frame using a two dual core Xeon CPU with 3.06 GHz and 2 GB of RAM. Moreover, the size of each scatter or smoothed randoms sinogram data set is ~2 GB per frame. In typical dynamic studies, each scan is divided into 14-20 dynamic frames. Therefore, it takes up to 10 hours just to compute the scatter and randoms estimates for all frames in a single study without even reconstructing any images, and a disk space of 80 GB is required just to store the scatter and smoothed randoms sinograms for a single study when one needs to reconstruct the whole study at once such as in the iterative kinetic parameter estimation [1]. Moreover, the quality and the quantitative accuracy of each frame image can be poor due to the low-statistics or low number of events (counts) acquired within each frame. For example in the conventional scatter correction (which is used to compensate the Compton interaction of the annihilation photons in the object), the biased scatter fraction can be as high as ~270% whereas the unbiased value is ~70% when the number of counts is ~2M for the HRRT as will be described in a later chapter. This bias in the scatter correction then propagates and results in the underestimation of radioactivity concentration.

The primary focus of this work is to accelerate the reconstruction task and also to improve the quantitative accuracy of dynamic brain imaging using the HRRT which is by far the most complex PET scanner to date. This work started with the implementation of a scatter correction based on the single scatter simulation (SSS) technique [2] for the list-mode reconstruction. Then once the same quantitative accuracy between the list-mode and histogram-mode reconstruction had been achieved, a dual reconstruction scheme was devised to accelerate the reconstruction task by making use of the time advantage of the list-mode and histogram-mode reconstructions. With regard to further improvement of quantitative accuracy in dynamic brain imaging, a novel scatter calibration technique was developed to compensate the overestimation bias and to reduce the noise in the axial scatter distribution in the conventional scatter estimation process when the random fraction is high and/or the number of acquired counts is low.
within the dynamic frames. To even further accelerate the reconstruction task, a practical scatter and randoms approximation technique was developed to make the estimation time for the scatter and randoms corrections much less dependent on the number of dynamic frames while maintaining or even potentially improving the quantitative accuracy of the images. The validations of the above techniques were carried out using phantom, non-human primate, and human studies.

1.3 Thesis Overview

The work described in this thesis which consists of seven chapters has resulted in a journal publication [3], a manuscript in preparation [4], and a number of presentations [5, 6, 7]. This dissertation first describes the principles and applications of PET imaging and then discusses the concepts, developments, and testing of the techniques which further accelerate the image reconstruction task and/or further improve the quantitative accuracy of the reconstructed images. The following paragraphs summarize the key contents of each chapter.

Chapter 1: Introduction to Positron Emission Tomography (PET)

This chapter provides the overview of this dissertation and the background materials associated with PET imaging. Basic principles and typical applications of PET imaging are described. The PET scanner instrumentation is then discussed. The chapter finishes with the corrections required to obtain quantitative PET images.

Chapter 2: Image Reconstruction

This chapter describes and compares the reconstruction algorithms with regard to linear versus non-linear, histogram-mode versus list-mode, and static versus dynamic PET imaging. It provides the basic principles of the image reconstruction methods used in this work. It then discusses the challenges in typical dynamic PET imaging and finishes with the image analysis techniques.
Chapter 3: The Most Widely Used Scatter Correction in 3-D PET

This chapter gives a summary of the scatter correction techniques for 3-D PET and talks about the most widely used scatter correction method in greater detail since the major part of this work focuses on further improving the accuracy and efficiency of the scatter correction for dynamic brain studies. The common biases in the conventional scatter estimation are described. Considerations for scatter estimation in dynamic imaging are discussed as well as the applications of scatter fraction.

Chapter 4: Implementation of the Scatter Correction in List-Mode Reconstruction with Validation and a Dual Reconstruction Scheme

This chapter discusses the implementation of the scatter correction previously described in Chapter 3 into the list-mode reconstruction algorithm which had not previously included any scatter correction. The validation was achieved by comparing the images obtained between the list-mode and the conventional histogram-mode algorithm. Figures of merit such as the axial, transaxial image profiles, contrast, noise, time activity curves (TAC), and binding potential (BP) values are compared. Excellent agreements are demonstrated, and a dual reconstruction scheme is developed based on these results. The time gained by applying the dual reconstruction scheme is presented for a representative phantom and non-human primate study.

Chapter 5: A Novel Scatter Calibration Technique and Experimental Validations

In this chapter, a calibration technique which improves the quantification of dynamic brain imaging in high resolution PET is presented. The chapter starts with descriptions on how high random fractions and low number of counts affect the scatter scaling process. The bias in the scatter scaling is then demonstrated using the Single Scatter Simulation (SSS) technique [2, 8] as described in Chapter 3 with a phantom study acquired with the HRRT. A new approach to scatter correction scaling: a novel scatter calibration technique which compensates the bias is discussed together with experimental validations based on phantom, non-human primate, and human studies.
Chapter 6: A Practical Scatter/Randoms Approximation Technique and Experimental Validations

In this chapter, an optimization technique which accelerates the image formation by making the scatter and the smoothed randoms estimation task much less dependent on the number of dynamic frames is described. As previously mentioned, one of the main challenges in dynamic PET image reconstruction is that the reconstruction time depends on the number of dynamic frames (e.g. frame-based scatter and randoms estimations). This chapter presents a practical approximate scatter and smoothed randoms estimation approach for dynamic PET studies. The basic idea of the approximation is to have a time-averaged scatter (randoms) estimate (i.e. scatter and randoms estimates obtained from a summed frame) and scale the averaged estimate according to the trues and randoms counts within each frame to obtain the individual frame scatter and randoms estimates. As a result, the scatter and randoms estimations do not need to be computed for every single temporal frame thus making the reconstruction time less dependent on the number of frames in the dynamic studies. The quantitative accuracy and efficiency (i.e. reconstruction time and data storage gain) of this method is discussed with experimental validations. Finally, the practical approximation is combined with the dual reconstruction scheme to optimize the efficiency of the image formation task.

Chapter 7: Conclusion

This chapter summarizes the key results obtained throughout the thesis and finishes with the final proposed reconstruction protocol and potential future applications.

1.4 Radioactivity and PET Tracers

1.4.1 Radioactivity and Radiation Dose in PET

Radioactivity is characterized by the number of atomic nuclei that undergo decay or 'disintegration' per second. One of the units used for radioactivity is the becquerel, Bq (1 Bq = 1 disintegration per second), which is named after Henri Becquerel. Another commonly used unit for radioactivity is the curie, Ci (1 Ci = 3.7 × 10^{10} disintegrations per second = radioactivity for 1 gram of radium), which is named after Pierre Curie.
Typical units used for PET tracers are the mega-becquerels (MBq) and milli-curies (mCi). A useful conversion is shown in Eq. (1.1):

\[ 37 \text{ MBq} = 1 \text{ mCi} \]  

Typical radioactivity of PET tracers used in oncology and neurology is 200-400 MBq. The injected amount is primarily governed by the allowed subject radiation dose exposure and the tracer principle in which the activity concentration of tracer injected should not perturb the function of the biological system of interest as well as the random fraction and the detector dead time as will be described later.

The energy deposited in the tissue due to ionizing radiation is characterized by the equivalent radiation dose in sievert, Sv (named after Rolf Sievert), which is used to reflect the biological effects of radiation. For acute full body equivalent dose, 1 Sv causes nausea, 2-5 Sv causes hair loss, hemorrhage, and death in many cases, and over 6 Sv survival is unlikely [9]. The commonly used unit for the effective dose in diagnostic medical procedures such as x-ray, CT, and PET is the milli-sievert (mSv). For PET tracers, the effective dose equivalent depends on primarily the amount of injected radioactivity, emitted-positron and photon energy, and the half-life of the radioisotopes. For example, the equivalent dose received from a typical $^{11}$C-raclopride scan is 3-4 mSv (about 0.01 mSv/MBq) [10, 11], whereas the equivalent dose from a typical FDG-PET scan is ~7 mSv since $^{18}$F has a relatively longer half-life than $^{11}$C. This can be compared to the worldwide average annual background radiation of 2.4 mSv [12], 0.02 mSv for a chest x-ray, up to 8 mSv for a CT scan of the chest, and 2-6 mSv per annum for aircrews [13].

1.4.2 PET Tracer Production and Limitations

PET tracers are formed by introducing radionuclides with relatively short half-lives such as $^{11}$C (20.4 min), $^{18}$F (109.8 min), $^{13}$N (~10 min), and $^{15}$O (~2 min) into molecules that bind to receptors or other sites of drug action, or into compounds normally used by the body such as glucose, ammonia, or water. For the radioisotopes with short half-lives, the tracers must be produced using an on-site cyclotron and a radiochemistry laboratory. As a result, limitations to the widespread use of PET arise from the high costs of
cyclotrons, the radiochemistry laboratory with specially adapted chemical synthesis apparatus, and the transportation of tracers (for the ones with relatively longer half-lives). Recently, micro-chips with the size of a coin are being investigated to replace the radiochemistry laboratory [14] as well as self-shielded, low energy cyclotrons as an effort to eliminate the transportation of tracers and to make the production of PET tracers more compact and cost-effective.

1.5 Applications of PET
Unlike standard radiological techniques which primarily outline the anatomical structures, PET is able to image biochemical reactions and physiological functions by measuring concentrations of radio-tracers that are partially metabolized in the body region of interest. Applications of PET include studies of oncological, cardiological, and neurological processes with the main clinical use in oncology. The radio-tracer that has had the most impact on clinical PET imaging is fluorine-18-labeled fluoro-deoxy-glucose (FDG) or C₆H₁₁¹⁸FO₅ which was first introduced in the late 1970s [15].

FDG, an analogue of glucose, is metabolized similarly to glucose in that FDG penetrates cell membranes, called the hematoencephalic barrier (HEB), with the aid of glucose transporter proteins. Afterwards, it initiates glycolysis due to hexokinase phosphorylation in carbon 6. Once phosphorylated, FDG-6-phosphate is metabolically trapped and accumulates within cells, whereas this is not the case for glucose-6-phosphate. This process is more intense in cells with a higher uptake of glucose, such as tumour cells. Furthermore, the phosphorylation is irreversible, in that FDG-6-phosphate can not go back to FDG. This fact allows studies of any organ or tissue with an increase in glucose metabolism, whether it is normal or pathological, given that cancerous cells consume more glucose than normal cells. Note that the primary exception to the metabolic trapping is in the liver, where the dephosphorylation of the FDG-6-phosphate and clearance of FDG from the liver occurs due to the large concentration of phosphatase enzymes [16]. Some of the specific applications of PET in each field with various tracers are listed in Table 1.1 and will be described next.
<table>
<thead>
<tr>
<th>Radioisotope</th>
<th>Tracer</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>$^{82}\text{Rb}$</td>
<td>$\text{RbCl}$</td>
<td>Blood flow</td>
</tr>
<tr>
<td>$^{13}\text{N}$</td>
<td>Ammonia: $\text{NH}_3$</td>
<td>Blood flow</td>
</tr>
<tr>
<td>$^{15}\text{O}$</td>
<td>Water: $\text{H}_2\text{O}$</td>
<td>Blood flow</td>
</tr>
<tr>
<td>$^{18}\text{F}$</td>
<td>FDG</td>
<td>Glucose metabolism</td>
</tr>
<tr>
<td>$^{18}\text{F}$</td>
<td>F-DOPA</td>
<td>Dopamine synthesis and storage</td>
</tr>
<tr>
<td>$^{11}\text{C}$</td>
<td>Raclopride</td>
<td>Dopamine D2 type receptor</td>
</tr>
<tr>
<td>$^{11}\text{C}$</td>
<td>DTBZ</td>
<td>Vesicular monoamine transporter type 2</td>
</tr>
<tr>
<td>$^{11}\text{C}$</td>
<td>PIB</td>
<td>Amyloid plaque</td>
</tr>
</tbody>
</table>

*TABLE 1.1: Applications of various PET tracers*

### 1.5.1 Oncology

The clinical applications of PET in oncology have been developed since 1980s. PET has been applied to investigations such as lung carcinoma, necrosis and recurrence in brain tumours, melanoma, head and neck carcinomas, and breast carcinoma. The advantage of using PET in oncology is the detection of the activity of very small masses of cancerous cells, which give an adequate indication of tumour activity with the use of FDG. The general indications of a PET scan in oncology include: initial diagnosis of cancer, differentiation between benign and malignant tumours, determination of the degree of tumour malignancy, staging the extent of the disease, confirmation of lesion significance detected by CT, MRI, and x-ray studies, treatment response follow-up monitoring, and detection of possible disease recurrence. The clinical applications of PET in oncology have been recently reviewed in [16].
1.5.2 Cardiology

PET has served an important role in cardiovascular research for nearly half a century. Radio-tracers such as rubidium-82-labeled rubidium chloride ($^{82}$RbCl) and nitrogen-13-labeled ammonia ($^{13}$NH$_3$) have been used to investigate blood flow and myocardial perfusion while FDG has been used to study glucose metabolism (i.e. viability of tissues) in the myocardium. Combined tracer technique applying polar mapping has been used to confirm the viability of an ischemic myocardium, in that perfusion defects can be identified with $^{13}$NH$_3$ and if there are still enough viable tissues remaining (obtained from the FDG image) in the region of the myocardium, a ‘by-pass’ surgery or revascularization would be chosen as treatment instead of a heart transplant since the tissues are likely to recover after the restoration of blood circulation.

1.5.3 Neurology and Psychiatry

One of the earliest applications of PET was in neurology for assessing the regional distributions of blood flow, neuro-transmitters, receptors, enzymes, and glucose metabolism. FDG-PET has been used to differentiate Alzheimer’s disease (AD), a progressive neurodegenerative disorder associated with gradual decline in cognition and behaviour, from other dementing processes and also to make early diagnosis of AD (in that AD greatly decreases brain metabolism of glucose; more specifically, a PET image shows relatively symmetric marked hypometabolism of the frontal, parietal and temporal lobes with sparing of the sensory motor strips, occipital lobes, cerebellum, and deep gray matter). Furthermore, Pittsburgh Compound-B (PIB), a novel probe developed at the University of Pittsburgh, has been used to visualize the amyloid plaques in AD thus contributing to aid in the development of novel anti-amyloid therapies.

PET has also been applied to investigate Parkinson’s disease (PD), a progressive neurodegenerative disorder characterized by the selective death of the dopaminergic neurons in the substantia nigra or nigrostriatal pathway. In particular, $^{18}$F-fluoro-DOPA has been used to image the loss of dopaminergic nerve terminals in the basal ganglia (putamen and caudate) of PD patients. Other radio-ligands such as $^{11}$C-raclopride and $^{11}$C-DTBZ have been developed for specific neuroreceptor and transporter (Dopamine
D2 receptor and vesicular monoamine transporter, respectively) to visualize and to correlate regionally reduced binding and the major motor features in PD. Binding Potential (BP), a quantity which is proportional to the regional density of the receptors/transporters, is usually calculated for these studies. In addition, radio-ligands which bind to neuroreceptors have also been used to examine the state of the receptors in patients compared to healthy controls in schizophrenia, substance abuse, mood disorders, and other psychiatric conditions.

1.5.4 Pharmacology

Other than the clinical human applications, PET has also been used in pre-clinical trials by radiolabeling new drugs and injecting them into animals such as rats and non-human primates. The uptake of the drugs, the tissues in which they concentrate, and their eventual elimination (washout), can be monitored far more quickly and cost effectively than the older technique of killing and dissecting the animals to obtain the same information. As a result, the distributions and the effects of the drugs can be studied in-vivo using PET.

1.6 Coincidence Detections and Data Acquisitions

1.6.1 Types of Detected events

As described in Section 1.1 and illustrated in Figure 1.1, the data acquired using a PET scanner are the coincidence events or pairs of approximately collinear 511 keV photons generated from the positron annihilations. A pair of photons is registered as a coincidence event when the two photons are detected nearly simultaneously (typically within a few nano-seconds depending on the capability of the detector). In other words, after a photon hits and gets recorded by a detector, the other detectors prompt for a second photon. If a second photon is detected within a coincidence time window, typically a few nano-seconds, one coincidence event also called a 'prompt' is registered and a corresponding line-of-response (LOR) is assigned. Otherwise, a single photon detection or a 'single' is registered if no second photon is detected.
The measured prompts consist of (i) the unscattered true coincidences or simply the true events, (ii) the scatter events, and (iii) the accidental coincidences or the random events. The three types of coincidences are illustrated in Figure 1.2.

\[ h\nu' = \frac{h\nu}{1 + (h\nu/m_0c^2)(1 - \cos \theta)} \]

where \( h \) is the Planck’s constant, \( \nu' \) is the frequency of the scattered photon, \( \nu \) is the frequency of the incident photon, \( h\nu' \) represents the energy of the scattered photon, \( h\nu \) represents the energy of the incident photon (which is equal to 511 keV for the
annihilation photons), $m_0$ is the mass of the electron, $c$ is the speed of light ($m_0c^2$ is the rest mass energy of the electron which is equal to 511 keV), and $\theta$ is the scattering angle.

A useful graph can be extracted from Eq. (1.2) by plotting the energy fraction ($h\nu'/h\nu$) as a function of the scattering angle for the 511 keV photons as shown in Figure 1.4. Given that the energy resolution of PET scanners such as the HRRT (see Section 1.7.4) is about 30%, the scattering angle ranging from $0^\circ$ to $\sim 55^\circ$ for the annihilation photons can still be detected as the 511 keV photons; this also depends on the energy window assigned for the coincidence measurement as will be elaborated shortly.

*FIGURE 1.3: The basic concept of the Compton Interaction*
The last type of the coincidence is the random events. As shown in Figure 1.2, two photons from different annihilation sites are registered as a coincident event when one of each pair of the annihilation photons (i) gets absorbed in the object, (ii) escapes the scanner’s field of view (FOV), or (iii) passes through the detector without getting recorded. Consequently, the coincident event is measured by ‘accident’ due to the finite coincidence timing window, and the assigned LOR most likely does not correspond to either of the annihilation sites. The detection of scatter and random events results in image blurring or degradation of PET image resolution, and quantitative corrections need to be applied to compensate for these physical interactions as will be discussed later. In addition, the tracers are typically injected from outside the FOV of the scanner, and the tracer activity outside the FOV during the early stage of the scan contributes to higher detections of scatter and random events.

1.6.2 Energy Window and Data Collection

The measurement of the coincident events is also governed by the energy window threshold assigned during the data acquisition. For example with the HRRT, the lower energy threshold is 400 keV and the upper energy threshold is 650 keV for emission
scans. Therefore, all photons detected with energy within the energy window are kept, and the energy window assignment is basically a compromise between increasing the detection sensitivity of the scanner and decreasing the detection of the scattered events.

For modern PET scanners, the coincident events or counts are stored in so called a ‘list-mode’ file in that the information such as the detection time, energy, and the address of the LOR of each event is stored one-by-one into a list as illustrated in Figure 1.5. Typically after the scan, the data are histogrammed or rebinned into sinograms (as will be elaborated next) for the purpose of image reconstruction.

FIGURE 1.5: The list-mode data acquisition

1.6.3 Sinograms

Sinograms are 2-dimensional histograms which contain the total number of counts or the projection data for each LOR. The address of each LOR for each tomograph plane is characterized by two parameters: the radial position, r, and the azimuthal angle, θ as depicted in Figure 1.6.
2D coincidence detector system

1D parallel projection $p(r, \theta)$

Projected data for each LOR stored in a 1D array

*FIGURE 1.6:* Illustration of the parameters characterizing the address of each LOR in a sinogram
Similar to a matrix, each sinogram element represents the total number of counts detected along a LOR using the two parameters as indices. Sinograms are typically shown with the azimuthal angle $\theta$ as the $y$-axis and the radial position $r$ as the $x$-axis as depicted in Figure 1.7. Each activity source point traces out a sinusoidal curve in the $\theta$-$r$ plane (the farther away the source point from the center of the detector ring, the higher the amplitude the sinusoidal curve) hence the data slice is termed ‘sinogram’. The axial position of the sinograms is characterized by the parameter, $z$, also referred to as the ‘plane’ number as shown in Figure 1.7. In 3-D PET, LORs with oblique axial angles which form the oblique sinograms are also acquired. The oblique angle is also referred to as ‘segment’ with segment zero sinograms corresponding to the direct (i.e. planes perpendicular to the axial direction of the scanner) sinograms. A segment number is assigned to each oblique angle, and sinograms within the same segment have the same oblique angle. The maximum (detector) ring difference determines the maximum oblique angle. In addition, the number of sinogram planes decreases with higher oblique angles as illustrated in Figure 1.8a. As a result, higher segments contribute less to the image reconstruction.
In order to reduce the large size of data sets from the 3-D acquisition, axial rebinning or data-mashing is typically performed prior to histogramming to form sinograms. Axial rebinning refers to summing the counts between nearby axial LORs which results in fewer LORs with more counts thus reducing the total number of sinograms. The level of axial-rebinning is described by the ‘span’. A span of 3 or span 3 refers to the rebinning scheme which mashes 2 LORs into 1 and 1 into 1 (i.e. 3 into 2), and a span of 9 or span 9 means rebinning 4 LORs into 1 and 5 into 1 (i.e. 9 into 2) as illustrated in Figure 1.8b. Although higher spans reduce more data size, they introduce more degradation in image resolution. A span of 3 has been used throughout the course of this thesis.

1.7 PET Scanner Design

1.7.1 Scintillation Detectors

PET scanners typically consist of rings of inorganic scintillators or scintillation crystals coupled with photo-multiplier tubes (PMT). The scintillation crystals convert the absorbed annihilation photons which excite the scintillation material into scintillation photons by de-excitation which in turn pass into the PMTs in order to amplify and form electronic signal outputs. Therefore, in order to effectively measure the annihilation
photons, ideal PET scintillation crystals should have (i) great ability to stop/record the annihilation photons within the crystals, (ii) a short decay time so that they can recover to count more annihilation photons, and (iii) a high light yield to obtain a well-defined signal.

The ability to stop/record the annihilation photons is governed by the effective atomic number, the mass density, and the linear attenuation coefficient which is the fraction of photons that interact per unit length or thickness of attenuating material. Higher atomic number and density (i.e. electron density) materials are desirable since the photons are more likely to interact with the electrons in the crystals thus resulting in higher probability of event or count detection. One of the dimensions (pointing towards the transaxial center of the detector rings) of the crystals is also made longer to ensure the interaction of the photons. A short decay or dead time allows the scintillation crystals to perform under high count rate situations as well as reducing the random coincidences as will be discussed later in the chapter. The light yield or the conversion efficiency is characterized by the number of light (scintillation) photons produced per energy deposited by the annihilation photons. A high light yield allows better energy resolution (i.e. the ability to resolve photons with different energies), thus contributing to better discrimination against the Compton scattered events. Other characteristics such as the index of refraction and the wavelength of the scintillation photons are also considered when coupling with PMTs. The characteristics of some typical scintillation materials are listed in Table 1.2 [17, 18, 19].
<table>
<thead>
<tr>
<th>Scintillator</th>
<th>Effective atomic number</th>
<th>Density (g/cm$^3$)</th>
<th>Attenuation coefficient at 511 keV (cm$^{-1}$)</th>
<th>Decay time (ns)</th>
<th>Light yield (photons/MeV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NaI</td>
<td>50</td>
<td>3.7</td>
<td>0.34</td>
<td>230</td>
<td>38000</td>
</tr>
<tr>
<td>BGO</td>
<td>74</td>
<td>7.1</td>
<td>0.96</td>
<td>300</td>
<td>9000</td>
</tr>
<tr>
<td>GSO</td>
<td>59</td>
<td>6.7</td>
<td>0.71</td>
<td>60</td>
<td>8000</td>
</tr>
<tr>
<td>LSO</td>
<td>66</td>
<td>7.4</td>
<td>0.88</td>
<td>40</td>
<td>26000</td>
</tr>
<tr>
<td>LaBr$_3$</td>
<td>46</td>
<td>5.3</td>
<td>0.47</td>
<td>15</td>
<td>60000</td>
</tr>
</tbody>
</table>

**TABLE 1.2: Characteristics of various scintillators**

PET scanners with bismuth germinate (BGO) detectors have been quite common in the past. Recently, lutetium oxyorthosilicate (LSO) has become more popular due to its fast decay time and high light yield though slightly radioactive due to the $^{176}$Lu isotope [20] (~100 true coincidences per second for the HRRT as will be described in Section 1.7.4), and LaBr$_3$ has been used for the 'time-of-flight' (TOF) PET as will be described later. Other than inorganic scintillators, PET scanners incorporating liquid xenon detectors which were claimed to have much superior energy resolution as compared to scintillators are currently under development [21].

### 1.7.2 Photo-multiplier Tubes (PMT)

The PMTs used to amplify the signals from the scintillation photons typically consist of a photocathode, a series of dynodes, and an anode. The amplification process starts with the scintillation photons passing through the photocathode and generating the primary photoelectrons. The number of primary photoelectrons generated per incident scintillation photons is termed 'quantum efficiency'. The primary photoelectrons then go through a series of dynodes which function as electron multipliers and form secondary and post-secondary electron showers. The amplified signals then reach the anode and form the readouts used in the energy discrimination window.
In the past, each crystal element was coupled with a single PMT. As a result, the size of the crystals which affect the image resolution depended primarily on the size of the PMTs. Moreover, the number of PMTs used was thus directly proportional to the number of scintillation crystals, which made half of the cost of the PET scanner and was not economical when typical scanners consist of more than thousands of crystals. Consequently, modern PET scanners typically incorporate the ‘crystal block’ design which was first introduced in the 1980’s [22]. Figure 1.9a shows a block of crystals (8 x 8) for the HRRT. The block design allows not only the use of smaller crystals which improve the image resolution but also causes a significant reduction on the number of PMTs.

*FIGURE 1.9: (a) A detector block for the HRRT and (b) a block coupled with 4 PMTs*

The gaps between the crystals are filled with white diffusive reflecting material [23] in order to trap the isotropic emission of the scintillation photons within each crystal. Conventionally, each block was coupled with 4 PMTs as depicted in Figures 1.9b and 1.10a (i.e. 4 detector blocks represented by the dashed-red squares are coupled with 16 PMTs which are represented by the blue-purplish squares). Recently, as an effort to further reduce the number and the cost of the PMTs, the ‘quadrant-sharing’ design
applying circular PMTs has been developed [24] as illustrated in Figure 1.10b. In this design, the same number of PMTs can be coupled with more detector blocks (i.e. 16 PMTs can be shared between 9 blocks instead of 4 as shown in Figure 1.10) with circular PMTs being cheaper and available with more sizes although the optimal shape of the detector rings is limited to 6-8 heads (i.e. panels of blocks) polygonal rings. For each detected annihilation photon within a block, the signals obtained from the 4 PMTs (identified as a, b, c, and d) are therefore inversely proportional to the distance between the PMTs and the detector element. The address (x, y) of the crystal can thus be positioned by using a simple analog ratio of the signals from the 4 PMTs as shown in Eq. (1.3):

\[
x = \frac{s_a + s_b - (s_c + s_d)}{s_a + s_b + s_c + s_d}, \quad y = \frac{s_a + s_c - (s_b + s_d)}{s_a + s_b + s_c + s_d}
\]

(1.3)

where \(s_a, s_b, s_c, \) and \(s_d\) are the signals from PMT a, b, c, and d respectively.

![Diagram](image)

**FIGURE 1.10:** (a) The conventional design for detector blocks coupled with PMTs and (b) the quadrant sharing design using circular PMTs.
1.7.3 Depth of Interaction (DOI)

**FIGURE 1.11:** The red arrows show the actual annihilation photon paths, and the black arrows show the assigned LOR

For the annihilation sites away from the axial or transaxial center of the detector rings, the pairs of photons intersect the detector crystals at higher oblique angles with respect to the longest dimension of the crystal. Due to the finite width of the crystal which is particularly the case for the small crystals in the block detector design, the incident annihilation photons with high oblique angles are likely to penetrate the first crystal of incident and getting detected by the adjacent or nearby crystals, and this is also more likely to happen for non-circular detector rings. This effect degrades the image resolution since the assigned LOR does not match the actual path of the annihilation photons as illustrated in Figure 1.11 (left) for the transaxial case. This is generally called the 'parallax effect', and the characteristic depth at which the annihilation photons are most likely stopped/absorbed in the crystal is referred to as the depth-of-interaction (DOI). As shown in Figure 1.11 (right), dual layer crystal design enables the use of the DOI information and improves the image resolution at off center regions as the assigned LOR matches more closely to the actual annihilation photon path. For the dual or multiple layer crystal design, the crystal material for each layer needs to have similar index of refraction and the gaps in between the layers are filled with index matching fluid in order to minimize the internal reflection of the scintillation photons and maximize the transmission towards the PMTs. Modern PET scanners such as the HRRT incorporate the dual layer LSO/LYSO (LYSO is LSO with an yttrium impurity)
detector design which also compensates for the extra parallax effect due to the non-circular octagonal detector ring geometry optimized for the block detector design.

1.7.4 The High Resolution Research Tomograph (HRRT)

The PET scanner used in this work is the second generation of Siemens high resolution research tomograph (HRRT) [25, 26] which is a dedicated 3D human brain scanner as shown in Figure 1.12. It has a octagonal detector ring design with each detector head consisting of a block panel of a dual 10mm layer of LSO/LYSO (each with different decay time so that the DOI information can be encoded). Each panel consists of $9 \times 13$ blocks coupled with $10 \times 14$ PMTs using the quadrant-sharing design as described previously. Each block consists of $8 \times 8 \times 2$ crystals with the size of the crystal being $2.4 \times 2.4 \times 10$ mm$^3$, and the total number of crystals is thus 119808. Each head is in coincidence with 5 opposing heads resulting in ~4.5 billion possible LORs (about 2-3
order of magnitude higher than most modern PET scanners) with no axial (rebinning) compression or ~460 million LORs in span 3 and a maximum ring difference of 67 with a sinogram data set size of 1.8 GB per frame for the scatter and random events (floats) and 0.9 GB per frame for the prompts (integers). The radial field-of-view (FOV) of the scanner is ~32 cm in diameter, and the axial FOV is ~25 cm. As illustrated in Figure 1.13, each panel head is separated by a 1.7 cm gap which results in ~10% of data loss and makes the image reconstruction task more challenging.

**FIGURE 1.13:** The cross section ring geometry of the HRRT with the dimensions in centimetres
The intrinsic background activity is ~100 true coincidences per second for the HRRT with an upper energy window of 650 keV and a lower energy window of 400 keV due to a cascade of gamma rays following the beta decay of $^{176}$Lu into $^{176}$Hf in the LSO crystals [20]. As a result, most protocols require the scan to stop before an activity threshold in order to minimize the contribution from the intrinsic LSO background activity. The increase in the scatter fraction (SF; see Section 3.6) due to the activity from outside the FOV has been reported to be ~5% [25] which agrees with our own findings. The detection sensitivity, which is the ratio of the measured true coincidence count rate to the activity of the source within the sensitive volume, is ~6% for the HRRT; this can be compared to the ~1.2% sensitivity for the ECAT 953B (a state of the art human brain scanner in the 1990s) and the ~2% sensitivity for small animal PET scanners such as the microPET R4. The image resolution, which is characterized by the full width at half maximum (FWHM) of the reconstructed point source image, is ~2.4 mm at the center of the FOV for the HRRT as compared to the ~6 mm resolution for the ECAT 953B. The data acquired are stored in the list-mode file which is then used for the histogram/list-mode reconstructions as will be elaborated in Chapter 2. The energy resolution is 25-30%, and the coincidence timing window can be set as low as 2 ns (though a 6 ns timing window is typically used) for the HRRT.

1.7.5 Time of Flight (TOF) PET

The idea of using the time of flight (TOF) information in PET reconstruction was first proposed in the 1960s. The concept of TOF PET is simply that by measuring the arrival time of each annihilation photon for each annihilation site, one can obtain the difference in distance travelled between the two photons along the LOR from the difference in the arrival time, and therefore, the annihilation event can be localized from the TOF information. As a result, one can pin-point the annihilation site exactly if the TOF information is known perfectly, and there is no need to reconstruct the images using any algorithms. However, this is not yet possible due to the limited timing resolution of the detectors and electronics.

Even though the timing resolution of the detectors and electronics is not good enough to fully localize the event thus significantly improving the image resolution, the TOF
information can be used to reduce the noise in the images when incorporated into the reconstruction task [19]. This can be understood by comparing the backprojection of the conventional PET and TOF PET as the conventional PET backprojects the events with equal probability along a LOR (thus introducing the ‘background noise’ along the whole LOR), where as the TOF PET backprojects the events within a portion of a LOR since the TOF provides a better localization of the range of the annihilation sites and consequently reduces the noise along the LOR (see Chapter 2 for more detail in image reconstruction). In addition, the bigger the object, the better the noise improvement can be achieved. The noise reduction can then be translated into a shorter scan time since the same image quality can be achieved with a lower number of counts as has been demonstrated by Philips GEMINI TF [19]. As the technology advances, faster detectors and electronics are developed and the TOF information can provide better and better localizations (smaller volume) of the events along the LOR thus reaching the goal of improving the image resolution (note that the best scintillators and electronics today can achieve ~300 ps timing resolution though a 3 ps resolution is needed to localize a event within 1 mm).

1.7.6 Multi-modality Scanners with PET Capability

One feature which PET images lack is the precise localization of the anatomical structures. Often anatomical information aids in the interpretation of the results, and the subjects need to undergo a separate CT or MRI scan in order to obtain the anatomical information. The functional images from PET are then co-registered or re-aligned with the anatomical images obtained from CT or MRI. Therefore, it would be advantageous to obtain both the anatomical and functional information at once to avoid the delay and transfer of the object in between the scans. PET/CT scanners have been developed and distributed by companies such as GE, Phillips, and Siemens. Promising results have been demonstrated [27], and PET/CT scanners have been used routinely. However, the PET/CT scanners do not truly perform the PET and CT scanning simultaneously, and the inaccuracy in the co-registration still exists due to object movement and intrinsic body motion such as the lung and heart motion. Moreover, the dose introduced from the CT scan is relatively high compared to PET and MRI (which has none). As a result, MR/PET scanners have been recently proposed and developed as an effort to minimize
the dose received by the objects and to perform simultaneous anatomical and functional scanning (this also reduces the total scan time). Note that the design of conventional PET scanners needs to be modified in order to combine with MRI since the PMTs do not function properly inside a magnetic field (as the field introduces drift motions on the photoelectrons).

1.8 Quantitative Corrections for the Physical Interactions in PET

In order to obtain quantitative PET images, a series of corrections needs to be applied in the reconstruction task to take into account the physical interactions. The corrections must be applied for radioactive decay, detector dead time, detector normalization, object attenuation, random coincidences, Compton scattered events, and subject motion during the scan. Additionally, resolution recovery with point spread function (PSF) has been used to model the image blurring caused by the physical limitations of PET (i.e. positron range and photon non-collinearity) as will be described shortly. The reconstructed images after incorporating all the corrections are then calibrated to obtain the absolute radioactivity concentration.

1.8.1 Decay Correction

As the radioisotope decays as time progresses, the rate of annihilation decreases thus reducing the number of the counted events. Consequently, the radioactive decay must be accounted for in order to make the images only sensitive to biological processes in dynamic imaging (as will be elaborated in Chapter 2). Given the well-known half-lives of the radioisotopes, the decay can be corrected by simply scaling the reconstructed images (counts) according to the global decay factors derived from the half-lives.

1.8.2 Dead Time Correction

The non-instant processing time of an event due to the scintillator decay and electronics limits the count rate of the measured coincidences. As one of the consequences, the additional annihilation photons mostly ‘pile up’ (which is referred to as the pulse pile-up effect) in the detectors and the signals are summed and appear as fewer events with higher energy (some of them can also go outside of the energy window and get rejected) during the data processing at high count rate. The system is ‘dead’ when it is busy
processing the data and writing on the disk. The count loss due to this is referred to as the system dead time effect, and it is one of the factors which limit the amount of radio-tracers injected. In order to recover the missing counts thus obtaining quantitative PET images, one needs to correct for the dead time effect. For the block detector PET scanners, the dead time correction is typically applied post-reconstruction as a global factor estimated from the measured true rate and singles rate. In dynamic PET imaging, the dead time correction factor is high for the early frames (high radioactivity), and it typically approaches 1 (i.e. zero dead time) for the later frames due to radioactive decay and tracer washout.

1.8.3 Detector Normalization Correction

The compensation for the variations in detector pair sensitivity/efficiency is referred to as the normalization correction in PET imaging. The detector pair sensitivity can be different due to variations such as in incident angle of the annihilation photons to the detectors, crystal imperfections, PMT gains, and other electronics. Therefore, a normalization correction factor needs to be assigned to each LOR (similar to a look-up table or matrix for the detector pairs). Typically, a normalization scan is performed with a rotating rod source, and the normalization correction factor for each LOR is obtained by dividing the average counts per LOR by the measured counts for each LOR. This process is referred to as the direct normalization. A normalization sinogram data set is obtained at the end of the process, and it is used before or during the image reconstruction task.

However, the normalization scan is very time consuming (60-120 hours) since the rotating source needs to be weak in order to minimize the dead time effect, and it also needs to cover all the detector pairs or LORs with sufficient number of counts (with modern PET scanners consisting of hundreds of millions of LORs). Furthermore, for the particular example of the HRRT frequent normalization scans need to be performed due to the drift in PMT gains over time. Additionally, multiple direct normalization scans need to be performed with different activity sources in order to compensate for the different pulse pile-up behaviours due to different count rates encountered in the dynamic studies. As a result, an alternative component-based normalization correction
method has been proposed and developed [28]. This approach is much more efficient since components such as the geometrical factors and detector efficiencies are measured and/or calculated separately for each detector (note that the number of detectors is much less than the number of LORs), and the normalization factor for each LOR is obtained from the combinations of separate detector components.

1.8.4 Attenuation Correction

The primary photon interactions in matter are the photo-electric absorption, the Compton scattering, and the pair production. Which one of them is likely to occur depends on the incident photon energy. For the case of the 511 keV annihilation photons, the dominant interaction in tissue is the Compton scattering as briefly described in Section 1.6. The results from the Compton scattering are (i) one or both of the annihilation photons get lost by either going outside the FOV or getting absorbed in the object when their energies become low enough for the photo-electric absorption after scattering and (ii) the annihilation photons are still detected but at a wrong LOR as shown in Figure 1.2. In both cases, the number of coincidence events along the original LOR is reduced, and this reduction in the measured counts is referred to as the attenuation. Attenuation is energy-dependent due to the energy dependence of the photon interactions. Although both attenuation and scattering in the coincidence measurements are mainly due to the Compton interaction, the corrections for them are different in that the scatter correction (as will be described shortly after and in Chapter 3) removes the scatter events, whereas the attenuation correction corrects for the missing counts due to the attenuation or Compton interaction in this case.

The probability of survival for a photon in the attenuating medium through a linear path is given by:

\[
P_{s01} = e^{-\int_{s0}^{s1} \mu(x)dx}
\]

where \(\mu(x)\) is the position or structure dependent linear attenuation coefficient (which increases with higher density or lower photon energy in the Compton scattering and
photo-electric absorption range), $x_0$ is the point of the annihilation, and $x_1$ is the edge point of the object along the linear direction $x$ as depicted in Figure 1.14.

![Figure 1.14: Photon attenuation path along a LOR](image)

The probability of survival for the partner annihilation photon is given by:

$$P_{s02} = e^{-\int_0^{x_2} \mu(x) dx} \quad (1.5)$$

where $x_2$ is the opposite edge point of the object along the linear direction $x$ (note that the probability stays the same if the photon travels in the opposite direction; i.e. from outside to the inside, so the limits of the integral is interchangeable).
The survival probability of both annihilation photons along the LOR is thus given by:

\[
P_{LOR} = P_{12} = e^{-\int_{0}^{2} \mu(x) dx - \int_{0}^{x} \mu(x) dx} = e^{-\int_{x}^{2} \mu(x) dx}
\]  

(1.6)

Therefore, the probability is independent of the point of the annihilation and only depends on the attenuation along the linear path.

The attenuation correction involves performing a transmission scan of the object using an external moving source and a blank scan. For example, a $^{137}$Cs source (a single photon emitter with energy of 662 keV) is used for the HRRT, and a blank scan in which the external source is used with the FOV being empty is subsequently performed. As the source orbits around the object during the transmission scan, single-photons transmit and get attenuated through the object. The photons are then recorded by the detectors as they pass through the object. The same measurement is repeated without the object during the blank scan. The attenuation correction factors (ACFs) can then be obtained by:

\[
ACF_i = \frac{n_{i}^{\text{blank}}}{n_{i}^{\text{trans}}}
\]

(1.7)

where $n_{i}^{\text{blank}}$ is the measured counts for the $i^{th}$ LOR in the blank scan, and $n_{i}^{\text{trans}}$ is the measured counts for the $i^{th}$ LOR in the transmission scan.

Although the ACFs can be obtained from this direct measurement, a long scan is required in order to have sufficient counts in each LOR to minimize the noise. Furthermore, the ACFs need to be scaled since the transmission source has a different energy from the 511 keV photons. As a result, the (improved) attenuation correction is typically done by reconstructing the attenuation values or the $\mu$-map from the measured ACFs as Eq.(1.6) is the inverse of Eq. (1.7) and the following relationship is obtained (similar to the projection data measured in the emission case):
The reconstructed $\mu$-map is subsequently used to determine the boundaries of different anatomical structures such as air, soft-tissues, and bones within the object. The specific known attenuation coefficient or $\mu$ value for the 511 keV photons is then assigned to each defined region, and this process is referred to as the segmentation. A noise-free $\mu$-map is generated after the segmentation, and the $\mu$-map is forward projected to form the attenuation or ACF sinogram data set. The sinogram data set is then used before or during the image reconstruction task to correct for the attenuation. In addition, further improvement in the attenuation correction can still be achieved by improving the initial estimate of the reconstructed $\mu$-map. Methods such as using a scaled CT image or improving the scatter correction in the transmission scan are currently under development [29, 30].

1.8.5 Randoms Correction

As mentioned in Section 1.6, random coincidence occurs when two photons from different annihilation sites are detected within the finite coincidence timing window thus being registered as a coincident event by accident. The inclusion of random coincidences causes degradation in image resolution as well as absolute quantitation, and therefore, it must be corrected. The rate of the random coincidences or randoms rate along the LOR associated with a detector pair $i$ and $j$ depends on the singles rate at each detector and the duration of the coincidence timing window. The relationship is given by:

$$R_{LORij} = 2\Delta t S_i S_j$$

where $\Delta t$ is the duration of the coincidence timing window, $S_i$ is the singles rate at detector $i$, and $S_j$ is the singles rate at detector $j$. The factor of 2 presented in Eq.(1.9) is due to the overlap of the timing windows associated with the two singles.
The singles rate is directly proportional to the amount of activity injected; the randoms rate is therefore proportional to the square of the activity. The random fraction (RF) which is a measure of the amount of randoms within the measured prompts is given by:

$$RF = \frac{\text{Randoms}}{\text{Prompts}} \times 100\% = \frac{\text{Randoms}}{\text{Randoms} + \text{Trues}} \times 100\% \quad (1.10)$$

The trues in this case refer to the true coincidences which include the scattered and the unscattered true events. Since the number of true events is directly proportional to the activity and the number of random events is proportional to the square of the activity, the RF increases with the activity. As a result, the RF is usually significantly higher in the early frames of the dynamic studies (this is also due to that initially the activity is mostly outside the FOV) for short lived radioisotopes such as $^{11}$C, and the RF decreases as time progresses. Consequently, the detection of randoms is one of the limiting factors for the maximum amount of radio-tracer injected. Although the randoms rate along each LOR can be estimated from the singles rates according to Eq.(1.9), this approach is not commonly used since it requires the singles rate to be known precisely at the crystal level, and the energy discrimination of the singles as well as the dead time difference between the singles and the coincidences needs to be taken into account. Alternatively, the randoms correction is typically done using the delayed coincidence measurement.

The basic idea of the delayed coincidence measurement is to introduce a delay in the coincidence timing window so that the probability of detecting the photons originating from the same annihilation sites (i.e. the trues) is zero. The delay is longer than the duration of the coincidence timing window (typically a few nano-seconds) so that the detected coincidences are all randoms. Moreover, the distribution of the randoms in the measured prompts is the same as that in the delayed coincidence measurement since the singles which produce randoms are not time correlated. The method provides an accurate estimate of the randoms when the counting statistics is high. For frames with low number of counts, the Poisson variance in the delayed coincidence measurement can significantly increase the noise in the reconstructed image. As a result, variance reduction techniques have been proposed and developed [31, 32] to minimize the noise.
in the delayed coincidence measurement. For the HRRT, a hybrid method [33], which iteratively calculates the singles rate from the delayed coincidence measurement and then estimates the randoms rate using Eq.(1.9), has been adapted since the typical variance reduction techniques cannot be applied directly to the HRRT data due to the various rebinning and compression schemes used to reduce the size of the data set (see Section 1.6.3).

The variance-reduced or ‘smoothed’ randoms sinograms are generated at the end of the process, and the sinogram data set is then used before or during the image reconstruction. In dynamic studies, the conventional method is to estimate the smoothed randoms for every single frame even though the randoms distribution is not sensitive to the change in the tracer distribution as will be discussed shortly with the scatter correction.

### 1.8.6 Motion Correction

In typical brain PET scanning, patients need to stay on the bed and inside the scanner for hours. As a result, patient movements during the scan are expected. Although head constraints such as thermoplastic masks have been applied to minimize the patient motion, they still do not eliminate it. With the continuous improvement in the PET image or spatial resolution, small movement has become a significant source of degradation in the image resolution and thus must be corrected. Typical motion correction methods involve using an external motion tracking device to monitor patient movements, and the motion data are incorporated into the reconstruction task.
A motion tracking system, Polaris™ (Northern Digital, Waterloo, Ontario, Canada), has been used for the HRRT as shown in Figure 1.15. Polaris™ consists of near infrared sources, cameras, and trackers. The trackers are retro-reflective spheres which are secured on a plastic panel, and at least 3 spheres which form a plane need to be used (4 spheres are typically used in case one of them goes outside the FOV). The panel is then attached to a neoprene cap which is worn by the patient during the scan. Polaris™ then tracks the movements or change in orientation of the patient with respect to a reference position (typically the position or axes of the scanner is used as the reference). The 3-D translations (x, y, and z) and rotations are generated and are subsequently used during the image reconstruction as the address of the LORs for the measured coincidence events is motion-corrected back to its original position (more accurate methods also involve modifying the system matrix and the sensitivity image [34, 35]).

Patient movements also influence the accuracy of the model-based scatter correction method (as will be discussed in Chapter 3) due to its use of the transmission (μ-map) and the emission data. Therefore, when there is a spatial mismatch between the transmission and the emission scan due to patient motion, the magnitude and the spatial distribution of the scatter estimate are very likely to be inaccurate as will be elaborated in greater detail in Chapter 3. For the HRRT, the implementation of accurate motion correction algorithm and compensation method is still work in progress. As a result, the
studies with the minimal amount of motion were chosen to validate the proposed quantification and optimization methods.

1.8.7 Scatter Correction

Scatter correction is a critically important but also challenging step in the task of quantitative PET imaging. This is particularly the case in 3D PET where the scatter fraction is typically as high as 30-50%. As mentioned previously, the purpose of scatter correction in PET is to remove the scattered events from the measured coincidences or prompts (in the emission case). A number of techniques have been proposed and investigated [36, 37, 38]. The most widely used scatter correction in 3D PET is the model-based single scatter simulation (SSS) which calculates the single scatter distribution using the Klein-Nishina equation based on the acquired transmission and emission data following by the tail-fitting to the measured true coincidences [2, 8]. This approach is attractive since it is based on the fundamental physical principles (though the tail-fitting process causes bias for some cases as will be elaborated in Chapter 3 and 5). Similar to the randoms estimation, the conventional scatter estimate is calculated for each temporal frame in the dynamic studies since the scatter distribution changes according to the change in the radio-tracer distribution. The scatter sinogram data set is thus generated for each frame and used before or during the reconstruction task. However, this results in the undesirable consequence that the time required to perform the scatter/randoms correction is directly proportional to the number of frames in the dynamic studies as well as poor scatter estimates due to the low number of acquired counts and/or high RFs within a single (short) dynamic frame. As a result, a novel scatter calibration (see Chapter 5) and a practical scatter and randoms approximation (see Chapter 6) technique has been developed. More details about the conventional scatter correction are presented in Chapter 3.

1.8.8 Positron Range, Photon Non-collinearity, and Resolution Modeling

The physical limitations of PET image or spatial resolution are the radioisotope dependent positron range and photon non-collinearity. As described in Section 1.1, the positrons travel a short distance which is referred to as the positron range before they reach thermal energies and annihilate with electrons (see Figure 1.1). The distances they
travel depend on their energies and are different from positron to positron since the emitted positrons have a wide range of energy due to the emissions of electron neutrinos. The positron range is also radioisotope dependent which is characterized by the maximum or mean positron energy emitted from different radioisotopes. Table 1.3 shows the maximum and mean positron energies together with the mean positron range in water for some of the commonly used radioisotopes. Consequently, the reconstructed PET image is blurred since the points of annihilation do not correspond to the points of decay perfectly thus resulting in the degradation in the spatial resolution.

<table>
<thead>
<tr>
<th>Radioisotope</th>
<th>$E_{\text{max}}$ (MeV)</th>
<th>$E_{\text{mean}}$ (MeV)</th>
<th>Mean Positron Range in Water (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$^{11}$C</td>
<td>0.96</td>
<td>0.33</td>
<td>1.1</td>
</tr>
<tr>
<td>$^{18}$F</td>
<td>0.63</td>
<td>0.20</td>
<td>0.6</td>
</tr>
<tr>
<td>$^{13}$N</td>
<td>1.20</td>
<td>0.43</td>
<td>1.5</td>
</tr>
<tr>
<td>$^{15}$O</td>
<td>1.73</td>
<td>0.70</td>
<td>2.5</td>
</tr>
<tr>
<td>$^{82}$Rb</td>
<td>3.40</td>
<td>1.39</td>
<td>5.9</td>
</tr>
</tbody>
</table>

*TABLE 1.3: Positron range and energy for typical PET radioisotopes*

The second fundamental constraint to the PET spatial resolution is the photon non-collinearity. The annihilation photons are perfectly collinear only if the total momentum of the electron-positron pairs before annihilation equals zero. Otherwise, a deviation from the collinear path is obtained as a consequence of the momentum conservation; this is referred to as the photon non-collinearity. This deviation shifts the assigned LOR from the actual path of the annihilation photons thus degrading the spatial resolution of the reconstructed images. The shift in LOR depends on the size of the scanner for the same angle of deviation as illustrated in Figure 1.16; i.e. the bigger the diameter the higher the degradation. The angle of deviation in water has a FWHM of 0.5° which corresponds to a resolution degradation or blurring of ~1.1 mm for two detectors separated by 50 cm.
As an effort to further improve the spatial resolution, resolution recovery using the point spread function (PSF) has been proposed and developed to model the blurring effect due to the physical limitations of PET [39]. The basic concept is that since the measured data are blurred by the physical limitations, one can model the effects by introducing blurring using PSF in the image reconstruction (whether modeling in the system matrix or blurring in the image estimate), and better spatial resolution has been achieved [39, 40, 41].

1.8.9 Image Quantitation Calibration
Once the images are reconstructed including all the aforementioned corrections, the last step is to convert the image intensity or count rates into the absolute radioactivity concentration. This calibration is done by scanning a uniform phantom (a cylinder...
which contains radioactivity with a uniform distribution) with a known activity concentration. The data acquired are then reconstructed, and the calibration factors are obtained by dividing the known concentration by the image intensity for the phantom. The calibration factors are subsequently used to convert the reconstructed images (e.g. from humans) into maps of activity concentration.
Chapter 2

Image Reconstruction
2.1 Chapter Overview
This chapter describes and compares the reconstruction algorithms with regard to linear versus non-linear, histogram-mode versus list-mode, and static versus dynamic PET imaging. It provides the basic principles of the image reconstruction methods used in this work. It then finishes with the potential challenges in typical dynamic PET imaging and the image analyses performed throughout the course of this work.

2.2 Introduction
Image reconstruction in PET refers to the operation which estimates the original activity distribution from the coincidence measurements of the scanner. The reconstruction algorithms are typically labeled in two categories: analytic image reconstruction (linear) and iterative statistical image reconstruction (mostly non-linear). In this chapter, the reconstruction methods are further divided into sub-categories such as list-mode and histogram-mode reconstructions as well as static and dynamic imaging.

2.3 Linear vs Non-linear
Linear reconstruction methods refer to algorithms which have a linear correspondence in operations between the measured data (sinograms) and the reconstructed image. For example, if one applies an operation on the sinogram data and then reconstructs the data using a linear reconstruction algorithm after the operation, the resultant reconstructed image will be the same as if one reconstructs the data first and then applies the operation on the reconstructed image. On the other hand, this is not the case for the non-linear reconstruction methods. One of the most widely used analytic (linear) reconstruction algorithms is the filtered back projection (FBP) which was used to validate the scatter and smoothed randoms sinogram data in this work (see Chapter 6) and will be described next.

2.3.1 Filtered Back Projection (FBP)
In analytic image reconstruction, the measured coincidence (projection) data are modeled by a line integral of a source distribution given by (see Figure 1.6 for reference):
\[ p(r, \theta) = \int_{-\infty}^{\infty} f(x, y) ds \]  \hspace{1cm} (2.1)

where \((r, s)\) is the rotated coordinate system of \((x, y)\) by the angle \(\theta\):

\[
\begin{bmatrix}
  x \\
  y
\end{bmatrix} =
\begin{bmatrix}
  \cos \theta & -\sin \theta \\
  \sin \theta & \cos \theta
\end{bmatrix}
\begin{bmatrix}
  r \\
  s
\end{bmatrix} \tag{2.2}
\]

The approach to obtain the measured source distribution is to 'back-project' the projection data by assigning equal weight along the LORs from each angle. The simplest back-projection is given by:

\[ b(x, y) = \int_0^{\pi} p(r, \theta) d\theta \]  \hspace{1cm} (2.3)

As shown in Figure 2.1, the streaks creating the star-like artifact decrease as the number of views of the back-projection increases. However, the back-projected image exhibits a \(1/r\) blurring due to the additional overlapping of the back-projected LOR around the source, where \(r\) is the radial distance from the source point.
Therefore, the back-projected image must be filtered in order to obtain the original source distribution, and the filtering can be easily applied in the frequency space. The reconstruction method which performs the filtering after the back-projection of the data is known as the backprojection-filtering (BPF) [43]. This method starts with the back-projection of the measured data followed by transforming the back-projection into the frequency space (Fourier transform) and then performs the inverse Fourier transform after applying the filter in the frequency space. Due to the linearity of the line integral, the back-projection step and the Fourier transform can be interchangeable. Consequently, the projection data are first Fourier transformed and filtered in the frequency space, and the back-projection is performed after the inverse-Fourier transform of the previous step. This alternative version of BPF is called the filtered back projection (FBP) [44]. For the continuous case, the FBP is given by:

\[
\begin{align*}
    f(x, y) &= \int_\theta \left( \int_v e^{2\pi i v r} |v| P(v, \theta) dv \right) d\theta \\
    P(v, \theta) &= \text{Fourier transform of } p(r, \theta) = \int_r p(r, \theta) e^{-2\pi i vr} dr
\end{align*}
\]

\[(2.4)\]
where $v$ is the spatial frequency and $|v|$ is the ramp filter applied to the Fourier transformed projection. FBP has been widely used in clinical settings due to its fast and low computational demands; however, it does not model the statistical noise and the discrete nature of the measured data nor the physical interactions such as the positron range and the photon non-collinearity. As a result, the iterative statistical reconstruction algorithms have been proposed. In this work, FBP was used to reconstruct and validate the scatter and randoms correction sinograms due to its fast nature and linearity as will be described in Chapter 6. In the next section, an iterative statistical algorithm will be discussed.

2.3.2 Iterative statistical Ordinary Poisson Expectation Maximization

In contrast to analytic image reconstruction, iterative statistical methods attempt to progressively update or refine the image estimate, rather than directly inverting the image transform as discussed above. Iterative algorithms typically start with an initial image estimate (e.g. a uniform low intensity image). The image estimate is then forward-projected to obtain the expected total counts for each LOR to compare with the measured data, and the resultant update factors are back-projected for each image volume element. The image is subsequently updated, and it enters the next iteration if the convergence or terminal condition is not met.

An iterative statistical algorithm typically consists of the following components:

- A finite parameterization of the image estimate (i.e. a finite number of discrete volume elements or voxels)
- A system matrix which describes the relationship between voxels and LORs and models the physical interactions of the system
- A statistical model (of noise) for how the measurements vary around their expectations
- An objective function to be maximized to find the image estimate
• An iterative algorithm which maximizes or progressively increases the objective function as it iterates

The finite parameterization of the image estimate is illustrated in Figure 2.2. The expected projection along a LOR is modeled by:

\[
\bar{n}_i = \sum_{j=1}^{J} p_{ij} \lambda_j = FP_i(\lambda_j)
\]

(2.5)

where \(\bar{n}_i\) is the expected (mean) total number of events or counts for the \(i^{th}\) LOR, \(p_{ij}\) is the system matrix, \(\lambda_j\) is the mean activity in voxel \(j\), and \(FP_i(\lambda_j)\) represents the forward-projection along the \(i^{th}\) LOR. The summation goes through all the voxels \((1, \ldots, J)\) along the \(i^{th}\) LOR.

![LOR i](image)

FIGURE 2.2: An example of the finite voxel parameterization of the image estimate

The system matrix \(p\) describes the probability of an emission from voxel \(j\) being detected along the \(i^{th}\) LOR and can be decomposed into three matrix components [39, 45]:


\[ p = w g b \]  

(2.6)

where \( w \) is the weighting assigned to each LOR to account for the sensitivity variation due to normalization and attenuation, \( g \) is the geometric factor which describes how voxels contribute to a given LOR when forward-projecting and can assign partial values to voxels that are not entirely within a given LOR when back-projecting, and \( b \) is introduced to model the blurring effects such as due the positron range and photon non-collinearity (i.e. resolution modeling).

As all counting experiments exhibit Poisson statistics (i.e. variance is equal to the mean), the Poisson likelihood function is employed and is given by:

\[
L(N \mid \lambda) = \prod_{i=1}^{I} e^{-\bar{n}_i} \frac{\bar{n}_i}{n_i!} \sum_{i=1}^{I} \frac{n_i}{n_i!} 
\]  

(2.7)

where \( \bar{n}_i \) is given by Eq.(2.5), \( n_i \) is the measured projection along the \( i^{th} \) LOR, and the summation from \( i = 1 \) to \( I \) is over all LORs. The Poisson likelihood function \( L(N \mid \lambda) \) describes the probability of obtaining a particular set of measurements \( N \) given a particular activity distribution \( \lambda \). It is maximized when the difference between the expected or estimated projection and the measured projection as known as the Kullback-Leibler distance is minimized. Alternatively, one can maximize the natural log of the likelihood function or the log-likelihood function since the calculations are simpler and both reach their maximum for the same image. The log-likelihood function is given by:

\[
l(N \mid \lambda) = \ln(L(N \mid \lambda)) = \sum_{i=1}^{I} \left[ -\bar{n}_i + n_i \ln(n_i) \right] 
\]  

(2.8)

where the term \( \ln(n_i!) \) is dropped since it does not depend on \( \lambda \) thus has no effect on the maximization.
An iterative expectation maximization (EM) algorithm has been developed [46] and can be shown to produce an image which increases the log-likelihood function as it iterates and eventually converge to a solution which maximizes both Eq.(2.7) and (2.8). The EM algorithm is given by:

\[
\lambda_{m+1}^j = \frac{\lambda_j^m}{\sum_{i=1}^l p_{ij}} \sum_{i=1}^l \left( p_{ij} \frac{n_i}{FP_i(\lambda_j^m)} \right)
\]  

(2.9)

where \( \lambda_j^m \) is the estimated activity in voxel \( j \) for the \( m^{th} \) iteration, and \( \sum_{i=1}^l p_{ij} \) is referred to as the sensitivity image and will be abbreviated as \( S_j \) from here on. The sensitivity image provides the normalizing factors for each voxel for the image estimate (corrections such as for the normalization and attenuation are embedded here). The forward and back-projection tasks are operated by the system matrix in the EM algorithm. They can be either voxel-driven in which the contributions from all LORs are computed for each voxel (many-to-one) or LOR-driven in which the contribution from a single LOR is assigned to all voxels along the LOR (one-to-many) when back-projecting.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{flowchart.png}
\caption{The flow chart of the iterative reconstruction algorithms}
\end{figure}

The flow of the algorithm is depicted in Figure 2.3. As mentioned previously, the initial image estimate is first forward-projected into the LOR space and compared to the measured projection data. The ratio is then computed for all LORs and back-projected.
into each voxel as the update factor. The image is subsequently updated (in the image space) and enters the next iteration if the terminal condition is not met. Although the resolution and contrast improves as the number of iteration increases, the noise in the image also increases due to the noise in the measured data. The maximum number of iterations or the terminal condition is thus a compromise between the resolution, contrast, and noise derived from phantom studies (with similar size and structures as the object). Regularization such as applying a Gaussian filter (if no resolution modeling) after image reconstruction is typically applied to reduce the noise.

The EM algorithm which takes the scattered and random coincidences into account is referred to as the Ordinary Poisson Expectation Maximization (OP-EM) and is given by:

\[ \tilde{\lambda}_j^{m+1} = \frac{\tilde{\lambda}_j^m}{S_j} \sum_{i=1}^{l} P_{ij} \frac{n_i}{FP_i(\tilde{\lambda}_j^m) + \bar{r}_i + \bar{s}_i} \]  

(2.10)

where \( \bar{r}_i \) is the expected mean random events or counts along the \( i^{th} \) LOR, and \( \bar{s}_i \) is the expected mean scattered events along the \( i^{th} \) LOR. The OP-EM algorithm is intuitive since the measured projection data or prompts contain the unscattered true, scattered, and random events. By incorporating the expected randoms and scatter estimates for each LOR in the denominator of Eq.(2.10), the image estimate which corresponds to the unscattered true events can be obtained. Due to the computation intensive nature of the 3D (which includes the oblique LORs into the calculation) OP-EM algorithm, it has not been widely used until recently. In the next section, the accelerated OP-EM algorithms will be described along with different modes of iterative reconstruction.

2.4 Histogram-mode vs List-mode

Up until now all the reconstruction algorithms mentioned previously are sinogram-based, or so called histogram-mode reconstruction. In this section, a different type of reconstruction algorithm which directly reconstructs the image from the list-mode file
will be introduced. The accelerated version of the histogram-mode as well as list-mode EM algorithms will be presented.

### 2.4.1 Histogram-mode Reconstruction

The general procedures for the histogram-mode reconstruction are illustrated in Figure 2.4. For modern PET scanners with list-mode acquisition capability, the measured coincidence events are stored in the list-mode file. The data are then ‘histogrammed’ into sinograms or projections in the LOR space, and the images are reconstructed from the sinograms.

![List-mode File](image)

**FIGURE 2.4:** The reconstruction procedures for the histogram-mode reconstruction (note that the grey-scale of the image is inverted)

In the original EM algorithm, the image estimate is updated after going through the whole data set in the sinograms, which is not very efficient considering that ~100 updates are typically required to produce a good image (i.e. to reach the optimum or terminal condition). As a result, the accelerated EM algorithm named ordered subset expectation maximization algorithm (OSEM) has been proposed [47]. In OSEM, the LORs with the same azimuthal angle $\theta$ are grouped into subsets as depicted in Figure 2.5.
A subset can also be a combination of multiple groups of LORs with evenly spaced θ depending on the number of subsets specified. Instead of updating the image after going through all the LORs in the sinograms, the OSEM algorithm updates the image estimate after going through the LORs in each subset and produces multiple updates per iteration. For example, if one sets 16 subsets with 6 iterations in the OSEM, a total of 96 image updates will result, and the image estimate has been found to be equivalent to running 96 iterations using the original EM algorithm provided that there is a sufficient number of counts in each subset [47]. As a result, a 16 times faster performance is achieved using OSEM in this case. The histogram-mode 3D-OP-OSEM is given by a very similar expression as Eq.(2.10):

\[
\lambda_{j}^{m,l+1} = \frac{\lambda_{j}^{m,l}}{S_{j}^{l}} \sum_{i \in s_{i}} P_{ij} \frac{n_{i}}{FP_{i}(\lambda_{j}^{m,l}) + \bar{r}_{i} + \bar{s}_{i}}
\]

(2.11)

where \(\lambda_{j}^{m,l}\) is the image estimate in voxel \(j\) for the \(m^{th}\) iteration and \(l^{th}\) subset, \(S_{j}^{l}\) is the sensitivity image obtained from the \(l^{th}\) subset, and the back-projection step goes through

\[\text{FIGURE 2.5: An example of the spatial angular subset grouping scheme used in the histogram-mode OSEM algorithm}\]
the LORs within the $t^{th}$ subset ($s_t$). For simplicity, the histogram-mode 3D-OP-OSEM will be further abbreviated as 3D-OP algorithm from here on.

One important observation from the histogram-mode algorithms (Eq. (2.9), (2.10), and (2.11)) is that the algorithms loop over the LORs which is independent of the number of acquired events, and the number of LORs is a fixed quantity. Consequently, the reconstruction time cost for the histogram-mode reconstruction is a constant (except the histogramming from the list-mode data to the sinograms takes a bit longer when there are more counts but is still short compared to the reconstruction time). In contrast, an algorithm which loops over the events will be described next.

### 2.4.2 List-mode Reconstruction

In list-mode reconstruction, the list-mode data are directly reconstructed to form the image estimate, bypassing the histogramming step or sinogram formation as illustrated in Figure 2.6.

![List-mode File](image)

**FIGURE 2.6:** The reconstruction procedures for the list-mode reconstruction

The list-mode OP-EM algorithm is given by:

$$
\mathcal{L}_j^{m+1} = \frac{\mathcal{L}_j^m}{S_j} \sum_{k=1}^{K} P_{ik} \frac{1}{FP_{ik} (\mathcal{L}_j^m + \bar{r}_{ik} + \bar{s}_{ik})} (2.12)
$$
where the differences in notation between Eq.(2.10) and (2.12) are that (i) the projection data \( n_i \) is replaced by 1 since the algorithm goes through the events one-by-one instead of the LORs, (ii) the summation goes through all the events from \( k = 1 \) to \( K \) instead of the LORs, and (iii) the index \( ik \) represents the \( i^{th} \) LOR associated with the \( k^{th} \) event. Eq.(2.12) is equivalent to Eq.(2.10) except that the back-projection for the histogram-mode can be either voxel-driven or LOR-driven (voxel-driven back-projection is typically used), whereas the back-projection for the list-mode has to be LOR-driven. The reason is that the projection data for all the LORs are not available (reminder: in histogram-mode reconstruction each LOR contains the corresponding total number of coincident events after the histogramming process) as the list-mode algorithm goes through each event one-by-one along the list-mode file. In addition, the reconstruction time cost is directly proportional to the number of acquired events in the list-mode reconstruction.

In subsetized histogram-mode reconstruction, the sinogram data are divided into angular subsets in order to speed up the convergence of the reconstruction and reduce the number of iterations required as discussed in Section 2.4.1. On the other hand, in subsetized list-mode reconstruction, the list-mode data acquired can be divided into 'temporal' subsets [39, 45]. In contrast to the (spatial) angular subsets in which the events are summed over the entire scan duration for each LOR in the histogram-mode reconstruction, the list-mode algorithm goes through the events in the time domain, and thus the list-mode file is divided in the time domain to form the temporal subsets.
As depicted in Figure 2.7, the temporal subsets are extracted from multiple divisions in the list-mode data in order to minimize potential time-varying inconsistencies in the activity distributions across the sequential temporal subsets thus consistently iterating towards an average activity distribution within the scan. The estimated image is therefore updated after processing through each regrouped temporal subset. The number of the divisions applied depends on the rate of variation of counts in the scan; i.e. it has been empirically found that when the number of counts changes rapidly within a scan (as is usually the case during the first frame of dynamic scans as will be elaborated in the next section), more divisions need to be applied in order to minimize the variation of the number of counts (eventually resulting in image intensity) in between the subsets thus resulting in unbiased image updates between the subsets. For example, it has been found for our dynamic non-human primate studies that at least 30 divisions need to be applied in the first dynamic frame and 10 divisions are sufficient for the rest of the frames [3] (see more detail on dynamic imaging in the next section).

The subsetized list-mode OP-EM algorithm (will be abbreviated as OP-LMEM from here on) is thus given by:
\[
\mathcal{X}_{j}^{m,l+1} = \frac{\mathcal{X}_{j}^{m,l}}{S_j} \sum_{k \in S_l} p_{ikj} \frac{1}{FP_{ik}(\mathcal{X}_{j}^{m,l}) + r_{ik}^{l} + s_{ik}^{l}}
\]  
(2.13)

where the back-projection step goes through all the events within the \(l^{th}\) subset \((S_l)\). Note that the scatter and randoms estimates need to be treated differently in each temporal subsets since each temporal subset is basically a shorter scan by itself. Therefore, the scatter and randoms estimates need to be rescaled in each temporal subset. The re-distribution or rescaling of the scatter and randoms sinograms in the temporal subsets will be presented in more detail in Chapter 4. The OP-LMEM algorithm is fairly new with regard to implementation, and the scatter-corrected algorithm has only recently been implemented and validated [3, 5] as a part of this dissertation and will be described in Chapter 4.

### 2.5 Static vs Dynamic

Static image reconstruction refers to reconstructing a single image from the data of the entire scan. Images from the FDG scans, for example, in oncology are typically static. Static images are useful in locating tumors and contain a high number of counts from the point of view of image reconstruction. In dynamic imaging, one wants to monitor the activity concentration as a function of time. The scan is thus divided into a number of ‘dynamic’ or temporal frames with specified durations for each frame in typical dynamic reconstructions. The dynamic frames are then reconstructed individually to form the dynamic images. In image analysis, the time activity curves (TACs) are generated by placing ROIs in each dynamic image, and biological parameters such as the binding potential (BP) values are extracted from the TACs. Figure 2.8 shows an example of a set of dynamic images with ROIs and the TAC for one of the ROIs (rt-cau: right caudate) obtained from images reconstructed using the 3D-OP and the OP-LMEM algorithms. For the tracers considered in this work, the scan can be differentiated into the early and later stages. During the early stage, the tracer distribution and the count rate changes very rapidly as the activity moves from outside the FOV into the brain, whereas the tracer distribution stays relatively constant and the count rate changes slowly during the later stage. Therefore, the reconstruction methods (such as the
temporal subset formation described above, the scatter calibration, and the practical approximation as will be presented in Chapters 5 and 6) need to be treated differently for the images within early stage of the scan.

As compared to static image reconstruction in which all the acquired events or counts from the entire scan are utilized to form a single image, the number of counts within each frame is lower and typically exhibit a wide variation along with the count rates in between the frames in dynamic image reconstruction. Furthermore, quantitative corrections such as the scatter and randoms estimations are performed for each dynamic frame thus rendering potentially inaccurate estimates due to the low number counts as well as the reconstruction time dependence on the number of dynamic frames. These are the aspects that are addressed in the present work.

2.6 Image Analysis

Once the images are reconstructed, they can be analyzed and compared in a number of ways depending on the context. When comparing similarities and differences in a single image reconstructed between different algorithms, analyses such as profile, contrast, and noise comparisons are typically performed. When comparing a set of dynamic images reconstructed, for example, in neurology between different algorithms, dynamic analyses such as TAC and BP value comparisons are performed. For instance, to validate the scatter-corrected list-mode reconstruction (OP-LMEM) algorithm in

\[ FIGURE 2.8: Dynamic imaging in PET \]
Chapter 4, the conventional histogram-mode reconstruction (3D-OP) which has been established to be quantitatively accurate [47] was used as the gold-standard, and the similarities and differences in between individual images as well as dynamic analyses were performed. These image analyses are summarized in the following sections.

2.6.1 Image Profile Analysis

The profile analysis can be performed along the transaxial and the axial direction of a PET image as depicted in Figures 2.9 and 2.10.

**FIGURE 2.9:** An example of the transaxial profile analysis for a phantom image with and without the scatter correction reconstructed with 3D-OP and OP-LMEM algorithms
The transaxial profile analysis requires a line profile placement on the cross section of the image, and the voxel values are displayed along the line profile across the object activity distribution. The axial profile, on the other hand, usually requires a ROI placement, and the mean voxel value within the ROI is displayed along the axial direction of the object activity distribution. The axial profile can also be done in the voxel-level. The profile comparisons are useful to examine the differences between the images when the differences are not obvious via visual inspection at the voxel-level. The profile analysis can also be performed to check the transaxial and axial uniformity of the object activity distribution.

2.6.2 Contrast and Noise Analyses

Other than the profile analysis, the contrast and noise analyses are typically performed when comparing individual images. A PET image is typically a combination of hot, cold, and background regions. A hot region is defined as the region which contains a relatively high activity concentration, whereas a cold region is referred to as the region with no activity concentration. Additionally, the background region is referred to as the region with relatively low activity concentration as compared to the hot region. According to the NEMA NU 2001 protocol [48], the contrast of a cold region is defined as:

**FIGURE 2.10: An example of the axial profile analysis**
\begin{equation}
Contrast_{\text{cold}} = \left(1 - \frac{C_C}{C_B}\right) \times 100\% \tag{2.14}
\end{equation}

where $C_C$ is the mean counts within the cold region, and $C_B$ is the mean counts within the background region in the image.

The contrast of a hot region is defined as:

\begin{equation}
Contrast_{\text{hot}} = \left(\frac{C_H}{A_H} - 1\right) \times 100\% \tag{2.15}
\end{equation}

where $C_H$ is the mean counts within the hot region in the image, $A_H$ is the measured activity for the hot region, and $A_B$ is the measured activity for the background region.

The noise in the reconstructed image is characterized by the coefficient of variation for the background region, and it is defined as:

\begin{equation}
C_v = \frac{\sigma}{\mu} \tag{2.16}
\end{equation}

where $C_v$ is the coefficient of variation, $\sigma$ is the standard deviation, and $\mu$ is the mean of the voxel or ROI values.

The coefficient of variation can be computed at the voxel-level as well as at the ROI-level, and they are typically referred to as the voxel-noise and the ROI-noise respectively. In addition, the contrast and noise analyses can be very useful when determining the terminal condition for the iterative reconstruction algorithms.

2.6.3 TAC and BP Value Comparisons
In dynamic brain imaging, the final goal of the image reconstruction is generally to extract the biological parameters such as the BP values from the TACs. For example in
typical $^{11}$C-raclopride studies (see Section 1.5.3), one places the ROIs around the striatum regions which govern the motor functions and balance in the brain and have specific receptor binding to the tracer as well as the cerebellum regions which have unspecific binding and are used as the reference regions [49]. The TACs are then generated for these ROIs and the corresponding BP values are obtained. For example, from the BP values, one can study how a drug or a placebo affects the neuro-receptor binding. Therefore, it is very important to compare the TACs and BP values between different reconstruction methods other than comparing the individual images separately.

As shown in Figure 2.8, the specific ROIs are drawn and placed in each dynamic frame, and the mean counts or activity within the ROI are computed for each frame to generate the corresponding TAC. The BP values are subsequently extracted from the TACs or directly from the reconstructed images using appropriate mathematical models [49].

The analyses described above were performed throughout Chapters 4, 5, and 6 when validating the list-mode reconstruction, a novel scatter calibration, and a practical scatter and randoms approximation as well as the combined method.
Chapter 3
The Most Widely Used Scatter Correction in 3-D PET
3.1 Chapter Overview
The chapter starts with an introduction on the effect of scatter and scatter correction in 3-D PET. A summary of various scatter correction methods is presented followed by the description of the most widely used scatter correction method: the single scatter simulation (SSS). The considerations for the scatter estimation in dynamic imaging are subsequently discussed, and this chapter finishes with the description and various applications for the estimated scatter fraction (SF).

3.2 Introduction
As described in Section 1.6, a portion of the scattered events is rejected by the use of the energy discrimination window. However, the measured true coincidences from brain studies still contain 30-50% scatter fraction (which is the fraction of the scattered events within the measured true coincidences; see Section 3.6) in 3-D acquisition due to the limited energy resolution of the PET scanner. The inclusion of the scattered events degrades the spatial resolution of the image as the assigned LOR does not correspond to the actual path of the annihilation photons (see Figure 1.2) and thus must be corrected. Figure 3.1 shows the image of a contrast phantom which consists of a hot and two cold spots (water and Teflon) with (right) and without (left) the scatter correction.

![Figure 3.1](image)

*FIGURE 3.1: A contrast phantom reconstructed with (right) and without (left) the scatter correction; the single scatter simulation (SSS) was used as the scatter correction in this case*
From the image without the scatter correction, one can clearly observe that more scatter-related counts are assigned to the Teflon (T) spot, which is a cold spot, due to its higher density as compared to the water (W) spot, and the central region of the phantom also contains more scattered counts as compared to the region near the boundary of the phantom. Other than the degradation in spatial resolution, the contrast loss especially in the cold regions in the reconstructed image is another consequence of the contamination of the scattered events in the measured coincidences. The radioactivity concentration obtained from the image without the scatter correction is also overestimated since the attenuation correction recovers the scattered and absorbed coincidence events with the intrinsic assumption that all the scattered events have been removed. With an appropriate scatter correction (the SSS method in this case), one can observe the recovery of contrast; i.e., similar contrast between the water and the Teflon spot as they are both cold spots. Uniformity for the background region is also achieved as depicted in Figure 3.1 (right).

### 3.3 Scatter Correction Techniques in PET

A number of scatter correction methods has been proposed and investigated in the literature [36, 37, 38, 50, 51, 2, 8]. They are categorized as the following:

**Tail-fitting method [36]:** The simplest one is the tail-fitting method in which the spatial distribution of the scattered events is assumed to be described by a polynomial shape or a Gaussian shape. This method also incorporates the observation that since there should be no activity outside the boundary of the object, the measured true coincidences outside (which are termed the ‘tail’ of the measured trues) are all contributed from the scattered events. Therefore, given the spatial distribution of the scattered events the magnitude of the scatter distribution can be obtained by fitting the tail of the estimated scatter distribution to the tail of the measured trues as illustrated in Figure 3.2. This method is easy to implement and accounts for the scattered events from outside the FOV. However, it does not account for the different anatomical structures and the asymmetrical scatter distribution in the objects.
Convolution-subtraction method [37]: The basic idea of the convolution-subtraction scatter correction methods is that the scatter estimate is a convolution of the unscattered true coincidences with a pre-defined scatter function (e.g. a mono-exponential) with scatter parameters such as the scatter fraction (obtained from the line source measurements in air and in the phantom). The iterative process starts with convolving the measured true coincidences which include both scattered and unscattered true events (since the unscattered trues are not known initially) with the scatter function to generate the scatter estimate in the first iteration and then subtracts the scatter estimate from the measured true coincidences to obtain the unscattered trues. The unscattered trues are then used in the next iteration and convolve with the scatter function to obtain the next update of the scatter estimate. The drawback of convolution-based methods is that the

FIGURE 3.2: Ideal projection data from the unscattered true events (left) and realistic projection data (red) which contains both unscattered and scattered true events (right); the magnitude of the estimated scatter projection distribution (pink) can be corrected by fitting the tail of the estimated scatter distribution to the tail of the measured trues close to the boundary of the object.
nature of the scatter function with the empirically derived scatter parameters can introduce bias in more complex non-uniform objects.

Energy-window method [38, 50]: The energy-window-based scatter correction techniques can be differentiated into the original dual energy window method (DEW), the estimation of trues method (ETM), and the multi-spectral method. The DEW method uses a lower energy window below the photo-peak and the original energy window in the photo-peak. The measured events from the lower energy window are used to estimate the scatter distribution, and the resultant scatter estimate is smoothed and scaled according to the scatter fraction obtained from the line source measurements in air and in the phantom. This method is also easy to implement; however, the scatter distribution obtained from the lower energy window only accounts for the large angle scattering which has a different distribution from the small angle scattering in the photo-peak measurement. The ETM technique uses a higher energy window above the 511 keV with the same upper energy threshold as the original photo-peak window. The measurement from the higher energy window is used to estimate the distribution of the unscattered true events, and the scatter estimate is obtained by subtracting the estimate of the unscattered true events from the original photo-peak measurement followed by scaling and smoothing. The main disadvantage of this method is the lack of statistics in the higher energy window measurement which introduces significant noise in the scatter estimate as well as in the reconstructed images. The multi-spectral method is the combination of the convolution-subtraction and the energy-window-based technique. A very large number of energy windows with the same upper energy threshold and different lower energy thresholds is used, and the acquired data from each window are fitted to mono-exponentials to differentiate the unscattered true, object scatter, and the detector scatter distributions. The position and energy dependent scatter function is then extracted from the fitted parameters and used in the convolution-subtraction method. The major drawback of this method is that it requires the hardware and software capabilities for multi-spectral data acquisition.

Monte Carlo simulation method [51]: Another technique to estimate the scatter distribution is the Monte Carlo (ray-tracing) simulations. The simulated annihilation
photons are propagated through the attenuating medium using random numbers to determine probabilistic locations and results of interactions. If a scatter interaction occurs, the new direction and energy of the scattered photon is determined using look-up tables derived from the physics of the system, and the history of all interactions is also recorded. Once both annihilation photons exit the object and are recorded by the simulated detectors, a coincidence event is recorded. The coincidence event is labeled as a scattered event if one or both of the photons contain any history of scatter interaction. Otherwise, an unscattered true event or a ‘primary’ event is recorded. After all the annihilations are simulated, the unscattered true and the scatter sinograms are generated. The sum of these two sets of simulated sinograms is then scaled to match the measured true coincidences from the experimental data. However, since the tracer distribution or distribution of the annihilation sites and the anatomical structures of the object are not known prior to the scan, these simulations require a reconstructed emission image and a transmission image as inputs. Furthermore, the input emission image should be scatter-corrected (a model-based scatter correction is typically used) in order to obtain accurate results from the simulation. As a result, the time cost to generate an accurate scatter estimate from the Monte Carlo simulation method is always longer than the model-based method or any method used to obtain the initial emission image. In addition, the simulation time is directly proportional to the number of simulated events, so it can be very time-consuming to simulate a fairly larger number of counts required to obtain an accurate estimate.

Model-based method: The most widely used scatter correction method is the model-based single scatter simulation (SSS) technique. A single scattered event refers to the event in which only one of the two annihilation photons has undergone the Compton scatter interaction. It has been shown that ~80% of the total detected scattered events are contributed from single scattered events for subject size of 20-30 cm [52, 2] (though smaller objects may contain more single scattered events). This method is typically used in conjunction with the tail-fitting method in order to compensate for the scatter from outside the FOV as the activity distribution and the attenuation map outside the FOV is unknown. It is highly favorable since it calculates the scatter distribution from the fundamental physical principles, is fast, and does not require any additional
measurements other than the transmission and emission scans. The details of the SSS method are described in the next section.

3.4 The Single Scatter Simulation (SSS)

The model-based SSS algorithm calculates the scatter distribution from an estimate of the activity and attenuation distribution within the object, a physical model of photon interaction, and scanner specific parameters [2, 8, 53]. Since the work presented in this thesis is based on this method, it will be described in more detail. The method typically consists of the following steps:

1. Generate a scatter-uncorrected image from the measured trues and define the attenuation distribution (μ-map) from the transmission scan.

2. Randomly distribute the scatter points within the scatter volume defined by the attenuation values.

3. For a LOR, calculate the number of events contributed from each scatter point along the photon path and sum over the contributions from all the scatter points to obtain the total scattered events along the given LOR.

4. Repeat for all LORs and obtain the scatter sinograms (note that typical SSS algorithms are optimized for speed, so the calculation is performed for a fewer number of LORs and the intermediate ones are obtained by interpolation while maintaining accurate results).

5. Scale the scatter sinograms by tail-fitting to the measured trues.
FIGURE 3.3: The illustration of the SSS calculation for a given LOR

The equation used to perform step 3 above is given by [2] (also see Figure 3.3):

\[
\overline{S}_{AB} = \int_{S} dV_{S} \left( \frac{\sigma_{\text{A}} \sigma_{\text{S}}}{4\pi R_{\text{AS}}^2 R_{\text{BS}}^2} \right) \frac{\mu}{\sigma_{\text{c}}} \frac{d\sigma_{\text{c}}}{d\Omega} \left[ I^{A} + I^{B} \right] \tag{3.1}
\]

where

\[
I^{A} = \mathcal{E}_{\text{AS}} \mathcal{E}^{'}_{\text{AS}} \left( \sum_{a} \frac{\mu^{a} ds}{\mu^{a} ds} + \sum_{b} \frac{\mu^{b} ds}{\mu^{b} ds} \right) \int_{S} \lambda ds,
\]

\[
I^{B} = \mathcal{E}^{'}_{\text{AS}} \mathcal{E}_{\text{SS}} \left( \sum_{a} \frac{\mu^{a} ds}{\mu^{a} ds} + \sum_{b} \frac{\mu^{b} ds}{\mu^{b} ds} \right) \int_{S} \lambda ds \tag{3.2}
\]
Vs is the scatter volume; $\sigma_{AS}$ and $\sigma_{BS}$ are the geometrical cross section of detector A and B for annihilation photons incident along AS and BS, respectively. $R_{AS}$ and $R_{BS}$ are the distances from detector A and B to the scatter point, respectively. $\mu$ is the linear attenuation coefficient, $\sigma_c$ is the total Compton scatter cross section per electron, $d\sigma_c/d\Omega$ is the Klein-Nishina scattering cross section, $\epsilon_A$ represents the detector efficiency for photons incident along AS, and $\lambda$ is the linear activity concentration obtained from step 1 above.

Note that the energy of a scattered photon depends on the scattering angle as given by Eq. (1.2), and the attenuation values as well as the detector efficiency are energy-dependent. In Eq. (3.2), all primed quantities are evaluated at the scattered photon energy, whereas the unprimed ones are evaluated at 511 keV. Eq. (3.1) and (3.2) can be understood by looking at the two photon paths. For the first path which corresponds to $f^A$ in Eq. (3.2), the annihilation sites occur between detector A and any one of the scatter points (i.e., $S_1$, $S_2$, $S_3$, or $S_4$). One of the annihilation photons is measured at detector A, and the other one undergoes Compton scatter and gets recorded at detector B as illustrated in Figure 3.3. The second path ($f^B$ in Eq. (3.2)) is the opposite of the first one in that the annihilation sites take place between detector B and the scatter point. One of the photons is measured by detector B, and the other one gets scattered at the scatter point and is recorded at detector A. The term $\mu/\sigma_c$ gives the electron density [54].

The incident photon flux on the scatter point contributed from the annihilation sites along the first path is given by $\int_{S}^{A} \lambda ds$. Therefore, by multiplying the incident photon flux by the Klein-Nishina scattering cross section, $d\sigma_c/d\Omega$, and the electron density, we obtain the total number of Compton interactions per scatter volume per solid angle. We then take the attenuation and the detector efficiency for the unscattered photon along AS and those for the scattered photon along BS into account by multiplying the previous product by $\exp(-\int_{S}^{A} \lambda ds)$ as well as $\epsilon_A \epsilon'_A$. Subsequently, we multiply the previous product by the solid angle term, $\frac{\epsilon_A \epsilon_B \sigma_{AS} \sigma_{BS}}{4\pi R_{AS}^2 R_{BS}^2}$. Finally, we sum the contributions over the
scatter volume $V_s$ to obtain the effective total scattered events along the first path for the given LOR. The effective total scattered events along the second path can be described accordingly. The total number of the scattered events for a given LOR is then obtained by adding the contributions from the first and the second path. Steps 4 and 5 are subsequently performed to form the scatter sinograms.

In step 5, the LOR or bin locations of the scatter tail (i.e. regions close to the boundary of the object used to perform the tail-fitting) are defined by the transmission scan. In terms of the attenuation sinogram values, the bins or LORs which do not go through any object contain values of 1s, and LORs or bins corresponding to the scatter tail region are thus defined by the attenuation sinogram value slightly higher than 1 [8]. The tail fitting process is done by a least squares fit for a scatter scaling factor, $f$, minimizing:

$$
\sum_b (t_b - f \cdot s_b)^2
$$  \hspace{1cm} (3.3)

where $b$ goes through all the bins or LORs within the tail defined from the $\mu$-map or attenuation sinogram, $t$ is the measured true coincidences obtained by subtracting the randoms estimate from the measured prompts thus making the RF play an important role here as will be discussed later, and $s$ is the scatter estimate from step 4. A single scatter scaling factor is then obtained for each scatter sinogram or plane since the scatter originating from outside the FOV contributes differently for different axial planes. In addition, an axial smoothness constraint (a boxcar filter) is applied to smooth the scatter scaling across the planes [8] as well as a positivity-constraint used to prevent the negative scaling of the scatter estimates as will be elaborated shortly. The above implementation of the SSS method has been applied for the HRRT by CPS/Siemens. Their implementation of the SSS method was also incorporated into the list-mode reconstruction by modifying the output of the SSS software as will be described in Section 4.2.

It has been suggested that the single scatter estimation from an overestimated emission image (i.e. without any scatter correction) compensates for multiple scattered events [2].
However, for objects with similar (e.g. human heads) or smaller size (e.g. small animals) compared to the mean free path of photon in the medium, almost all the scatter events are contributed from single scatter; therefore, single scatter estimated from an overestimated emission image is likely to overestimate the overall magnitude of the scatter distribution. As a result, the tail-fitting process is typically required not only to account for the scatter from outside the FOV but also to make the estimated scatter commensurate to the measured trues. An example of the typical scatter scaling factors across the planes for a human study acquired on the HRRT is shown in Figure 3.4 for a frame with 400M prompts and ~30% RF.

\[ \text{FIGURE 3.4: The scatter scaling factors (y-axis) across the direct planes (x-axis); the raw scaling factors are depicted in red, and the ones after applying the axial smoothness constraint (boxcar filter) are in green} \]

Note that the scaling factors obtained from the tail-fitting are generally less than 1.0 which reflects the initial overestimation of the scatter. Although the model-based SSS method works very well in typical static studies, challenges in dynamic imaging still remain to be overcome.
3.5 Considerations for Dynamic Imaging

In typical dynamic imaging, scattered (and random) events are generally estimated on a frame-by-frame basis. The justification for this approach is that the tracer distribution changes as a function of time, and therefore the scatter distribution changes accordingly. However, given the often highly variable number of events per frame typical of dynamic scanning and variable count rates (RFs) at which the acquisition is performed as depicted in Figure 3.5, the number of counts in each frame might not always be sufficient to produce an accurate scatter scaling across the planes.

Consequently, noisy scatter scaling and noisy axial scatter distribution is typically obtained from the tail-fitting even after applying the axial smoothness constraint for frames with a low number of counts due to the statistical variation of counts across the planes (see Figures 3.6 and 5.6), and the scaling is also imprecise due to the noisy tail of the measured data. One can improve the noisy scaling by using the trues obtained from subtracting the variance reduced randoms from the measured prompts during the tail-
fitting process; however, it does not eliminate the noisy scaling since the number of measured prompts is still low (see Figures 5.24, 5.25, and 5.27). Furthermore, a positivity-constraint is used to prevent the negative scaling of the scatter estimate when tail-fitting the scatter to the trues. Negative scaling occurs frequently for frames with a low number of counts and with a relatively high RF due to the negative true counts obtained from subtracting the measured random coincidences from the measured prompts. As a result, a pre-set positivity-constraint is typically applied on the scaling factors, which is equivalent to setting the negative true counts into some pre-set positive threshold during the tail-fitting process. However, this constraint is prone to introducing an overestimation bias into the scatter estimate. An example of the noisy scaling and the positivity-constraint is demonstrated in Figure 3.6 for a frame with 2M prompts and 94% RF in a human dynamic brain study.

**FIGURE 3.6:** Scatter scaling factors across the direct planes for a frame with 2M prompts and 94% RF; the raw scaling factors are depicted in red, and the ones after applying the axial smoothness constraint (boxcar filter) are in green.
Note that the truncation of the scaling factor at 0.25 with an upper threshold at 2.0 due to the pre-set positivity constraint empirically tested and provided by the manufacturer (CPS/Siemens) of the HRRT [55], and a 350% estimated SF was obtained from the above example. Eventually, the overestimated scatter leads to an underestimated activity concentration in the emission image. Alternatively, one might think that the tail-fitting process can be performed by adding the estimated scatter to the randoms and fitting the tail of that sum to the tail of the measured prompts which contains no negative counts. This approach can avoid the use of the positivity-constraint; however, it is not accurate since the estimated randoms distribution does not require any rescaling but only the scatter thus introducing a bias on the scatter distribution. Furthermore, the noisy and imprecise scaling can still occur due to the low number of prompts measured. The effects of high RF and/or with low number of counts in the scatter tail-fitting will be described in greater detail in Chapter 5. To remedy the problems in the tail-fitting as described above, a novel scatter calibration technique has been developed (see Chapter 5).

In addition, the conventional frame-based approach results in the undesirable consequence that the time and data storage required to perform the scatter/randoms correction is directly proportional to the number of frames in the dynamic studies. As a result, a practical scatter and randoms approximation technique has been proposed and developed in this work (see Chapter 6).

3.6 Scatter Fraction (SF) and Its Applications

Another useful quantity which can be extracted for the scatter estimate is the scatter fraction (SF) which is given by:

$$SF = \frac{Scatter}{Scatter + Unscattered\ Trues} \times 100\% = \frac{Scatter}{Prompts - Randoms} \times 100\% \quad (3.4)$$

As can be observed from Eq.(3.4), realistically SF should be always less than or in the worst case equal to 100% (which means all the measured trues are scattered events). The regional SF values depend on the size and density of the anatomical structure; e.g. high density regions such as the bones contain high SF values, whereas low density
regions such as air and water contain low SF values. The impact of the scatter estimate in the reconstructed emission image depends on the global and regional SF values, in that the reconstructed activity is more sensitive to the scatter estimate for regions with high SF values. For example, a change of 10% in the regional scatter estimate would correspond to a change of ~5% in the region of the emission image when the regional SF is ~50%; a change of 50% in the regional scatter estimate would correspond to a ~0.5% change in the region of the emission image when the regional SF is ~1%. For the practical scatter approximation as will be presented in Chapter 6, the limitation of the method can thus be checked using the cold regions with different densities or SF values.

With regard to the global SF value, any estimated global SF value higher than 100% would imply that the scatter estimate is inaccurate. Moreover, in dynamic human and animal brain studies, the global SF value is approximately constant as a function of time since the density of the brain tissues is very uniform and the anatomical structures do not change over time. Therefore, the global SF value can be used as a quality control (QC) for the scatter estimation in each dynamic frame. As mentioned in Section 3.4, the model-based SSS method uses the transmission scan (μ-map) as the definition for the regional scatter points as well as the location of the scatter tail. Consequently, the scatter estimate is sensitive to patient movement since the motion introduces mismatch between the transmission and emission scan. An example of a human brain study with motion data and the corresponding SF values is shown in Figure 3.7.
FIGURE 3.7: The temporal SF values with the corresponding motion data with and without motion compensation

The patient motion is shown as the displacement in the x, y, and z directions with respect to the scanner’s axis. One can observe an almost one-to-one correspondence between the displacement and the SF values. In addition, after applying a motion compensation, the SF values become much more consistent over time. As a result, the estimated SF can be used as a check for motion compensation techniques as well as a general check for any bias in the scatter estimate due to low number of counts and/or high RF. The proposed scatter calibration and approximation also make use of the approximately constant temporal SF value in the dynamic studies as will be described in Chapter 5 and 6.
Chapter 4

Implementation of the Model-Based Scatter Correction in List-Mode Reconstruction with Validation and a Dual Reconstruction Scheme
4.1 Chapter Overview

In this chapter, an Ordinary Poisson List-Mode Expectation Maximization (OP-LMEM) algorithm with a sinogram-based scatter correction method based on the Single Scatter Simulation (SSS) technique, and a randoms correction method based on the variance-reduced delayed-coincidence technique is described. This work has been published in Physics in Medicine and Biology [3]. The quantitative accuracy achieved using OP-LMEM was compared to that obtained using the histogram-mode 3D Ordinary Poisson Ordered Subset Expectation Maximization (3D-OP) algorithm with similar scatter and randoms correction methods. Excellent agreements were achieved and a dual reconstruction scheme was developed based on these results.

4.2 Implementation of the Scatter Correction in List-Mode Reconstruction

In this work, the expected scatter estimate in our OP-LMEM algorithm is based on the Single Scatter Simulation (SSS) technique in which the scatter sinograms are calculated on a frame by frame basis as described in Chapter 3. This approach requires an intermediate histogramming step in the list-mode reconstruction to obtain the scatter and randoms estimates. As mentioned in Section 2.4.2, the temporal subsets in list-mode reconstruction are basically short scans by themselves. Therefore, the scatter and randoms estimates should be ideally calculated for each subset individually as shown in Eq.(2.13):

\[
\lambda_{j+1}^{m,l} = \frac{\lambda_j^{m,l}}{S_j} \sum_{k \in s_i} P_{ij} \frac{1}{FP_{ik}(\lambda_j^{m,l}) + \gamma_{ik} + s_{ik}}
\]  

(2.13)

However, in a single temporal subset the number of counts would be too low to provide a reliable scatter (as discussed in Section 3.4 and will be further demonstrated in Chapter 5) and randoms estimates for dynamic studies. Moreover, the time cost for estimating the scatter and the smoothed random events would be proportional to the number of the list-mode subsets, which is undesirable. As a result, a scatter and a randoms estimate are calculated using the data within each whole dynamic frame.
instead of within each temporal subset. It will be addressed in Chapters 5 and 6 that even the frame-based scatter and randoms estimates are still time consuming and may typically contain a low number of counts, and therefore, a novel scatter calibration and a more practical scatter and randoms approximation techniques are introduced.

In our OP-LMEM algorithm, the outputs from the scatter and smoothed randoms software provided by CPS/Siemens have been modified to be incorporated into the list-mode temporal subsets; i.e. the estimated frame scatter and random events are rescaled amongst the temporal subsets, taking into account a different number of counts within the individual subset, and the dead time and decay correction factors are applied to the image obtained from the subset as well as to the scatter and randoms estimates during the reconstruction. Note that unlike the histogram-mode methods in which data subsets span the same time durations, data subsets in list-mode reconstruction are temporally divided and interleaved as shown in Figure 2.7. For the particular example of the HRRT (Section 1.7.4) scanner, in which scatter and randoms estimates are performed in span 3, the list-mode event coordinates were binned into the closest LOR in span 3 in order to use the corresponding scatter and randoms estimates, and the span 3-binned list-mode events were reconstructed to compare with the histogram-mode reconstruction in span 3.

To validate the accuracy of the implemented scatter correction, one of the options is to use simulation data where the truth is known exactly though the parameters in the simulation need to be adjusted to match the real experimental conditions so that the results obtained from the simulation agree with the experimental ones. Another option to validate the accuracy of the implemented scatter correction is to use the experimental data directly. As long as the experimental results are consistent and reproducible within reasonable expectations, experimental data can be used as validation tools since the methods are tested directly under realistic conditions. The latter option has been taken to validate the proposed methods throughout Chapters 4, 5, and 6. In later sections, all comparisons presented were performed with experimental data and were designed to test the performance of the scatter-corrected list-mode reconstruction (OP-LMEM) in realistic scanning situations based on the practical uses of the reconstructed images (e.g.
to obtain the time activity curves (TACs) and the binding potential (BP) values). The following experiments were performed to validate the scatter-corrected list-mode reconstruction using the image analyses described in Section 2.6.

4.3 Experiments

4.3.1 Methods

The following methods (acronyms) which will be described later in this section were tested and shown in the results section:

a) Histogram-mode reconstruction with the scatter and randoms estimation (3D-OP)

b) List-mode reconstruction with the scatter and randoms estimation (OP-LMEM)

Phantom study: A 20 cm long, 10 cm radius phantom was used. The phantom has three 5 cm diameter cylindrical inserts, one of them was filled with water, another one was Teflon (two 'cold' inserts), and one with an initial $^{11}$C radioactivity concentration of 47 kBq/ml ('hot' insert). The rest of the phantom was filled with an initial $^{11}$C concentration of 11.7 kBq/ml ('background'), giving a hot to background ratio of 4.0. Twenty four dynamic frames of 360s each were reconstructed. The measurement thus covered 7 radioisotope half-lives, and the counts per frame ranged between 200M and 2M with the corresponding count rate between 555 and 5.5 kcps.

Non-human primate study: A non-human primate underwent a 60 min $^{11}$C-dihydrotetrabenazine (DTBZ – a vesicular monoamine transporter VMAT2 marker) scan on the HRRT after a 6 min transmission scan and a 5 mCi bolus injection. Data were acquired in list-mode and then framed into a $5 \times 1$ min, $5 \times 5$ min and $4 \times 7.5$ min framing sequence. The number of counts/frame in this study ranged from 140M to 13M and the count rate ranged between 600 and 50 kcps.

Comparison schemes for validating the scatter-corrected list-mode reconstruction: Dynamic phantom images were reconstructed using (i) 3D-OP and (ii) OP-LMEM both with and without the scatter correction in span 3 and with a maximum ring difference of
67. Table 4.1 outlines the differences between the 3D-OP and OP-LMEM algorithms as described in Chapter 2.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Data compression</th>
<th>Backprojector</th>
<th>Subset scheme</th>
</tr>
</thead>
<tbody>
<tr>
<td>3D-OP</td>
<td>Span 3</td>
<td>Voxel-driven</td>
<td>Spatial angular</td>
</tr>
<tr>
<td>OP-LMEM</td>
<td>Span 3</td>
<td>LOR-driven</td>
<td>Temporal</td>
</tr>
<tr>
<td>OP-LMEM</td>
<td>Pure list-mode</td>
<td>LOR-driven</td>
<td>Temporal</td>
</tr>
</tbody>
</table>

*TABLE 4.1: Differences between the 3D-OP and OP-LMEM algorithms; note that the last row shows the pure list-mode approach without any data compression which has not been considered here*

The following comparisons were performed:

1.1) *Transaxial profile comparisons (with and without the scatter correction):* A line profile was placed along a summed transaxial plane (summed over 100 axial planes). The summation was performed so as to minimize the contribution of statistical errors and emphasize the potential presence of systematic differences. The line traverses the hot spot and one of the cold spots (teflon) as shown in Figure 4.1. The transaxial profiles of all 24 frames were compared between the histogram and list-mode reconstructions with the same number of subsets and iterations (4 iterations with 8 subsets). The transaxial profile comparison was also performed on the images without the scatter correction to ensure the equivalence of the list-mode and histogram-mode reconstruction algorithms before adding the scatter correction. In addition, another profile was arbitrarily placed on a single plane to check the statistical variations between the reconstructions.
1.2) Axial profile comparisons (with the scatter correction): Mean voxel intensities within a ROI placed on each axial plane were plotted as a function of the axial position. The placed ROI covered the entire cross section of the phantom (i.e. including the hot, cold, and background regions). The axial profiles of all frames were compared between the histogram and list-mode reconstructions.

1.3) Contrast, ROI, and voxel noise vs number of iterations comparisons (with the scatter correction): The percent contrast of the cold cylinder as defined in Eq.(2.14) and the coefficient of variation $C_r$ as defined in Eq.(2.16) (for both ROI-level and voxel-level as referred to ROI-noise and voxel-noise later) of the background were calculated for all frames according to the NEMA protocol [48]. Contrast and noise were plotted as a function of number of iterations, and comparison was performed between histogram and list-mode reconstructions both with and without a 2mm 3D Gaussian filter.

1.4) Time activity curve (TAC) comparisons (with the scatter correction): The TAC, with the ROI described in 1.2) which covers the entire cross section of the phantom (i.e. including the hot, cold, and background regions) on a single plane, was plotted over all frames for 3D-OP and OP-LMEM reconstructions for the phantom study as well as for the non-human primate study with the striatal and cerebellum ROIs.

FIGURE 4.1: The transaxial (line) profile across the summed plane (a) without the scatter correction and (b) with the scatter correction.
1.5) Binding potential (BP) comparisons (with the scatter correction): The binding potential for the non-human primate study was computed using the tissue input Logan graphical approach [49] using the cerebellum as reference region for 3D-OP and OP-LMEM reconstructions.

1.6) Reconstruction time cost comparisons (with the scatter correction): The reconstruction time cost was plotted as a function of the number of counts in each frame for both 3D-OP and OP-LMEM reconstructions. In all cases, the same number of computer processors (Dual Xeon 2.8 GHz blade processors with 2.5 GB RAM and gigabit Ethernet networking with jumbo-frames option) was used. The time costs for the histogramming process and estimating the scatter and randoms corrections are excluded in the comparison since they are the same for both 3D-OP and OP-LMEM.

4.3.2 Results

Profile comparisons (1.1 and 1.2) between OP-LMEM and 3D-OP: In all cases, the transaxial and axial profiles obtained with 3D-OP and OP-LMEM virtually overlap for all frames of the phantom study. An example is depicted in Figure 4.2. However, when examining the transaxial profiles on a single plane, it has been observed that there are some minor statistical variations between images reconstructed using 3D-OP and OP-LMEM when the number of counts within the frame is lower than 25M as shown in Figure 4.3; the cause of this variation will be discussed in Section 4.3.3.

FIGURE 4.2: (a) The transaxial or radial profile on a summed plane and (b) the axial profile comparison between 3D-OP and OP-LMEM image reconstruction on a representative frame
FIGURE 4.3: The transaxial profile comparison on a single plane for a frame with ~15M true counts

*Contrast and noise comparisons (1.3) between OP-LMEM and 3D-OP:* The percent contrast, ROI noise, and voxel noise vs number of iterations comparisons for frame 10 are shown in Figure 4.4a to 4.4d, respectively. Frame 10 was arbitrarily chosen here, and other frames all show similar results.
FIGURE 4.4: (a) Contrast vs number of iteration comparison, (b) ROI noise vs number of iteration comparison; there is virtually complete overlap between the results obtained with the two algorithms, (c) Voxel noise vs number of iteration comparison, and (d) Voxel noise vs number of iteration comparison after applying a 2mm 3D Gaussian filter.

Here the contrast also agrees between histogram and list-mode reconstruction. In addition, the ROI noise comparison shows a good agreement between the two algorithms. The voxel noise however shows a slight difference between 3D-OP and OP-LMEM as depicted in Figure 4.4c, which is likely due to the difference between a voxel-driven backprojection (3D-OP) and a LOR-driven backprojection (OP-LMEM): for the voxel-driven backprojection a voxel can get contributions from millions of LORs, whereas for the LOR-driven backprojection a LOR can only contribute to a few voxels as discussed in Section 2.4.2. The difference in voxel noise has also been confirmed not to be contributed from the difference in the subsets by using exactly the same data (i.e. no subsets) in both 3D-OP and OP-LMEM. As shown in Figure 4.4d, after applying a 2mm 3D Gaussian filter, which is commonly used when reconstructing HRRT data, the voxel noise for the 3D-OP reconstruction agrees very closely with that for the OP-LMEM reconstruction thus minimizing the practical difference between the images reconstructed using each algorithm.
Accuracy of TACs comparisons (1.4) between OP-LMEM and 3D-OP: The TAC analysis showed very good agreement between results obtained with the histogram-mode and list-mode reconstructions for the phantom study (i.e. the differences in the ROI values are within ~2% with respect to the 3D-OP ROI value) as depicted in Figure 4.5a. A representative TAC comparison for the non-human primate study (where the counts per frame vary from 140M to 13M) between 3D-OP and OP-LMEM is shown in Figure 4.5b. The TACs agree within ~5%, and the difference is expected to be due to the aforementioned statistical variation between 3D-OP and OP-LMEM (see Figure 4.3).

![Graph](image)

**FIGURE 4.5:** (a) TAC comparison between 3D-OP and OP-LMEM for the phantom study and (b) TAC comparison for the non-human primate study; the TACs for the right caudate (rt-cau) and the left cortex (lf-cortex) regions are shown here

Binding potential (BP) comparisons (1.5) between OP-LMEM and 3D-OP: In all cases examined, we found at most a ~3% difference in the BP estimated when comparing analyses performed on data reconstructed with 3D-OP and OP-LMEM. The quantitative binding potential values of a representative study for each algorithm are shown in Table 4.2.
<table>
<thead>
<tr>
<th></th>
<th>3D-OP</th>
<th>OP-LMEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Right-caudate</td>
<td>2.70</td>
<td>2.77</td>
</tr>
<tr>
<td>Left-caudate</td>
<td>2.93</td>
<td>2.89</td>
</tr>
<tr>
<td>Right-putamen</td>
<td>3.07</td>
<td>3.16</td>
</tr>
<tr>
<td>Left-putamen</td>
<td>3.01</td>
<td>3.01</td>
</tr>
<tr>
<td>Right-ventral striatum</td>
<td>2.42</td>
<td>2.40</td>
</tr>
<tr>
<td>Left-ventral striatum</td>
<td>2.55</td>
<td>2.49</td>
</tr>
</tbody>
</table>

* TABLE 4.2: BP value comparison between 3D-OP and OP-LMEM

*FIGURE 4.6: The reconstruction time cost v.s. the number of counts using the same number of processors for both 3D-OP and OP-LMEM reconstructions

*Reconstruction time cost comparisons (1.6) between OP-LMEM and 3D-OP:* The reconstruction time cost as a function of number of counts is depicted in Figure 4.6. The
intercept between the two curves shown in Figure 4.6 determines the threshold value for the dual reconstruction scheme (which will be discussed in Section 4.4), and in this case a threshold value of 50M counts was obtained; i.e. with our current implementation of the list-mode reconstruction, OP-LMEM is more efficient than 3D-OP when the number of counts within the frame is less than 50M.

4.3.3 Discussion
A very good agreement between 3D-OP and OP-LMEM was obtained for most figures of merit examined; only statistical variations were observed when the number of counts within the frame was lower than 25M. These variations were attributed to the fact that although the two methods use the same overall data, the subsets are not identical (i.e. the difference between the histogram-mode spatial angular subsets and the list-mode temporal subsets; also see Table 4.1): in a situation of low counts, statistical differences in the subsets will produce a different reconstructed image, and up to a 5% difference has been observed in the examined ROI values and at most a 3% in the binding potential values calculated with TACs obtained from images reconstructed with the two algorithms. Considering the scan-to-scan BP estimate variability is approximately 10% for most methods, a 3% difference due to the reconstruction algorithm is considered acceptable.

Since very similar quantitative accuracy between the list-mode and the histogram-mode reconstruction has been achieved, a dual (histogram/list-mode) reconstruction scheme was developed to reduce the time cost for dynamic PET imaging using the HRRT scanner which often has a high number of LORs/number of counts ratio as will be discussed next.

4.4 A Dual (Histogram/List-Mode) Reconstruction Scheme
List-mode reconstruction for high-resolution PET imaging in which an increasing number of lines-of-response (LORs) are possible has been established to be particularly advantageous for studies with a low number of counts; data are reconstructed on an event by event basis thus rendering the reconstruction times proportional to the number of acquired events. In contrast, conventional histogram-mode reconstruction has a fixed
time cost for every frame since data are reconstructed on a line-of-response (LOR) basis. Histogram-mode reconstruction is thus time-wise more efficient for studies with a large number of counts. As a result, especially in dynamic imaging, it would be practically advantageous to have the ability to choose the time-wise optimal reconstruction mode for each frame on the basis of the number of events acquired during that temporal frame. Since the same quantitative accuracy between 3D-OP and OP-LMEM has been achieved based on the results obtained above, a dual (histogram/list-mode) reconstruction scheme can be applied. The basic concept of the dual reconstruction scheme is shown as a flow chart in Figure 4.7.

![Flowchart for dual reconstruction scheme](image)

**FIGURE 4.7:** The flow chart for the dual reconstruction scheme

The applicability of the dual reconstruction scheme depends on the optimization of the histogram and list-mode algorithm, the number of counts within the dynamic frames,
and the corrections applied (i.e. scatter, randoms, and motion corrections) to both reconstruction algorithms. In other words, the dual reconstruction scheme is not applicable when one algorithm is faster than the other for all the dynamic frames as well as when the same or equivalent corrections are not applied to both algorithms.

For the case of our HRRT scanner, there are ~460M LORs in span 3; therefore, ideally if the number of events in a dynamic frame is less than the number of LORs (a threshold value of 460M), then one should use the event-based list-mode reconstruction and vice versa given that both reconstruction algorithms are equally optimized. The intercept between the two curves shown in Figure 4.6 determines the threshold value for the dual reconstruction scheme, but in this case a threshold value of 50M counts was obtained since our OP-LMEM is not optimized as compared to 3D-OP.

The following experiments have been designed to test the efficiency of the dual reconstruction scheme.

4.4.1 Experiments
The same phantom and non-human primate studies as described in Section 4.3.1 were used to check the time gained by applying the dual reconstruction scheme; i.e. frames with more than 50M counts were reconstructed with 3D-OP, and frames with fewer than 50M counts were reconstructed with OP-LMEM. For the case of the phantom study, 7 frames were reconstructed with 3D-OP, and 17 frames were reconstructed with OP-LMEM. For the case of the non-human primate study, 4 frames were reconstructed with 3D-OP, and 10 frames were reconstructed with OP-LMEM. Both studies were also reconstructed with pure 3D-OP and OP-LMEM.

4.4.2 Results
The result of the dual reconstruction scheme applying to the phantom study is shown in Figure 4.8. Here the frame number was sorted in the descending order of the number of counts, and the reconstruction time was plotted as a function of the frame number in order to easily visualize the time gained by applying the dual reconstruction scheme. In the case of the phantom study, the time gained is ~40% as compared to either 3D-OP or OP-LMEM reconstruction as depicted by the shaded area.
For the case of the non-human primate study as shown in Figure 4.9, the time gained by applying the dual reconstruction scheme is ~20% as compared to 3D-OP and ~15% as compared to OP-LMEM. OP-LMEM performs faster in this case since more frames contain a low number of counts.

FIGURE 4.8: The result of the dual reconstruction scheme from the phantom study
4.4.3 Discussion
Recently a further accelerated 3D-OP algorithm which improves the reconstruction speed by an order of magnitude has been developed [56]. Currently, we are collaborating with the group which developed the algorithm to further accelerate our OP-LMEM algorithm, and the new threshold value for the dual reconstruction scheme will be determined once both of the accelerated algorithms have been proven to be equivalent.

4.5 Conclusion
Results from phantom and non-human primate studies look excellent from the above comparisons. In particular the excellent agreement reached between OP-LMEM and 3D-OP indicates that the implementation of this scatter correction method into the list-mode algorithm is appropriate thus reaching the same level of quantification accuracy as
3D-OP. For the current version of our 3D-OP and OP-LMEM algorithms, we have found that, for frames with less than 50M counts, OP-LMEM is more advantageous time-wise, and that indicates the threshold value for the dual reconstruction scheme to be 50M. A time gain of 40% was obtained from the phantom study and a 20% time gain was obtained from the non-human primate study by applying the dual reconstruction scheme as discussed above.
Chapter 5

A Novel Scatter Calibration Technique and Experimental Validations
5.1 Chapter Overview

This chapter describes a calibration technique which improves the quantification of dynamic brain imaging in high resolution PET. As previously mentioned, one of the challenges in dynamic imaging is that the scatter estimate for each dynamic frame might be potentially inaccurate due to the short frame duration and various acquisition count rate (i.e. low number of acquired counts and high RF). To avoid confusions on methods described in Chapters 5 and 6, a summary table on all the proposed methods is given in Section 5.2. In Section 5.3, we first describe how high RFs and low number of counts affect the scatter scaling process and demonstrate the bias in the scatter tail-fitting using the Single Scatter Simulation (SSS) technique [2, 8] for frames with high RFs and low number of counts corresponding to typical dynamic human and non-human primate studies using a phantom study acquired with the HRRT. We then present a new approach to scatter correction scaling: a novel scatter calibration technique which compensates the aforementioned bias together with experimental validations. The experiments performed in this chapter are categorized by (i) phantom, (ii) non-human primate, and (iii) human studies. By using phantom studies where the truth is known and the tracer activity distribution does not change over time, the overestimation bias in the scatter estimate, the increase in the SF due to pulse pile-up, and the noisy axial scatter distribution across the planes can be identified. Phantom studies are also good for demonstrating the improvements achieved by the scatter calibration methods. On the other hand, the non-human primate studies typically contain enough counts in each frame to bypass the bias in the scatter estimate, and the tracer activity distribution changes over time in the non-human primate studies thus providing a good reference to check the validity of the assumptions made for the calibration methods. Once the problems are identified and the assumptions are validated, the calibration method is applied to the human studies, and the results are correlated to the phantom studies.

5.2 Summary Table for the Proposed Methods in Chapters 5 and 6

The following table summarizes the methods described in the Chapters 5 and 6.
Table 5.1: Summary of the proposed methods in Chapters 5 and 6

<table>
<thead>
<tr>
<th>Method</th>
<th>Observations/Assumptions</th>
<th>Descriptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global-level Scatter Calibration (GSC)</td>
<td>None</td>
<td>Calibrates the individual (biased) scatter estimate globally</td>
</tr>
<tr>
<td>Segment-level Scatter Calibration (SSC)</td>
<td>The pattern of segment SF is consistent within the reference frame</td>
<td>Calibrates the individual (biased) scatter estimate in the segment-level</td>
</tr>
<tr>
<td>Global-trues-ratio Plane-level Scatter Calibration (GPSC)</td>
<td>Constant axial scatter distribution as a function of time within the reference frame</td>
<td>Calibrates the individual (biased) scatter estimate in the plane-level using the global trues ratio</td>
</tr>
<tr>
<td>Segmented Plane-level Scatter Calibration (SPSC)</td>
<td>Consistent pattern in segment SF and constant axial scatter distribution as a function of time within the reference frame</td>
<td>Calibrates the individual (biased) scatter estimate in the plane-level using the trues ratio from the segment-level</td>
</tr>
<tr>
<td>Practical Scatter and Randoms Approximation</td>
<td>The spatial scatter and randoms distributions are very smooth and not sensitive to the change in the activity distribution; constant global SF within the summed frame</td>
<td>Calculating the averaged scatter and randoms estimates from a summed frame and scaling those estimates according to the true and random counts in each dynamic frame to obtain the individual frame scatter and randoms estimates</td>
</tr>
</tbody>
</table>

5.3 Effect of High RF and/or with Low Number of Counts in the Scatter Tail-Fitting Process

As described in Section 3.4, the last step of the model-based scatter estimation process is to scale the magnitude of initial calculated scatter distribution to account for scatter originated from outside the field of view (FOV) and to make it commensurate to the measured true coincidences. The scaling process is done by fitting the tails of the estimated scatter distribution to the tails of the measured true coincidences. The true coincidences are obtained by subtracting the measured random events: either the
delayed coincidences or the variance reduced (smoothed) randoms from the measured
prompts with the later producing a less noisy estimate of the trues. Therefore, when the
RF is high and/or the number of acquired counts is low, the estimate of true
coincidences for many sinogram bins or lines-of-response (LORs) may result into
negative numbers especially for high resolution scanners which have hundreds of
millions of LORs thus having fewer counts per LOR. Typically, in order to prevent the
negative scaling in the scatter, a positivity-constraint is applied as described in Section
3.5. However, the positivity-constraint is prone to introducing an overestimation bias
into the scatter estimate which, in turn, leads to an underestimated emission image.
Dynamic scanning with short lived radiotracers is particularly susceptible to such biases
for the early frames [57] which tend to be short and acquired at high count rates such as
the $^{15}$O and $^{11}$C studies. In particular, due to the very short half-life of $^{15}$O, a large
amount of activity is injected initially thus introducing a high RF, and the dynamic
frame duration is also short, which makes the number of counts uncertain to bypass the
bias in the scatter scaling due to the positivity-constraint at a high RF. Another example
of studies which are sensitive to the bias in the scatter estimate is the image derived
input functions for dynamic brain studies where the desired frame duration is between 5
to 10 seconds [58, 59]. The image derived input functions technique is potentially useful
as it eliminates the need to take blood samples from the patients. Figure 5.1 shows a
simple example on how high RF may result in negative counts in the trues sinogram
using a sinogram with 4 bins or LORs.

\[
\begin{array}{ccc}
4 & 4 & - \\
1 & 1 & - \\
3 & 1 & - \\
2 & 2 & - \\
1 & 3 & - \\
-1 & -1 & - \\
\end{array}
\]

**Prompts**
10

**Randoms**
8

**Trues**
2

*FIGURE 5.1: Although the net true counts are 2, there are -2 counts in the trues sinogram in this
case; the positivity-constraint sets the negative values in the plane to some pre-set positive threshold
during the scatter scaling process (overestimation).*
Another effect of low number of counts in the scatter tail-scaling is that the tail of the true or prompts sinogram is very noisy thus making the scaling imprecise. Moreover, the scaling is also artificially variable across the axial planes due to the statistical variation of counts across the planes. In typical dynamic imaging as mentioned in Section 3.5, the scatter events are generally estimated on a frame-by-frame basis. The justification for this approach is that the tracer distribution changes as a function of time, and therefore the scatter distribution changes accordingly. However, given the often highly variable number of events per frame typical of dynamic scanning and variable count rates at which the acquisition is performed (see Figure 3.5), the number of counts in each frame might not always be sufficient to produce an accurate scatter scaling across the planes. The bias due to the high RF and/or with a low number of counts is demonstrated with a phantom study which will be described next.

5.4 Phantom Experiment

5.4.1 Method

A 20 cm long, 20 cm diameter cylindrical phantom was filled uniformly with an initial $^{11}$C radioactivity concentration of 60.9 kBq/ml. Data were grouped into sets of frames with similar numbers of counts for various RFs. The scatter estimates were computed for these frames to evaluate the accuracy of the SF for the (conventional) tail-fitting method with and without the positivity-constraint provided by CPS/Siemens and also to study the increase in the SF values at high count rate or RF due to the pulse pile-up effect (see Section 1.8.2). At high count rate, some of the low energy photons which are typically rejected are summed to have energy acceptable to the energy discrimination window. The net result is an increase in the global SF value due to the pulse pile-up.

The framing scheme is as the following: 3 frames with ~2M prompts, 3 frames with ~5M prompts, 3 frames with ~10M prompts, 3 frames with ~20M prompts, and 1 frame with ~50M prompts (the summed frame which includes all frames with 2M, 5M, and 10M prompts) for RFs of 20%, 40%, 60%, and 75% (i.e. 13 frames for each RF). An additional frame with ~2M prompts and 90% RF was also examined (no other number
of prompts would correspond to a 90% RF in this study). These numbers of counts and RFs were chosen since they are fairly representative of the number of counts encountered in human and non-human primate receptor imaging. The multiple frames with similar number of counts and RF were used to check the reproducibility of the SF values, and no frame data are overlapped with one another except that the ones with ~50M counts are the summed frames of the 2M, 5M, and 10M counts frames. The count rate for these frames was at least 500 times higher than the intrinsic LSO true coincidence rate (intrinsic LSO true coincidence rate is about 100 cps) in order to exclude the effect of the intrinsic LSO background contribution to the SF value. In order to minimize the difference in scatter due to pulse pile-up, the scatter estimated from a frame with 75% RF and 300M prompts was used as the gold standard when comparing the frames with 75% RF with and without the positivity-constraint in the scatter tail-fitting process. The gold standards for the other RFs were also obtained with at least 300M prompts and with similar RFs except for the frame with 90% RF.

5.4.2 Results

![Figure 5.2: The global SF as a function of RF for various numbers of counts using the conventional scatter tail-scaling with and without the positivity-constraint](image)

FIGURE 5.2: The global SF as a function of RF for various numbers of counts using the conventional scatter tail-scaling with and without the positivity-constraint
The SF values are reproducible with the maximum variation to be less than 5% in SF within the 3 scatter estimates obtained from similar RF and number of counts for all the frames in this study. The over-scaling bias caused by the positivity-constraint in the conventional scatter tail-scaling is demonstrated in Figure 5.2. One can also observe the overlap of the SF values for all frames obtained without the positivity-constraint independent of the number of counts within the frame. The increase in the SF at high RF without the constraint is due to the pulse pile-up effect, and the information about the pulse-pile up is embedded in the tail of the measured trues at high count rate thus producing a higher magnitude in the scatter estimate through the tail-fitting process. As expected, the SF value obtained with the positivity-constraint gets closer and closer to the SF value without the constraint as the number of counts increases and as the RF decreases since it is less likely to get negative true counts. The SF values obtained without the constraint also agree well with the gold-standard SF values obtained from more than 300M prompts for the corresponding RFs. The worst case here shows a biased global SF as high as 270% obtained from a frame with 90% RF and 2M counts. As a result, the true emission images are globally underestimated due to the over-subtraction of the scatter (see Figure 5.28 for a human example). In general, frames with a RF higher than 50% and with less than 20M prompts were found to suffer significantly from the over-scaling bias. Note that for $^{11}$C tracers in typical human brain studies, most frames contain less than 50M prompts.

A short summary of the number of prompts required to produce a relatively accurate global SF is shown in Table 5.2. Though the relation between the measured prompts and the random fraction would depend on the size of the object, the size of the phantom used here is similar to that of human heads. Consequently, the results from the phantom study give a good guide line for the human scans. A similar guide line can be obtained for different PET scanners by performing a similar experiment.


<table>
<thead>
<tr>
<th>RF (%)</th>
<th>Prompts</th>
<th>Accuracy in global SF</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>~5M</td>
<td>within 2% to the gold standard</td>
</tr>
<tr>
<td>40</td>
<td>~10M</td>
<td>within 2%</td>
</tr>
<tr>
<td>60</td>
<td>~20M</td>
<td>within 3%</td>
</tr>
<tr>
<td>75</td>
<td>~50M</td>
<td>within 3%</td>
</tr>
</tbody>
</table>

**TABLE 5.2:** The number of prompts required to produce a relatively accurate global SF for various RFs

![Graph showing SF estimated from each segment](image)

**FIGURE 5.3:** The SF estimated from each segment (i.e. segment SF) of the sinogram for the frame with 90% RF and 2M prompts with and without the positivity-constraint in the scatter tail-scaling process
Figure 5.3 shows the SF obtained from each segment of sinograms as defined in Section 1.6 (i.e. segment SF) with and without the positivity-constraint for the frame with 90% RF and 2M prompts. No reference was obtained for this frame due to the lack of counts for the 90% RF. One can observe that the estimated SF is higher for the more oblique segment which contains fewer counts (i.e. it is more likely to get negative counts in the trues thus showing significant bias due to the positivity-constraint). The segment SF obtained without the constraint shows a more reasonable value (mostly under 100%); however, the negative scaling of the scatter estimate is allowed without the constraint as can be seen from the negative SF values. Figure 5.4 shows the segment SF (average over the planes within the same segment) obtained with the constraint as a function of the number of true counts per plane for each segment. As expected the estimated SF increases as the number of counts decreases.
Figure 5.5 shows a similar analysis to Figure 5.3 for a frame with 75% RF and 2M prompts with and without the p-constraint. A reference was obtained from a frame with 75% RF and 300M prompts. The bias due to the p-constraint in the SF values is less severe in this case, and also note that the SF values obtained without the constraint oscillate around the reference SF values.

An example of the noisy scaling and noisy axial scatter distribution across the planes due to the low number of counts is shown in Figure 5.6 where (a) is the segment zero scatter sinogram estimated from a reference frame (300M prompts), and (b) is the segment zero scatter sinogram estimated from the frame with only 2M prompts (the same phantom).
In conclusion, the phantom study demonstrated the overestimation bias and noisy axial scatter distribution across the planes in the scatter estimate from tail-fitting the scatter to the measured trues for frames with a high RF and a low number of counts. The increase in the SF values due to the pulse pile-up was also shown. Additionally, the removal of the positivity-constraint was demonstrated to eliminate the bias in the estimated global SF values. The segment SF values obtained without the constraint however have been shown to introduce a significant amount of variability and were thus found to be not sufficiently accurate. A less noisy estimate of the true coincidences can be obtained from subtracting the smoothed randoms from the prompts. By using the less noisy trues in the scatter tail-fitting process a less biased and less noisy scatter estimate can be obtained though the improvement is not very significant as compared to using the trues obtained from subtracting the raw delayed coincidences from the prompts as will be demonstrated in Figures 5.24, 5.25, and 5.27. A novel scatter calibration technique which addresses simultaneously the noisy axial scatter distribution and the bias problems will be described next.

FIGURE 5.6: The $z-r$ view of the segment zero scatter sinogram which shows all 207 planes (i.e. the vertical axis is along the tomograph axis, the horizontal axis is the radial distance from the tomograph’s centre (see Section 1.6.3), and segment zero refers to the direct planes or sinograms with zero oblique angle) (a) for a frame with 300M counts and (b) for a (biased) frame with 2M counts
5.5 A Novel Scatter Calibration Technique

5.5.1 A Global-level Scatter Calibration (GSC)

5.5.1.1 Overall Concept

The basic idea of this calibration is to have a reference scatter estimated from a frame with a relatively low RF and with a relatively high number of counts so that the estimated scatter is free from the aforementioned bias. For example, in dynamic studies the scatter estimated from a summed frame (with longer duration or data summed over multiple dynamic frames) can be used as the reference. This reference scatter estimate is then used to calibrate the scatter in each dynamic frame within or for some cases outside the reference frame. The calibration is based on the following observation/assumption:

- Both scattered and unscattered true events are proportional to the radioactivity thus making the SF approximately constant as a function of time (as will be referred to as the temporal SF from here on) as long as (i) the anatomical structure of the object does not change over time, (ii) it has a relatively uniform density distribution (e.g. head) and (iii) there is no drastic change in the radioactivity distribution (see Figure 5.7).

This can be stated in another useful way:

- The amount (magnitude) of scatter in each dynamic frame is proportional to the number of true counts in the frame. This is the case at least in the global and segment level as will be discussed shortly.
Consequently, by summing over all the scatter counts in the reference scatter estimate and scaling the sum according to the number of true counts in the frame whose scatter estimate needs to be calibrated, one can obtain the expected total scatter counts in the target or biased frame. As such in this Global-level Scatter Calibration (GSC) technique, the calibration factor is obtained by dividing the expected total scatter counts by the total scatter counts from the biased scatter estimate as shown in Eq. (5.1).

\[
S_{cal}^{LORs in (G)} = F_{cal}^{(G)} S_b^{LORs in (G)} , \text{where } F_{cal}^{(G)} = \left(\frac{c_{sb}}{c_{sr}}\right)^{(G)} \left(\frac{c_{sr}}{c_{sb}}\right)^{(G)} \text{ for GSC (5.1)}
\]

where \( S_{cal} \) is the scatter estimate after calibration, \( S_b \) is the (biased) scatter estimate from the original or target frame, \( F_{cal} \) is the scatter calibration factor used globally here, but as will be discussed later, can also be used at the segment, or even plane level, \( c_{sr} \) is the number of scatter counts in the reference frame, \( c_{sb} \) is the number of (biased) scatter counts in the original target frame, \( c_{tb} \) is the number of true counts in the original
(biased) frame, \( c_{tr} \) is the number of true counts in the reference frame, and \((G)\) represents the quantity at the global level. The trues ratio \((c_{ab}/c_{tr})\) is used to scale the reference scatter counts \( c_{tr} \) to obtain the expected scatter counts, and the calibration factor \( F_{cal} \) is obtained by dividing the calibrated scatter counts by the scatter counts from the biased frame \( c_{ab} \) as mentioned previously. The global SF in the biased frame is thus calibrated to be the same as that in the reference frame. Although one can always set a pre-determined limit for the estimated global SF such as < 100% when performing the scatter tail-fitting to the trues, the GSC method determines the object-dependent global SF from the reference frame, which is more appropriate since the reference frame is a frame with a longer duration in the corresponding study.

Since the RF or the count rate is typically chosen to be lower in the reference frame as compared to the target or biased frames to minimize the bias due to high RF during the high count rate period, the SF assigned to the biased frame is expected to be lower due to less pulse pile-up in the reference frame. To account for the pulse pile-up effect for frames with high count rate, an additional correction term needs to be added to the calibration formula as shown in Eq.(5.2):

\[
S_{f, cal}^{LORs \ in \ (G)} = \left( \frac{c_{usb}}{c_{csh}} \right)^{(G)} S_{cal}^{LORs \ in \ (G)}, \text{ where } c_{csh} = \sum S_{cal}^{LORs \ in \ (G)}
\]

where \( S_{f, cal}^{LORs \ in \ (G)} \) is the final calibrated scatter estimate, \( c_{usb} \) is the total scatter counts from the (unbiased) scatter estimate obtained from performing the tail-fitting to the trues without the positivity-constraint for the target or biased frame, \( c_{csh} \) is the total scatter counts from the original calibrated scatter estimate for the biased frame; it is obtained by summing the counts over all LORs in the original calibrated scatter sinogram as depicted in Eq.(5.2). Although the scatter estimate without the positivity-constraint allows the negative scaling and thus is likely to contain negative scatter values, the global SF is unbiased and includes the pulse pile-up information. The additional correction term globally compensates for the increase in SF due to pulse pile-up. One may notice that the reference scatter estimate is not necessarily needed in the GSC.

\[108\]
method when the unbiased total scatter counts are known as one can directly scale the biased scatter estimate according to the ratio of the unbiased total scatter counts and the biased total scatter counts. Consequently, no assumption is needed in the GSC method. However, the reference scatter estimate is essential when calibrating in the segment and plane-level as will be described in the later sections.

5.5.1.2 Experimental phantom results for GSC

The same phantom study as described previously was used to evaluate the GSC method. The reference scatter estimate shown in Figure 5.5 was used to calibrate the frame with 75% RF and 2M prompts. The segment SFs were calculated from the scatter estimates obtained using the conventional method with the positivity-constraint and the GSC method. The segment SFs were also calculated for the reference frame as depicted in Figure 5.8. One can observe that even though the GSC method corrects the global magnitude of the scatter, the bias pattern in the segment SF is still maintained.

![Graph showing segment SFs](image)

**FIGURE 5.8:** Segment SFs obtained from the frame with 75% RF and 2M prompts using the conventional, GSC methods, and that obtained from the reference frame is also shown
The $z-r$ view of the segment zero scatter sinograms obtained from the aforementioned biased frame using the conventional and the GSC methods are shown in Figure 5.9. As expected, the noisy axial scatter distribution across the planes is still present since the GSC method is just a global compensation for the scatter estimate.

![Image](image.png)

**FIGURE 5.9:** The $z-r$ view of the segment zero scatter sinogram obtained from the frame with 75% RF and 2M prompts using (a) the conventional with p-constraint and (b) the GSC method

In conclusion, the global scaling is not sufficient to compensate for the bias in the segment and plane level.

### 5.5.2 A Segment-level Scatter Calibration (SSC)

#### 5.5.2.1 Description of Technique

This calibration can be performed, as shown above, at the global level (i.e. a single correction factor) thus preserving the original calculated frame or global spatial scatter distribution and corrects for the overall scaling. It can also be applied to each segment of the scatter sinogram since the segment SF is also approximately constant across the dynamic scan as depicted in Figure 5.10.
FIGURE 5.10: The estimated segment SF values for 5 representative segments as a function of time for a typical dynamic brain study.

The segment-level calibration preserves the original calculated spatial scatter distribution at the segment level and corrects the scaling for each segment. This Segment-level Scatter Calibration (SSC) method can be written as:

\[
S_{f,\text{cal}}^{\text{LOR}}(S) = \left( \frac{c_{sb}}{c_{cb}} \right)^{(G)} F_{cal}^{(S)} S_b^{\text{LOR}}(S), \text{where } F_{cal}^{(S)} = \left( \frac{c_{sb}}{c_{sr}} \right)^{(S)} \left( \frac{c_{sr}}{c_{sb}} \right)^{(S)} \text{ for } SSC
\]

where (S) represents the quantity in the segment level; also note that the scatter counts in the LORs within each segment have a separate calibration factor.

5.5.2.2 Experimental phantom results for SSC

The same phantom study as described previously was used to evaluate the SSC method. As shown in Figure 5.11, a much better agreement in the segment SF between the SSC method and the reference was achieved. Although the SSC method does not modify the noisy axial scatter distribution across the planes, it does not assume anything about the
tracer or scatter distribution, and it can therefore be applied to all frames independent of the early and the later stages of the scan. Note that the SF should not be confused with the spatial scatter distribution. The experimental results of applying the SSC method on a representative human study will be presented in Section 5.5.6. Next, the plane level scatter calibration will be discussed.

![Figure 5.11: Segment SF obtained from the frame with 75% RF and 2M prompts using the SSC, GSC methods, and that obtained from the reference frame is also shown.](image)

5.5.3 Applying the Scatter Calibration in the Plane Level

According to the results obtained previously, the calibration needs to be applied on a plane-by-plane basis in order to minimize the noise in the axial scatter distribution across the plane. This approach preserves the original calculated plane or transaxial spatial scatter distribution and corrects the scaling for each plane. However, the difficulty for this technique to be applied on a plane-by-plane basis is to obtain a reasonable ratio of the trues (i.e. $c_d/c_r$) on a single plane level since the total number of true counts for each plane in a frame with a low number of counts may be statistically too uncertain to obtain a reliable ratio. Moreover, the scaling according to the trues ratio
at the plane level can only be performed when the temporal SF estimated on a plane level is constant which is not quite true since the spatial scatter distribution is very smooth and is not very sensitive to the change in the tracer distribution. As a result, dividing the slowly-changing scatter by the temporally fast-varying trues for each plane (i.e. temporal plane SF) will not be very constant temporally. In other words, the temporal plane SF is more sensitive to the change in the tracer distribution. This will be demonstrated in Section 5.5.3.1 using a non-human primate study. A solution to make the calibration work at the plane level is either (i) using the global trues ratio (GPSC: Eq. (5.4)) or (ii) using the trues ratio obtained from each segment (SPSC: Eq.(5.5)) since the temporal global and segment SF is much more constant or not sensitive to the change in the tracer distribution in the dynamic scans. An additional assumption that the axial scatter distribution is smooth and approximately constant within the reference frame is thus required as the \( \left( \frac{c_{tr}}{c_{sb}} \right)^{(P)} \) term replaces the noisy axial scatter distribution across the planes with the smoother one from the reference. Because of this assumption, the plane level scatter calibration needs therefore to make use of the relatively constant tracer distribution thus giving a constant scatter distribution during the later stage and is not suitable to calibrate the frames within the early stage of the scan where the tracer and the corresponding scatter distributions change rapidly.

\[
S_{f, \text{cal}}^{\text{LORS in (P)}} = \left( \frac{c_{sb}}{c_{sb}} \right)^{(G)} F_{\text{cal}}^{(P)} S_{b}^{\text{LORS in (P)}}, \text{ where } F_{\text{cal}}^{(P)} = \left( \frac{c_{tb}}{c_{tr}} \right)^{(G)} \left( \frac{c_{sr}}{c_{sb}} \right)^{(P)} \text{ for GPSC}
\]

(5.4)

\[
S_{f, \text{cal}}^{\text{LORS in (P)}} = \left( \frac{c_{sb}}{c_{sb}} \right)^{(G)} F_{\text{cal}}^{(P)} S_{b}^{\text{LORS in (P)}}, \text{ where } F_{\text{cal}}^{(P)} = \left( \frac{c_{tb}}{c_{tr}} \right)^{(S)} \left( \frac{c_{sr}}{c_{sb}} \right)^{(P)} \text{ for SPSC}
\]

(5.5)

where \( (P) \) represents the quantity in the plane level, \( (G) \) represents the quantity in the global level, and \( (S) \) represents the quantity in the segment level.
The choice of using the trues ratio obtained from the segment-level or global-level may seem arbitrary; however, they were chosen as the candidates since in both levels the whole object is covered by the projection data thus giving a more consistent temporal SF, whereas using the trues ratio from the other combinations of the planes may likely (i) not have enough counts and (ii) not have a constant temporal SF within the summed planes. However, a potential drawback of scaling the scatter according to the global trues ratio is the fact that it preserves the biased pattern in the segment SF as will be demonstrated in the non-human primate study (see Figure 5.15). Consequently, scaling according to the trues ratio obtained from the segment level was determined to be the more accurate method (SPSC). A non-human primate study which was designed to demonstrate that the SPSC is more accurate than the GPSC method will be discussed next.

5.5.3.1 Non-human Primate Experiment

5.5.3.1.1 Methods

A $^{11}$C–PIB non-human primate study was arbitrarily chosen for this test. The entire scan was histogrammed into two frames (first half and last half) in order to obtain frames with enough counts in each plane as well as different tracer distributions between the frames. Each frame contains ~300M counts, and the scatter was calculated for these two frames as well as for the entire scan (three data sets). The plane SFs (i.e. scatter fraction obtained from data within each plane) were then calculated for all three data sets, and the ratios of the plane SFs between different data sets were computed to check the difference (temporally) in between the data sets. The first frame was also calibrated with GPSC and SPSC using the scatter estimated from the entire scan as the reference, and the ratios of the plane SFs were calculated using the original plane SFs obtained from the first frame in the denominator. The segment SFs were also computed and compared between the conventional, GPSC, and SPSC methods.

5.5.3.1.2 Results

Figure 5.12 shows the plane SF ratio comparison between the first half, last half, and the entire scan of the non-human primate study. One important observation is that for different tracer distributions (e.g. between the first half and last half of the scan) the
plane SF is not very constant temporally, and the variation gets smaller when comparing the average tracer distribution (i.e. obtained from the entire scan) to the distribution in each frame. A ~20% variation in temporal plane SF ratio was observed in between the first and last half of the scan, and a ~10% variation was observed in between the entire scan and each individual frame (note that it is the ‘percent difference’ of the SF; not the absolute difference in percentage. i.e. if the global SF is 40%, a 20% difference in SF ratio would translate to a 8% absolute difference in the global SF which is the difference one would observe in the true emission image). Consequently, if one wants to perform the calibration by scaling the scatter in each plane according to the plane trues ratio (i.e. the PPSC method as defined in Eq.(5.6)) between the individual frame and the average (i.e. reference frame), similar variation (i.e. ~10% as shown in Figure 5.12) in the temporal plane SF will be observed.

\[
S_{f,\text{cal}}^{\text{LORs in } (P)} = \left( \frac{c_{ab}}{c_{cb}} \right)^{(G)} F_{\text{cal}}^{(P)} S_{b}^{\text{LORs in } (P)}, \text{where } F_{\text{cal}}^{(P)} = \left( \frac{c_{ib}}{c_{ir}} \right)^{(P)} \left( \frac{c_{sr}}{c_{sb}} \right)^{(P)} \text{ for PPSC}
\]

(5.6)

![Figure 5.12: The plane SF ratio between the different (temporal) frames](image)
Figure 5.13 shows the plane SF ratio between the frame calibrated with GPSC/SPSC and the original frame. The red curve shows the difference between the entire scan and the first frame; this is equivalent to calibrating using the true ratio in the plane level (PPSC) when there are enough counts. A very good agreement in temporal plane SF was achieved by both GPSC and SPSC since the true ratio in the calibration is obtained from the entire frame and each segment respectively (due to the fact that the temporal global and segment SF is much more constant thus making the scaling more valid; see Figures 5.7 and 5.10).

\[ \text{FIGURE 5.13: The plane SF ratio between the calibrated and original target frame} \]
Figure 5.14 shows the segment SF comparison between the conventional, SPSC, and GPSC methods for the first half of the non-human primate scan which contains ~300M counts. Both GPSC and SPSC agree very well with the conventional when there are a lot of counts and the RF is not very high (~45%).
FIGURE 5.15: Segment SF comparison between SPSC and GPSC for a frame with 94% RF and 2.6M prompts

However, when comparing using a frame which contains ~2.6M counts and with a RF of 94% from a human study, SPSC shows a more consistent SF for all the segments than GPSC as shown in Figure 5.15. Note the only difference between GPSC and SPSC is the way we obtain the trues ratio (i.e. GPSC uses the same trues ratio for all the segments, whereas SPSC obtains the trues ratio from each segment). Therefore, GPSC only ensures the global SF in the calibrated frame is as consistent as the reference, but the variation in the segment SF stays similarly to that in the original biased frame (also see Figure 5.3 for the biased pattern in the segment SF for the phantom study), whereas SPSC ensures both the global and segment SF in the calibrated frame are as consistent as those in the reference while both of them preserve the same temporal plane SF as the original. As a result, SPSC is the more robust method for the dynamic scatter calibration.

As described previously, the advantage of the SPSC method over the SSC method is that it can improve the noisy axial scatter distribution across the planes with an additional assumption that the axial scatter distribution is constant within the reference
frame. Two outstanding questions are thus: (i) When is this assumption valid, and (ii) what difference does SPSC make in the emission image by having a smoother axial scatter distribution? The next section will address the first question using a non-human primate study.

5.5.4 Validation of Constant Axial Scatter Distribution: Early and Later Stages of the Scan

It is expected that the tracer distribution varies very rapidly during the early stage, and the scatter distribution can therefore be quite different, whereas the tracer distribution changes relatively slowly during the later stage. As a result, the assumption of a constant axial scatter distribution for the SPSC method is expected to be not valid during the early stage but valid during the later stage. In this section, a dynamic non-human primate study with sufficient number of counts for all frames thus having the scatter estimates free from the overestimation bias and the noisy estimated axial scatter distribution is used. Those scatter estimates can thus be used as the gold-standard when performing comparisons. The SPSC method is then applied and compared to the gold-standard.

5.5.4.1 Methods

A non-human primate underwent a 60 min $^{11}$C-dihydrotetabenazine (DTBZ – a vesicular monoamine transporter VMAT2 marker) scan on the HRRT (after a 6 min transmission scan and a 5 mCi bolus injection). Data were acquired in list-mode and then framed into a $5 \times 1$ min, $5 \times 5$ min and $4 \times 7.5$ min framing sequence. The number of counts/frame in this study ranged from 140M to 13M with the RF ranging from 10% to 40%, and the count rate ranged between 600 and 50 kcps. The minimum number of counts per frame in this study (as in other non-human primate studies) was higher than the typical minimum number of counts in human studies, and the size of the non-human primate is smaller than that of humans thus containing more prompts per plane and being less susceptible to the bias in the scatter scaling. Consequently, the conventional method was able to produce the scatter estimate free from the bias caused by low number of counts and high RF. The conventional method was thus used as the gold-standard when evaluating the performance of the calibration methods.
FIGURE 5.16: The global trues rate vs time plot or the global TAC; the frames within the red circle are grouped to form the reference frame; the SSC and SPSC methods are applied to the scatter estimate in the target frame, and the differences in the emission image between the gold-standard and the images reconstructed with the calibrated scatter estimates are compared.

Approach for the early stage of the scan: The early and later stages of the scan can be identified using the dead time and decay corrected global trues rate versus time plot which will be referred to as the global TAC from here on. The frames with similar trues rate adjacent to the first frame in the early stage were grouped to form the reference frame for the calibration as shown in Figure 5.16. The first few frames during the early stage in dynamic brain studies typically contain a low number of counts and a high RF; thus, including these frames in the reference only increases the overall RF. As the result, although it is not the case for the non-human primate study used here, the first frame during the early stage of the scan was not included in the reference formation in order to simulate the typical condition. The SSC and SPSC methods were applied to the first frame in the early stage (target) using the aforementioned reference frame as depicted in Figure 5.16. The emission image of the target frame was reconstructed using 3D-OP with the gold-standard, SSC, and SPSC scatter estimates. The difference in the emission image was investigated by dividing the ones obtained from the scatter calibration to the gold-standard image.
Approach for the later stage of the scan: The first frame during the later stage was used as the target frame since it is the frame with the most different tracer distribution as compared to the rest of the frames in the later stage of the scan. Here, one would like to study how constant the axial scatter distribution is between the target frame and the reference frame and also considering different reference formation scheme. Therefore, the SPSC method was applied to the target frame using reference frames with the grouping scheme shown in Figure 5.17. The emission images were then reconstructed with those calibrated scatter estimates as well as the gold-standard using 3D-OP. The difference in the axial distribution was investigated between the gold-standard and those calibrated with different reference frames as well as the ratio difference in the emission image. An axial profile was examined with the ROI covering the object, and the percent difference in the axial profile was plotted for the various reference frames. This comparison shows the impact of the calibrated scatter estimates which have the same axial scatter distribution as the reference scatter estimates on the emission images for the later stage of the scan. Both of the above comparisons demonstrate if the calibration methods introduce any significant errors in the reconstructed images when there are enough counts to bypass the bias in the scatter estimates.
FIGURE 5.17: The global true rate vs time plot or the global TAC; the frames within the red circles are grouped to form the reference frames, and the difference in the axial distribution is investigated between the gold-standard and the images reconstructed with the calibrated scatter estimates.

5.5.4.2 Results

Early stage of the scan: The first row of Figure 5.18 shows the gold-standard emission image of the non-human primate for the first frame in the early stage of the scan. The second row depicts the ratio of the emission image reconstructed with the SSC method to the gold-standard. The colour scale displays the difference within 10% of the gold-standard. The image obtained from the SSC method agrees very well with the gold standard inside the object as shown by the ratio value of 1's. The third row shows the ratio of the image reconstructed with the SPSC method to the gold-standard. One can observe differences higher than 10% inside the object. Note that the higher differences were found around the relatively cold regions (i.e. regions with relatively low activity) in the emission image since they are more sensitive to the change in the scatter estimate. Nevertheless, the image obtained from the SSC method shows a very good agreement even in those relatively cold regions thus demonstrating that the SSC method is suitable.
for the early stage of the scan since it does not assume anything about the tracer or scatter distribution.

![Diagram](image)

**FIGURE 5.18:** The first row shows the gold-standard image for the first frame in the early stage of the scan; the second row shows the ratio of the SSC image to the gold-standard; the third row shows the ratio of the SPSC image to the gold-standard; the transaxial, coronal, and sagital views are shown from left to right respectively

The SPSC method on the other hand assumes a constant axial scatter distribution between the target and the reference frame thus showing a higher difference in the ratio comparison. This also indirectly demonstrates that the axial scatter distribution is not very constant between the target and the reference frame in the early stage of the scan.
In addition, another reference frame obtained from the entire scan was used in the SSC method, and an almost identical result was obtained thus showing that the reference frame formation for the SSC method is very flexible.

Later stage of the scan: The percent difference in the mean axial distribution between the images reconstructed with the calibrated scatter estimates and the gold-standard is depicted in Figure 5.19; only the axial planes which contain the object are shown. As expected, the SPSC image using the longest reference frame shows the highest difference, whereas the one using the shortest reference frame shows the lowest. However, even in the worst case scenario when using the entire later stage as the reference frame, the differences of less than 5% are only noticeable near the end planes. This comparison demonstrates that the assumption of a constant axial scatter distribution is reasonably valid during the later stage of the scan.

**FIGURE 5.19:** The percent difference in the axial profile between the images reconstructed with the SPSC calibrated scatter estimates and the gold-standard using the reference frames defined in Figure 5.17; the 3300s reference frame covers the entire later stage of the scan
The image ratio comparison between the SPSC image using the entire later stage of the scan as the reference and the gold-standard is shown in Figure 5.20. The first row depicts the gold-standard image, and the second row shows the image ratio. Again, higher differences were found at the relatively cold regions near the end planes.

**FIGURE 5.20:** The first row shows the gold-standard image for the first frame in the later stage of the scan; the second row shows the ratio of the SPSC image using the entire later stage of the tracer as the reference to the gold-standard; the transaxial, coronal, and sagital views are shown from left to right respectively; the axial boundaries (plane 80 to 200) as shown in Figure 5.19 are depicted by the two line profiles in the coronal and sagital views in the second row.

All in all, the SSC method has been demonstrated to be robust even in the early stage of the scan while the SPSC method is only suitable during the later stage. Next, the improvement achieved by the SPSC method will be presented using a phantom study.
5.5.5 Experimental Phantom Results for the SPSC Method

5.5.5.1 Methods

The same uniform phantom study as described in Section 5.4 was used. The frames with 40% RF were chosen to simulate the frames during the typical later stage of the scan. Since the SPSC method modifies the axial scatter distribution, the improvements are expected to be in the axial uniformity and the voxel noise in the phantom images. The following comparisons are performed:

1.1) **Scatter sinogram comparisons:** The scatter sinograms obtained from the frame with 40% RF and 3M prompts using the conventional, SSC, and SPSC methods and that obtained from the reference frame were compared visually to check the axial scatter distribution across the planes. A reference frame with 40% RF and 400M prompts was used.

1.2) **Axial uniformity comparison:** The emission image for the frame described in comparison 1.1 was reconstructed with the conventional, SSC, and SPSC scatter estimates. The axial uniformity of the phantom images was examined by using the axial profile with a ROI covering the whole cross section of the phantom. The emission image of the reference frame was also reconstructed, and its axial profile was used as the gold-standard. The ratio comparison was performed by dividing the axial profiles obtained from the images using the conventional, SSC, and SPSC scatter estimates by the gold-standard. In this case, the method which shows the more constant ratio across the axial direction achieves better axial uniformity. The standard deviation of the ratio over the mean ratio was computed for each method to demonstrate the variability of the axial profile.

1.3) **Voxel noise comparison:** The coefficient of variation in the voxel-level as defined in Section 2.6.2 was computed for a thousand voxels arbitrarily chosen from ten axial planes for the emission images reconstructed with the conventional, SSC, and SPSC scatter estimates. The number of prompts ranges from 3M to 50M. The voxel noise was plotted as a function of the number of prompts for each method.
5.5.5.2 Results

Comparisons 1.1: The scatter sinogram comparison is shown in Figure 5.21. One can observe the smooth axial scatter distribution across the plane in the reference, while the distribution is noisy and biased in the conventional scatter estimate. The SSC method corrects the global magnitude and the biased pattern in the segment SF, and the SPSC method produces a smoother axial scatter distribution closer to the reference.

![Figure 5.21](image)

*FIGURE 5.21*: The z-r view of the segment zero scatter sinogram which shows all 207 planes (a) for the reference with 400M prompts and 40% RF, (b) for the frame with 3M prompts and 40% RF, (c) after calibrating with SSC and (d) SPSC

Comparisons 1.2: The axial uniformity comparison is depicted in Table 5.3. The standard deviation over the mean gives a measure of the relative variation from the mean. The SSC method compensates for the bias and also produces a lower variation of the ratio from the mean or a better axial uniformity. In addition, the SPSC method demonstrates a slightly better uniformity as compared to the SSC method as can be observed from the lower standard deviation over the mean.
Comparisons 1.3: The voxel noise comparison is shown in the Figure 5.22. One can observe that the SSC reduces the noise by compensating the overestimation bias in the scatter, and the SPSC method produces a slightly less noisy image as compared to the SSC method. As expected, the difference in the voxel noise between these three methods decreases as the number of prompts increases since the scatter estimates reach minimal difference at high number of counts.

The advantages of the SPSC method over the SSC are demonstrated above. Next, an alternative simplified method which shows similar results as the SPSC method will be discussed.
5.5.5.3 SPSC Method Leading to a Practical Scatter Approximation

As described previously, the primary assumption of a constant axial scatter distribution used in the SPSC method limits the use of the method in the later stage of the scan. However, since the tracer distribution changes relatively slowly in the later stage, one can also assume that the 3-dimensional scatter distribution is fairly constant since the scatter distribution is smooth with no high frequency components and is not very sensitive to the change in the tracer distribution. Furthermore, the temporal global SF value is also fairly constant as shown in Figure 5.7. Therefore, one can simply take the averaged scatter distribution obtained from a summed frame and scale it according to the ratio of the true counts between the individual dynamic frame and the summed frame. For example, the reference scatter estimate shown in Figure 5.21a can be directly used in the reconstruction of the target frame after the scaling. This method will be referred to as a practical scatter approximation and will be described in greater detail in Chapter 6. This approximate method not only improves the efficiency of the image formation task as it is not required to compute the scatter estimate for every single dynamic frame but also shows similar results as the SPSC method as will be briefly demonstrated below.

<table>
<thead>
<tr>
<th>Method:</th>
<th>SPSC</th>
<th>Approximate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Std/mean:</td>
<td>0.04781</td>
<td>0.04792</td>
</tr>
</tbody>
</table>

*TABLE 5.4: The standard deviation over the mean of the axial profile ratio comparison*
FIGURE 5.23: The voxel noise comparison between the phantom images reconstructed with the SPSC and the approximate scatter estimates

The same analyses as described in Section 5.5.5.1 were done for the approximate method using the phantom study. As shown in Table 5.4 and 5.23, the approximate method shows very similar results to the SPSC method. Since the approximate method is much more efficient than the SPSC method as it only requires the scatter estimate from the summed frame while still maintaining similar results as the SPSC method, the approximate method is chosen to be the candidate for the frames when the scatter distribution varies minimally as will be discussed in Chapter 6. For the early stage of the scan, the SSC method should be applied, and the results on a representative human study will be presented in the next section.

5.5.6 Experimental Human Results for the SSC Method

5.5.6.1 Methods

In order to exclude the effect of patient motion without performing any motion compensation, human studies with minimal amount of motion (~1-2 mm) were selected for the following validations.
A human subject underwent a 60 min $^{11}$C- raclopride (RAC – a D$_2$ receptor marker) scan on the HRRT (after a 6 min transmission scan and a 10 mCi Harvard pump injection). Data were acquired in list-mode and then framed into a 4 × 1 min, 3 × 2 min, 8 × 5 min, and 1 × 10 min framing sequence. The number of counts/frame in this study ranged from 57M to 2.6M with the RF ranging from 10% to 94%, and the count rate ranged between 370 and 25 kcps. The following comparisons were performed:

2.1) Scatter sinogram comparisons: The first frame of this study which contains a 94% RF and 2.6M counts was used to examine the effect of the SSC method. The segment zero scatter sinograms and the axial scatter profiles which were placed at the centre of the scatter sinograms as shown in Figure 5.26 for the conventional method with and without the positivity-constraint were compared with those obtained by the SSC method. For this comparison, an addition case, in which the conventional method was performed using the measured true obtained from subtracting the variance reduced randoms from the prompts instead of the raw delayed coincidences in the scatter tail-fitting process, was examined. The bias in the scatter estimate and the noisy axial scatter distribution are expected to be improved when the variance reduced randoms estimate is used.

2.2) Segment SF comparison: The segment SF as a function of the number of true counts per plane was plotted for a frame with 90% RF and 2M prompts obtained from a phantom study as well as for the frame obtained from the human study as described in the above comparison (2.1). The Segment SF values were also compared between all the scatter estimates described in comparison (2.1).

2.3) Emission image comparison: The fully-corrected emission images were reconstructed with the conventional and the SSC methods using 3D-OP. The visual comparison between the two methods is shown for one of the frames which suffered from the biased scatter scaling.

5.5.6.2 Results

Comparisons 2.1 and 2.2: Figure 5.24 shows the z-r view of the segment zero scatter sinogram for the first frame of the human study which contains a RF of 94% and 2.6M counts. The sinograms obtained with both conventional methods (using prompts –
delayed coincidences with and without the positivity-constraint) show very noisy axial scatter distribution across the planes. As expected, the prompts – variance reduced (smoothed) randoms approach shows lower bias and less noisy axial scatter distribution since the tail of the trues is less noisy than the one obtained from subtracting the delayed coincidences from the prompts. The sinogram calibrated with the SSC method shows a less noisy distribution across the planes with a global SF of ~47%. The axial scatter profile is shown in Figure 5.25. The overestimation bias can be easily observed for the conventional methods with the positivity-constraint as the estimated global SF is ~350% for the prompts – delayed coincidences approach and ~280% for the prompts – smoothed randoms approach; clearly there are very few counts left in the emission image after subtracting too much scatter, which agrees with our findings from the phantom study. The segment SF as a function of the number of true counts per plane also shows good agreement between the phantom and human study with similar RF and number of prompts as shown in Figure 5.26. Moreover, even though the global SF for the conventional methods without the constraint are ~47% which is much more accurate than the one with the constraint, the negative and noisy scaling still makes it suboptimal as demonstrated in Figures 5.24b, d, and 5.25.

**FIGURE 5.24:** The z-r view of the segment zero scatter sinogram for the first dynamic frame of the human study (a) using (prompts – delayed coincidences) with the positivity-constraint, (b) using (prompts – delayed coincidences) without the constraint, (c) using (prompts – smoothed randoms) with the positivity-constraint, (b) using (prompts – smoothed randoms) without the constraint, (e) calibrated with SSC, and (f) the reference used for the calibration.
FIGURE 5.25: The profiles across the planes defined in Figure 5.24 (a), (b), (c), (d), and (e) together with the corresponding global scatter fraction (SF)
FIGURE 5.26: Comparison of segment SF as a function of the number of true counts per plane within each segment using the conventional scatter estimate between the phantom and human study.

The segment SF comparison between all the scatter estimates described in Figure 5.25 is depicted in Figure 5.27. The two methods which apply the p-constraint show the biased pattern in the segment SF, whereas the segment SF values obtained from the two methods without the p-constraint oscillate around the SSC calibrated segment SF values.
Comparison 2.3: Figure 5.28 shows the reconstructed emission image for a frame which suffers severely from the biased scatter estimate (with a corresponding estimated global SF of 99%) obtained from the conventional method with the p-constraint (middle 3 images). One can clearly observe the 'missing regions' as indicated by the arrows due to the biased scatter when comparing to the same image after the calibration (bottom 3 images). The image reconstructed without any scatter correction is also shown (top 3 images). This frame is shown instead of the one depicted in Figures 5.24 and 5.25 because of the relatively better anatomical layout of the image.
5.5.7 Applying SSC to Frames with a Very Short Duration

As mentioned in Section 5.3, applications (such as the image derived input function method) which require very short frame durations for the early stage of the scan are very sensitive to the bias in the conventional scatter estimation. In this section, the improvement which can be achieved by the SSC method on these short frames is investigated.

*FIGURE 5.28:* An example which demonstrates the effect of the biased scatter on the (true) emission image (middle 3), and the same image after applying the calibration (bottom 3). The image reconstructed without any scatter correction is also shown (top 3). The transaxial, coronal, and sagittal views are shown from left to right respectively.
The same human study as described in Section 5.5.6 was used. The first two minutes of the scan was divided into $12 \times 10$ seconds framing sequence. The number of prompts ranges from $10k$ to $4M$, and the RF ranges from 50 to 95% for these short frames. The emission images were reconstructed with the conventional, the SSC methods, and also without any scatter correction using 3D-OP. A representative TAC with a ROI covering most of the activity distribution in the middle planes of the subject is shown in Figure 5.29.

\[ \text{FIGURE 5.29: A representative TAC comparison between the conventional, the SSC methods, and without any scatter correction (nosc).} \]

As expected, the overestimation bias in the conventional scatter estimate occurs in all the frames; in particular, there is nothing much left in the emission image for the first 70 seconds. As a result, the TAC obtained from the conventional method is very inaccurate. After assigning a more accurate scatter fraction in each segment through the SSC method, a more accurate ROI value which leads to a more quantitative determination of the activity concentration can be obtained. Consequently, a significant improvement in the TAC is achieved using the SSC method in this case.
5.6 Discussions and Conclusion

A scatter calibration technique was developed to reduce the noisy axial scatter distribution across the planes and to compensate for the over-scaling bias in the scatter estimation for scans with high RF and/or low number of counts, a situation often encountered in dynamic imaging on scanners with a large number of LORs such as the HRRT. For example in the phantom study, an estimated SF of as high as 270% was obtained using the conventional method from a frame with 90% RF and 2M counts (similar results were obtained from human studies with similar RF and counts as well), whereas the unbiased SF was ~70%. A much more consistent global SF was obtained by GSC, SSC, and SPSC as compared to the conventional method, and a smoother scatter estimation across the planes was achieved by applying SPSC. It has been identified that the SSC method can be applied independent of the early and later stages of the scan, while the SPSC method should be only applied during the later stage. However, since the tracer and the corresponding scatter distribution varies relatively slowly in the later stage of the scan, a practical scatter approximation can be applied instead of the SPSC method as will be described in Chapter 6. The calibration technique was also applied to human studies and shown both more qualitative and quantitative results. The proposed method is thus a solution to the overestimation bias described in [57], and it allows one to have more freedom on making framing and scanning protocols when the bias in the scatter at low number of counts is a major concern. In addition, the proposed scatter calibration method is particularly beneficial for low count $^{15}$O studies and for the image derived input functions for dynamic brain studies where one is interested in having short frames with only 5 to 10 seconds image duration [58, 59].
Chapter 6

A Practical Scatter and Randoms Approximation Technique and Experimental Validations
6.1 Chapter Overview

In this chapter, an optimization technique which further accelerates the image formation task is described, and it has been published in Physics in Medicine and Biology along with the work described in Chapter 4 [3]. As previously mentioned, one potential challenge in dynamic PET image reconstruction is that the reconstruction time depends heavily on the number of dynamic frames (e.g. frame-based scatter and smoothed randoms estimations). In this work, we present a practical approximate scatter and smoothed randoms estimation approach for dynamic PET studies based on a time averaged scatter and randoms estimate followed by scaling according to the global numbers of true coincidences and randoms for each temporal frame. The quantitative accuracy and efficiency (i.e. reconstruction time and data storage gain) of this method will be discussed with experimental validations. Finally, the practical approximation is combined with the dual reconstruction scheme to optimize the efficiency of the dynamic image reconstruction.

6.2 Introduction

As previously described in Chapter 5, a scatter calibration technique was developed to improve the quantitative accuracy of the dynamic images. It is based on using a reference scatter estimate obtained from a reference frame which is typically formed by summing over multiple dynamic frames thus containing more counts and relatively lower RF to minimize the overestimation bias and noisy axial scatter distribution in the scatter tail-fitting process as introduced in Chapter 5. The reference scatter estimate is then used to calibrate the scatter estimate in the target frames to improve the quantitative accuracy. In this Chapter, we investigate a different approach named the practical scatter approximation, which directly applies the scatter estimate from the summed frame into the reconstruction task. Furthermore, it has been demonstrated that the practical scatter approximation provides similar improvements as the SPSC method when the tracer/scatter distribution changes minimally in Section 5.5.5.3. The same approach was also applied to the smoothed randoms estimate thus resulting in a frame-independent computation time and storage requirements for the dynamic scatter and smoothed randoms estimates. As an example, for the HRRT the scatter estimation as well as the smoothed randoms estimation takes about 15 minutes per frame in span 3

140
using a two dual core Xeon CPU with 3.06 GHz and 2 GB of RAM. Moreover, the size of each scatter or smoothed randoms sinogram is ~2 GB per frame. In typical dynamic studies, each scan is divided into 14-20 dynamic frames. Therefore, it takes up to 10 hours just to compute the scatter and randoms estimates for all frames in a single study without even reconstructing any images, and a disk space of 80 GB is required just to store the scatter and randoms sinograms for a single study when one needs to reconstruct the whole study at once such as in the iterative kinetic parameter estimation [1]. The proposed practical approximation is thus much more efficient and beneficial to dynamic studies since the time and storage requirements are independent of the number of dynamic frames.

6.3 A Practical Scatter and Randoms Approximation Technique

This technique is based on the following observations/assumptions:

- The scatter and randoms distributions are not very sensitive to the changes in the tracer distribution (i.e. the spatial scatter and randoms distributions are spatially much smoother compared to the distribution of the unscattered true events which correctly identifies the tracer distribution).

- The amount (magnitude) of scatter in each dynamic frame is proportional to the number of true coincidences (the pair of gamma rays originated from the same annihilation; i.e. the unscattered true events + the scattered events) in the frame, and likewise that of randoms is proportional to the global randoms counts in each frame.

In addition, the randoms estimate typically becomes less and less significant relative to the trues in the later frames of dynamic studies since the RF drops as the square of the activity thus increasing the ‘tolerance’ of the randoms estimate. By using a time-averaged scatter and randoms estimation followed by scaling the distribution according to the frame true coincidences and randoms, higher computational efficiency can be achieved. Furthermore, a potentially more accurate and less noisy scatter estimate can be obtained since it is derived from a data set with a higher number of counts when the spatial tracer distribution changes minimally as demonstrated using phantom studies in Section 5.5.5.3.
After incorporating this approximation technique into the histogram-mode algorithm (details of which are elaborated shortly), the equation for the 3D-OP becomes:

\[
\lambda_{j,f}^{m,l+1} = \lambda_{j,f}^{m,l} \sum_{i \in s_i} P_{ij} \sum_{j=1}^{J} P_{ij} \lambda_{j,f}^{m,l} + \frac{n_i}{c_{rf} \bar{r}_i^A + \frac{c_{rf} - A}{c_{rt} - s_i}} (6.1)
\]

where a subscript \( f \) has been added to represent the mean activity in voxel \( j \) for each individual frame (i.e. \( \lambda_{f,f}^{m,l} \)), \( c_{rf} \) is the number of random events (counts) in each individual frame, \( c_{rt} \) is the total number of random counts in the summed frame, \( c_{rf} \) is the number of true coincidences in each individual frame, \( c_{rt} \) is the total number of true coincidences in the summed frame. The true coincidences (which contain scatter) are obtained by subtracting the random counts from the total measured prompts for the given frame. \( \bar{r}_i^A \) is the randoms estimate for the summed frame (i.e. time-averaged estimate), and \( \bar{s}_i^A \) is the scatter estimate for the summed frame. In a similar fashion, this approximation can be incorporated into the list-mode algorithm.

This technique can be used reliably (as quantitatively demonstrated later) by using the plot of global true coincidence rate vs time, i.e. the dead time and decay corrected global Time Activity Curve (TAC which can be easily obtained before data reconstruction), in order to guide the decision-making procedure as to which frames are to be grouped together for the scatter/randoms-estimation tasks. Although different tracers might have different correspondences between the count rates and the spatial tracer distributions, it has been indeed found empirically that whenever there is a large change of slope or count rate in the global TAC, there is generally a larger change in the spatial radioactivity distribution for all the tracers examined in this work, and a separate scatter and randoms estimate needs to be performed for those frames. Additionally, all the frames in the later stage of the scan can be grouped to form the average estimate since the tracer activity distribution changes relatively slowly and the corresponding scatter and randoms distributions change minimally.
FIGURE 6.1: A global TAC for a $^{11}$C-dihydrotetrabenazine dynamic non-human primate study. The solid line connects the measured data points and aids in determining larger changes in the true coincidence rate.

Figure 6.1 shows a global TAC for a $^{11}$C dynamic non-human primate study. Data from the frames grouped by the ellipses are summed to perform the average scatter and randoms estimates. Subsequent sections will demonstrate how this technique works on the real data obtained from tracers, which have a relatively rapid change of their spatial distribution at least during the first part of a dynamic scan, such as $^{11}$C-dihydrotetrabenazine (DTBZ) and $^{11}$C-raclopride (RAC). The difference in the spatial scatter distribution between the conventional and the approximate method will be examined using the conventional non-human primate study as the gold-standard as will be discussed in later sections.
6.4 Experiments

6.4.1 Methods

The following methods (acronyms) which will be described later in this section were tested and shown in the result section:

a) Filtered Back Projection of the conventional scatter estimation (FBP-S-conventional)

b) Filtered Back Projection of the practical scatter approximation with one estimate from the entire scan (FBP-S-approx1e)

c) Filtered Back Projection of the practical scatter approximation using the global TAC as guidance (FBP-S-approx)

d) Histogram-mode reconstruction with the practical scatter and randoms approximation using the global TAC as guidance (3D-OP approx)

e) List-mode reconstruction with the practical scatter and randoms approximation using the global TAC as guidance (OP-LMEM approx)

f) 3D iterative OP reconstruction (both histogram-mode and list-mode) with the practical scatter and randoms approximation with one estimate from the entire scan (OP-approx1e)

g) 3D iterative OP reconstruction (both histogram-mode and list-mode) with the practical scatter and randoms approximation using the global TAC as guidance (OP-approx)

Non-human primate study: The non-human primate study described in Section 5.5.4.1 was used.

Comparison schemes for validating the practical scatter/randoms approximation technique: The quantitative accuracy of this technique was checked on two levels. First the spatial scatter (randoms) distribution and TACs themselves were compared between
the conventional and the approximate method using the non-human primate data (with rapid varying spatial tracer distribution). Since the estimated scatter and randoms sinograms contain many segments (direct and oblique planes), it is not very efficient to compare the sinograms directly. In Chapter 5, sinograms were used when comparing the scatter estimates since the scatter tail-fitting process is done for each sinogram plane thus clearly demonstrating the axial scatter distribution and the over-scaling bias across the planes using sinograms. However, one needs to also compare the 3-dimensional spatial scatter and randoms distributions in this case. As a result, estimated scatter and smoothed random events were reconstructed using FBP due to its linearity and fast nature (see Section 2.3.1) for each frame separately (i.e. the conventional method), for a summed frame with the entire scan duration (i.e. the approximate method with only one estimate), and for a number of summed frames determined according to the behaviour of the global count rate curve (i.e. the approximate method with three estimates: see Figure 6.1) followed by scaling according to the global number of true coincidences or randoms in each frame. As discussed in Sections 1.7.4 and 2.3.1, the HRRT contains gaps between the detector heads which in turn introduce gaps in the measured data. Moreover, FBP requires continuous data set across the FOV, and therefore, the missing data in the gaps need to be filled in order to use FBP. For the case of scatter and randoms, only the randoms sinograms contain gaps which need to be filled, whereas the scatter estimate is calculated for the entire sinogram space and does not require any gap-filling. The gap-filling was performed using the constrained Fourier space method [60], and FBP was used with a ramp filter at the Nyquist frequency. On a second level, the effect of the approximate scatter and randoms correction on the quantitative accuracy of the final images was investigated by reconstructing the emission data using 3D iterative OP algorithm (both histogram-mode and list-mode) in span 3 for both the conventional and approximate scatter/randoms estimations. The following comparisons were performed:

1) Spatial scatter and smoothed randoms distribution comparisons: For the non-human primate study, the reconstructed scatter and smoothed randoms images obtained from the approximate method were subtracted from the ones obtained using the conventional method (gold-standard images), and the difference was then divided by the gold-
standard images. This relative spatial difference was examined visually. A radial profile was also placed at the plane with the greatest difference in the scatter images between the conventional and the approximate method.

2) Time activity curve (TAC) comparisons: The time activity curves, with the ROIs (as shown in Figure 6.2) in the striatum (right/left caudate, putamen, and ventral striatum), cerebellum, and a number of selected cold regions, were plotted over 14 frames of the non-human primate scan for the scatter, randoms, and the emission (unscattered trues) images without the decay and dead time corrections. Two kinds of cold spots are defined here; one is defined (t-cold, s-hot) as the cold spot in the unscattered trues image with a corresponding hot spot in the scatter image; i.e. higher density region with no or very low radioactivity, for example bones for most tracers. The other one (t-cold, s-cold) is defined as the cold spot in the unscattered trues image with a corresponding cold spot in the scatter image; i.e. lower density region with very low or no radioactivity such as air and water as shown in Figure 6.2. Those two kinds of cold spots were introduced to check the limit of the approximate method since the radioactivity estimated in the cold spots is sensitive to small changes in the scatter estimate. The TACs were then compared between the images reconstructed using the conventional and the approximate methods by estimating the ratio of their values relative to each frame (i.e. ROI_{approximate} / ROI_{conventional} for each frame). In addition, the same TAC analysis was also performed in voxel-level for the emission images, and the voxel based ratio was calculated between the approximate and the conventional method (similar to the ROI ratio defined above).
3) Binding potential (BP) comparisons: The binding potential for the non-human primate study was computed using the tissue input Logan graphical approach [49] using the cerebellum as reference region for the conventional and the approximate methods.

### 6.4.2 Results

Spatial scatter distribution comparisons: The relative difference in the spatial scatter distribution between the approximate (with 3 estimates) and conventional (used as the gold-standard as discussed previously) method was first examined visually. Large differences were only observable outside the image of the object. When limiting the comparison to the actual image of the object, a maximum difference or spatial inaccuracy of approximately 10% was observed as shown in Figure 6.3. Similar results have been found from the smoothed randoms images. The contribution of this difference to the TAC analysis will be addressed shortly.

*FIGURE 6.2: ROIs for the selected hot and cold spots in the striatum plane placed (a) in the emission image and (b) in the scatter image which is obtained by reconstructing the scatter events using the Filtered Back Projection.*
FIGURE 6.3: The radial profile comparison in the scatter images of the non-human primate study between the conventional (FBP-S-conventional) and the approximate method with three estimates obtained using the global TAC as guidance (FBP-S-approx) in the plane with the greatest observed difference.
**FIGURE 6.4:** The scatter TAC (without decay and dead time corrections) comparison for the right caudate region of the non-human primate study between the conventional (FBP-S-conventional) and the approximate method with one (FBP-S-approx1e) and three estimates obtained using the global TAC as guidance (FBP-S-approx).

*Scatter and randoms TACs comparisons between the conventional and the approximate scatter and randoms correction method:* The TAC and ratio of TAC comparisons of the ROIs placed on the scatter images for the non-human primate study are shown in Figures 6.4 and 6.5.

Hot spots: As shown in Figure 6.4 with the corresponding TAC ratio shown in Figure 6.5a, the scatter TAC of the conventional method agrees with the approximate method (one estimate; i.e. FBP-S-approx1e) within 10% except for the first frame which is during the early stage of the scan where there was a significant change in the tracer distribution. The ratio of the TAC basically shows the variation in the distribution between the frame-based estimate and the time-averaged estimate. Figure 6.5a also demonstrates the improvement achieved with the approximate method using the global TAC as guidance (FBP-S-approx) according to the framing scheme described in Section 6.3 (Figure 6.1); i.e. the conventional and approximate methods agree within 5% for all
frames. The TAC for the right-caudate was arbitrarily chosen, and similar results were obtained from all other ROIs. The same analysis for the smoothed randoms images was performed, and the conventional method agrees with the approximate method within 5% for both one and three estimates; the difference of up to 5% was observed only in the last three frames, where the RF is lower than 15%.

Cold spots: As shown in Figure 6.5b, the approximate method works quite well in the t-cold, s-hot regions, while somewhat higher variation was observed for t-cold, s-cold regions. Again, significant improvement achieved using the global TAC as guidance is shown for the cold spots. Very similar results were obtained for the smoothed randoms TAC in the selected cold spots as well.

**FIGURE 6.5:** (a) The ratio of the scatter TAC comparisons for the right-caudate region of the non-human primate study between the conventional and the approximate method with one (FBP-S-approx1e) and three estimates obtained using the global TAC as guidance (FBP-S-approx), (b) The ratio of the scatter TAC comparisons for the selected cold spots as shown in Figure 6.2 which contain low radioactivity but correspond to higher density regions, (c) The ratio of the TAC comparisons for the right-caudate region of the non-human primate study between the conventional and the approximate method with one (OP-approx1e) and three estimates obtained using the global TAC as guidance (OP-approx), and (d) The ratio of the TAC comparisons for the selected cold spots as shown in Figure 6.2 with three estimates.
TACs comparisons between the conventional and the approximate scatter and randoms correction method: The ratio of the TAC comparisons between the conventional and the approximate method with one and three estimates (using the global TAC as guidance) for the non-human primate study are shown in Figure 6.5c. The net contribution from the difference in the scatter (as mentioned in the spatial scatter distribution comparison) and randoms estimates between the conventional and the approximate method with one estimate (the worst case scenario) is shown to be less than 3% in the TAC as depicted in Figure 6.5c. The improvement achieved with three estimates is also shown: the TAC using the approximate method with three estimates agrees with that using the conventional method within 1%. The same analysis has also been done for the selected cold spots, and the TAC using the approximate method (three estimates) agrees with that using the conventional method within 4% as shown in Figure 6.5d. Here a better ROI ratio was obtained for the t-cold, s-cold spot (very low activity) as compared to that of the t-cold, s-hot spot since the SF for the t-cold, s-cold spot is much lower than that of the t-cold, s-hot spot; thus, the scatter estimate contributes much less to the emission image for the t-cold, s-cold spot. The ratio of the voxel TAC comparison also showed similar and consistent results; the voxel TAC using the approximate method with three estimates agrees with that using the conventional method within 1% for the hot voxels and within 4% for the cold voxels in the object as depicted in Figure 6.6.
**FIGURE 6.6:** The ratio of the voxel TAC comparisons for the representative hot and cold voxels with three estimates

<table>
<thead>
<tr>
<th></th>
<th>3D-OP</th>
<th>3D-OP approx</th>
<th>OP-LMEM</th>
<th>OP-LMEM approx</th>
</tr>
</thead>
<tbody>
<tr>
<td>Right-caudate</td>
<td>2.70</td>
<td>2.70</td>
<td>2.77</td>
<td>2.78</td>
</tr>
<tr>
<td>Left-caudate</td>
<td>2.93</td>
<td>2.92</td>
<td>2.89</td>
<td>2.90</td>
</tr>
<tr>
<td>Right-putamen</td>
<td>3.07</td>
<td>3.07</td>
<td>3.16</td>
<td>3.17</td>
</tr>
<tr>
<td>Left-putamen</td>
<td>3.01</td>
<td>3.01</td>
<td>3.01</td>
<td>3.01</td>
</tr>
<tr>
<td>Right-ventral striatum</td>
<td>2.42</td>
<td>2.42</td>
<td>2.40</td>
<td>2.40</td>
</tr>
<tr>
<td>Left-ventral striatum</td>
<td>2.55</td>
<td>2.55</td>
<td>2.49</td>
<td>2.49</td>
</tr>
</tbody>
</table>

**TABLE 6.1:** Binding potential (BP) comparisons between the conventional and the approximate scatter and randoms correction method

*Binding potential (BP) comparisons between the conventional and the approximate scatter and randoms correction method:* The BP analysis also showed an excellent agreement (within ~0.2%) between the conventional and the approximate method (using...
the global TAC as guidance) as shown in Table 6.1. Both histogram-mode and list-mode reconstructions show the same agreement between the conventional and the approximate method. Next, the practical approximation will be combined with the dual reconstruction scheme.

6.5 Combining the Practical Scatter/Randoms Approximation with the Dual Reconstruction Scheme

The previously described non-human primate study was used to demonstrate the accuracy and efficiency of the combined approximated dual reconstruction method as shown in Figure 6.7. The 3D-OP and OP-LMEM algorithms were applied according to the scheme described in Section 4.4.1 for the non-human primate study.

As shown in Figure 6.7 (left), the TACs agree very well between the conventional 3D-OP and the approximated dual reconstruction. The variation in the TAC ratio (up to 5%) as shown in Figure 6.7 (right) is mainly contributed from the statistical difference between the histogram-mode angular subset and the list-mode temporal subset as described in Chapters 2 and 4. Figure 6.7 (right) also demonstrates the accuracy of the practical scatter and randoms approximation: the TACs of the 4 frames (as circled) reconstructed using the approximated 3D-OP agree with those using the conventional 3D-OP within 1%. The binding potentials obtained from the conventional 3D-OP agree
with those obtained from the approximated dual reconstruction scheme within 3% as shown in Figure 6.8 (left).

![Figure 6.8: The BP value comparison (left) and the reconstruction time cost comparison (right; the number of counts ranges from 140M to 13M, and the frames were sorted by number of counts in a descending order.)](image)

Figure 6.8 (right) shows the reconstruction time cost for both histogram and list-mode OP algorithms. 3D-OP shows a constant time cost for each frame since the number of LORs stays constant, whereas OP-LMEM shows a decrease of time cost at frames with a lower number of counts since the time cost in list-mode is proportional to the number of counts in the frame. The time we gained by applying the dual reconstruction scheme is depicted by the shaded area; in the case of the non-human primate study, we gained about (i) 20% of time as compared to reconstructing every frame with 3D-OP and (ii) nearly 4 times faster performance (75% gain in time) on the scatter and randoms estimations (i.e. computing 3 estimates instead of 14) and (iii) 4 times less storage cost by applying the practical approximation. The overall time gain depends on the number of dynamic frames and the relative difference in time cost between the reconstruction process and the scatter/randoms estimations. In this case, an overall gain of about 35% in time was achieved since the reconstruction task takes longer than the scatter and the smoothed randoms estimation processes.
6.6 Conclusion

In conclusion, the approximate method was demonstrated to be equivalent to the conventional method for studies with high number of counts, and it performs potentially better than the conventional method for studies with low number of counts. Furthermore, by using the proposed method a 75% gain in the estimation time and data storage was achieved when computing the scatter and randoms estimates for the non-human primate study presented here; i.e. estimating the scatter and randoms takes about 10 hours and a storage size of 80 GB per study using the conventional frame-based method, whereas it only requires 2.5 hours and 20 GB of storage using the proposed method. Additionally, the combined approximated dual reconstruction method has been demonstrated to be accurate and efficient.
Chapter 7

Conclusion
7.1 Summary

With regard to the scatter-corrected list-mode reconstruction and the dual reconstruction scheme, results from phantom and non-human primate studies look excellent from the comparisons presented in Chapter 4. In particular the excellent agreement reached between OP-LMEM and 3D-OP indicates that the implementation of the scatter correction method into the list-mode algorithm is appropriate thus reaching the same level of quantification accuracy as 3D-OP. For the current version of our 3D-OP and OP-LMEM algorithms, we have found that for frames with less than 50M counts OP-LMEM is more advantageous time-wise, and that indicates the threshold value for the dual reconstruction scheme to be 50M. A time gain of 40% was obtained from the phantom study and a 20% time gain was obtained from the non-human primate study by applying the dual reconstruction scheme as discussed in Chapter 4.

Furthermore, a scatter calibration technique was developed to reduce the noise in the axial scatter distribution across the planes and to compensate for the overestimation bias in the scatter correction for scans with high RF and/or low number of counts, a situation often encountered in dynamic imaging on scanners with a large number of LORs such as the HRRT. For example in the phantom study, an estimated SF of as high as 270% was obtained using the conventional method from a frame with 90% RF and 2M counts (similar results were obtained from human studies as well), whereas the unbiased SF was ~70%. A much more consistent global SF was obtained by GSC, SSC, and SPSC as compared to the conventional method, and a smoother scatter estimation across the planes was achieved by applying SPSC. It has been identified that the SSC method can be applied independent of the early and later stages of the scan and its reference formation is very flexible, while the SPSC method should be applied only during the later stage. However, since the tracer and the corresponding scatter distribution varies relatively slowly in the later stage, a practical scatter approximation can be applied instead of the SPSC method. The calibration technique was also applied on human studies and showed more accurate results. The proposed method is thus a solution to the overestimation bias described in [57], and it allows one to have more freedom on making framing and scanning protocols when the bias in the scatter at low number of counts is a major concern. In addition, the proposed scatter calibration method is
particularly beneficial for low count $^{15}$O studies and for the image derived input functions for dynamic brain studies where one is interested in having short frames with only 5 to 10 seconds image duration [56, 57].

As for the practical approximation, in a situation of a low number of counts per frame the approximate scatter estimate has been demonstrated to produce a better axial uniformity and less noisy emission images similar to the improvements obtained by the SPSC method from the phantom studies. For the non-human primate studies in which there was a sufficient number of counts within each frame to obtain an accurate scatter and smoothed randoms estimate for the conventional method (i.e. good as an absolute reference), a very good agreement in the spatial scatter and smoothed randoms distribution, TAC, and BP comparisons between the conventional and approximate method was obtained. Moreover, for tracers with relatively uniform spatial distribution and phantom studies, the approximate method is expected to be more accurate than the conventional method due to the higher statistical quality of the data used in the scatter and randoms estimates. Furthermore, the time and storage costs for scatter and randoms estimates in the dynamic studies became largely independent of the number of frames by applying the proposed method. We achieved a nearly 4 times reduction in computational demands (both time and storage) related to the scatter and randoms estimates for the studies presented here; i.e. calculating 3 estimates instead of 14 for the non-human primate study as depicted in Figure 6.1. Additionally, the combined method which incorporates the practical approximation and the dual reconstruction scheme has been demonstrated to be accurate and much more efficient than the conventional method.

7.2 Concluding Remarks

The final proposed reconstruction protocol for the dynamic brain imaging in high resolution PET is as follows:

- Apply motion correction to the acquired data

- Use the global TAC as the guidance to determine the framing scheme for the practical scatter and randoms approximation method (tracer specific)
• Once the averaged scatter estimate in the later stage of the scan or the one with the highest statistical quality is obtained, use it as the reference and apply the SSC method to calibrate the other scatter estimates (targets) as shown in Figure 7.1

• Scale the averaged scatter and smoothed randoms estimates using the individual to summed frame trues and randoms ratio, respectively, to obtain the individual frame scatter and smoothed randoms estimates

• Check the number of counts in each dynamic frame and apply the dual reconstruction scheme

• Obtain the fully corrected emission images

![Figure 7.1: An example of the scatter calibration scheme in the practical scatter approximation](image)

As described in Sections 1.8.6 and 3.6, object motion compensation needs to be applied before the scatter calibration and the practical approximation as the motion significantly affects the scatter estimate in each frame as well as in the reference frame.
In conclusion, both the scatter calibration and the practical approximation can be applied to brain scans, whereas the dual reconstruction scheme is object-independent. They are modular and not scanner specific. Therefore, these methods can be applied to any PET scanners and are particularly beneficial for high resolution scanners which have a large number of LORs such as the HRRT. All in all, these studies contribute to further improving the quantitative accuracy thus extending the current protocols, decreasing the computational burden, and enabling reconstructions of multiple studies within a shorter time for dynamic brain imaging in high resolution PET.
REFERENCES


[29] D. Keator. Personal communication.


