

## Review

## System models for PET statistical iterative reconstruction: A review

A. Iriarte<sup>a,\*</sup>, R. Marabini<sup>b</sup>, S. Matej<sup>c</sup>, C.O.S. Sorzano<sup>a,d</sup>, R.M. Lewitt<sup>c</sup><sup>a</sup> Department of Information and Telecommunication Systems, University CEU San Pablo, 28668 Madrid, Spain<sup>b</sup> Escuela Politécnica Superior, Universidad Autónoma de Madrid, 28049 Madrid, Spain<sup>c</sup> Department of Radiology, University of Pennsylvania, Philadelphia, PA 19104, USA<sup>d</sup> National Center of Biotechnology (CSIC), 28049 Madrid, Spain

## ARTICLE INFO

## Article history:

Received 29 January 2015

Received in revised form 9 October 2015

Accepted 9 December 2015

## Keywords:

Nuclear imaging

PET

Statistical reconstruction

System matrix

System model

System response model

## ABSTRACT

Positron emission tomography (PET) is a nuclear imaging modality that provides *in vivo* quantitative measurements of the spatial and temporal distribution of compounds labeled with a positron emitting radionuclide. In the last decades, a tremendous effort has been put into the field of mathematical tomographic image reconstruction algorithms that transform the data registered by a PET camera into an image that represents slices through the scanned object. Iterative image reconstruction methods often provide higher quality images than conventional direct analytical methods. Aside from taking into account the statistical nature of the data, the key advantage of iterative reconstruction techniques is their ability to incorporate detailed models of the data acquisition process. This is mainly realized through the use of the so-called system matrix, that defines the mapping from the object space to the measurement space. The quality of the reconstructed images relies to a great extent on the accuracy with which the system matrix is estimated. Unfortunately, an accurate system matrix is often associated with high reconstruction times and huge storage requirements. Many attempts have been made to achieve realistic models without incurring excessive computational costs. As a result, a wide range of alternatives to the calculation of the system matrix exists. In this article we present a review of the different approaches used to address the problem of how to model, calculate and store the system matrix.

© 2015 Elsevier Ltd. All rights reserved.

## Contents

1. Nuclear medicine computed imaging .....	31
2. PET imaging applications and challenges .....	31
2.1. Applications .....	31
2.2. Challenges for PET technology .....	31
3. PET reconstruction overview .....	32
3.1. Algorithms .....	33
3.2. System matrix for iterative reconstruction .....	34
3.3. Image-degrading effects to be modeled in iterative reconstruction .....	34
4. Approaches for calculating the system matrix .....	35
4.1. Empirical methodologies .....	36
4.2. Monte Carlo (MC) methodologies .....	36
4.2.1. Accuracy .....	36
4.2.2. Time .....	36
4.2.3. Storage .....	36
4.3. Analytical methodologies .....	37
4.4. Hybrid approaches .....	38

\* Corresponding author. Tel.: +34 656230847.

E-mail address: [ana.iriarte.ruiz@gmail.com](mailto:ana.iriarte.ruiz@gmail.com) (A. Iriarte).

5.	Modeling of individual effects and their combination into the system model .....	38
5.1.	Effects separation .....	38
5.2.	System matrix factorization for resolution modeling .....	40
5.2.1.	Factorization for resolution modeling using both projection and image blurring .....	40
5.2.2.	Factorization for resolution modeling with emphasis on projection blurring matrix .....	40
5.2.3.	Factorization for resolution modeling with emphasis on image blurring matrix .....	41
5.2.4.	Alternative uses of the blurring matrices .....	41
5.3.	Alternative to factorization: multivariate integral formulation .....	42
6.	Topics of current research .....	42
6.1.	Efficient implementations of iterative reconstruction with system modeling .....	42
6.2.	Performance characteristics of iterative reconstruction with system modeling .....	43
7.	Discussion .....	43
8.	Summary .....	44
	References .....	44

## 1. Nuclear medicine computed imaging

Nuclear medicine imaging techniques, whose two major subbranches are single-photon emission computed tomography (SPECT), and positron emission tomography (PET), provide images reflecting physiologic functions unlike other traditional techniques, such as X-ray computed tomography or ultrasound, which yield anatomical structures. In order to image properties of the body's physiology, both the PET and SPECT techniques begin with the injection of a radiopharmaceutical into the subject under study. A radiopharmaceutical consists of a tracer compound that interacts with the body and a radioactive label that, by way of the emission of gamma rays, is used to track the flow and distribution of the tracer compound. A wide range of radioactive-labeled compounds have been synthesized as tracers that target biologically interesting markers or pathways thus permitting the measurement of quantities that go from glucose metabolism [1] to gene expression [2].

SPECT uses radionuclides that emit single gamma rays. Such radionuclides require the use of a collimator between the patient and the detectors in order to define the direction from which the gamma rays are coming, thereby providing spatial information. SPECT collimation does an inefficient use of the emitted radiation, as the collimation strongly limits the number of detected gamma rays. On the other hand, PET studies do not require a physical collimation, but use a scheme of detection called “electronic collimation” (shown in Fig. 1), allowed by detection of collinear photon pairs generated by positron emitting radioisotopes.

## 2. PET imaging applications and challenges

The first PET scanners were introduced in the mid 70s. Since then, they have been used for a broad range of applications, which are covered in Section 2.1. The requirements that these applications have been imposing have translated into several challenges, as described in Section 2.2. In order to overcome these challenges, the PET technology has evolved and improved in a variety of ways. These include the mathematical procedures for image reconstruction, and the statistical models on which these are often based, which constitute the topic addressed by this article.

### 2.1. Applications

Nowadays, the majority of applications of PET in clinical (i.e., human patient) medicine are found in oncology imaging, either for cancer diagnosis or for cancer treatment planning and guidance [3–6]. However, there are other uses of clinical PET that do not involve oncology [7,8]. Some of them have been around for a long time such as cardiac imaging with PET [9], others are relatively new and provide new insights into how to make diagnosis of diseases where other techniques fail to provide sufficient information.

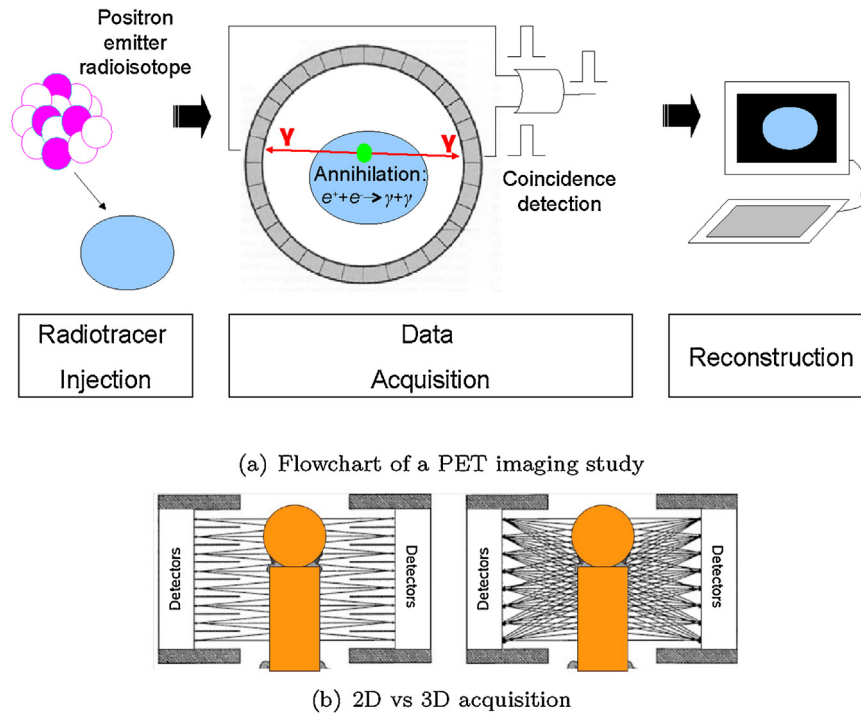
Although initially conceived for whole-body clinical purposes, the earlier clinical systems were also used for brain and animal (predominantly in the larger laboratory animals such as nonhuman primates) imaging [10]. These two areas have steadily gained importance in the clinical and research environments, the former (brain imaging) due to the explosive interest in the early diagnosis of neurodegenerative disorders and in the study of cognitive and affective processes, the latter (animal imaging) due to the increased use of small-animal models for studying disease and for designing and evaluating new drugs in the pharmaceutical research and development pipeline. As technological advances have allowed whole-body PET scanners to settle in the clinical practice, there has been significant motivation to design dedicated units that maximize the PET performance in brain and small animal studies. Given the characteristics of the objects being imaged in these fields, the resolution and sensitivity of whole-body human PET are often insufficient for a broad range of applications, and specific devices for brain and small-animal PET have been developed. The improved ability of brain-dedicated PET scanners to provide reliable quantitative assessment of the activity in small cerebral regions involved in specific functional pathways has been already demonstrated [11–13]. Similarly, the development of small animal PET scanners has allowed the clear visualization and precise quantification of the biodistribution of PET tracers in small structures within rodents (see [14–16] and the references therein).

### 2.2. Challenges for PET technology

Despite the major advances mentioned above, the use of PET in a wider range of applications will be dictated, along with other limitations (cost-effectiveness, availability of PET tracers, user friendliness, etc.) by the requirements these applications impose on the intrinsic performance of the PET tomographs. The challenges that should be confronted depend on the final target.

Improving the typical spatial resolution of around 5 mm of the human images is not a main goal in the current routine clinical PET applications. However, there is a need to improve the signal-to-noise ratio of whole-body studies. The mean photon path distance from emission to detection is greater than in other modalities, which increases the probability of attenuation and/or Compton scattering inside the patient's body, especially in larger persons. This increases the amount of noise, produces quantization error and reduces contrast in the reconstructed images.

On the other hand, animal PET scanners are designed to perform studies on small subjects (typically mice and rats), which often demand to assess concentrations of radioactive nuclide in structures with dimensions of a millimeter or less. The parameter getting the most attention is thus spatial resolution, with the objective of achieving sub-millimeter isotropic volume resolution. Whether this range is able to be reached is not clear since there are several complex factors interacting to limit spatial



**Fig. 1.** (a) After the radiopharmaceutical injection, positrons are emitted within the subject's body, combine with nearby electrons and become annihilated. The result of the annihilation is a pair of 511-keV gamma photons being released in opposite directions. PET scanners use pairs of radiation detectors to measure the nearly simultaneous, coincident interaction of the 511-keV photons. The line that connects any two detectors in the PET scanner is called a line of response (LOR). The data recorded by the scanner is a collection of coincidence detections. In the reconstruction step, a mathematical procedure has to be implemented to convert the acquired data into tomographic images. (b) In two-dimensional scanners (left), multiple planes of detectors surround the patient with dense material, or septa, so that coincidence events are collected only between pairs of detectors that are either in a single plane or in a pair of adjoining planes. In contrast, the three-dimensional scanners (right) have no septa so that coincidence photons can be detected between any two planes.

resolution during data formation and collection (positron range, pair non-collinearity, crystal penetration, inter-crystal scatter, detector decoding errors, etc.). [17] performs a quantitative analysis of these effects and estimates that the ultimately achievable spatial resolution for pre-clinical PET cameras with appropriate detector designs is 0.83 mm.

Furthermore, since the amount of injected dose is rather limited for small animals (in order to not perturb their biological systems), a high level of sensitivity is required to achieve an acceptable number of coincidences.

For neuro-imaging, the desired spatial resolution is slightly higher compared to the whole-body imaging (around 3 mm<sup>3</sup> is normally sufficient). However, in applications in which the structure of distribution of tracer within the organ is important for diagnosis, higher resolution might be needed. Furthermore, brain applications constitute an area of wide use of kinetic modeling methods in PET [18]. The kinetic studies are based on dynamic PET data where short frames, with low signal-to-noise ratio, are acquired in order to capture the temporal distribution of the activity concentration after the radiotracer injection. This makes high sensitivity a particularly important requirement in most brain applications.

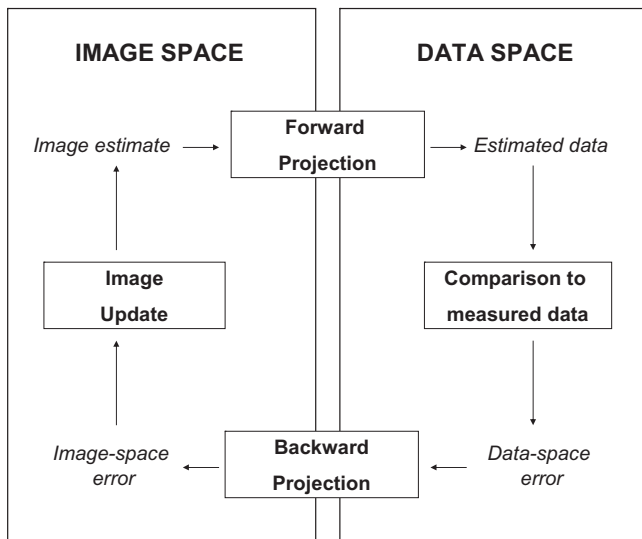
In all the modalities, the requirements are first addressed at the stage of design and development of instrumentation. For example, in body systems, the introduction of the time-of-flight (TOF) systems, that are able to approximately locate the annihilation event along the path of the two photons, has led to remarkable improvements in the final SNR in the images. In small animal devices, the measurement of the depth of interaction (DOI) of photons within detectors reduces the magnitude of the parallax effect, which is known to significantly degrade the resolution of the images [19].

These advanced technologies greatly alleviate some of the PET issues, but there is a point at which the instrumentation encounters some limitations. For example, the requirements of sensitivity of the small animal applications can be satisfied using detector modules with long crystals, placed in a gantry with the smallest possible diameter [20]. However, both long pixelated crystals and small gantries lead to a significant resolution loss due to the parallax effect. Moreover, the increase of the gantry diameter to reduce the penetration effect (at the cost of reducing the sensitivity) would lead to another form of spatial resolution degradation due to the non-collinearity of the pairs [21].

Under these paradoxical circumstances, other means to meet the demanding requirements the PET studies impose have been explored. A particularly important area is the reconstruction step, in which the events collected are converted to three-dimensional images. Certain algorithms are capable of preserving and even enhancing (as shown in [22] for the parallax error correction) the hardware advantages.

### 3. PET reconstruction overview

As pointed out in the previous section, the achievement of the ultimate PET performance will require the help of advanced reconstruction techniques [23]. In this section, we introduce the different approaches to estimate the three-dimensional radiotracer distribution from the recorded data (Section 3.1). Then, we focus on the successful type of reconstruction methods known as iterative algorithms, specifically on the statistical model these methods make use of (Section 3.2) and on the effects of the PET acquisition process they should take into account and compensate (Section 3.3).



**Fig. 2.** Flow diagram of an iterative reconstruction algorithm. Starting with an initial image guess, the algorithms iteratively compute new image estimations based on the comparison of the estimated and measured data.

### 3.1. Algorithms

The early techniques of PET image reconstruction were analytical approaches based on the method of Filtered Back Projection (FBP) [24]. FBP is a mathematical technique based on an idealized model of PET, which ignores many significant physical factors. Specifically, FBP assumes that the number of events measured in each LOR (line of response, see Fig. 1) is equal to the line-integral of the radio-isotope distribution along the line connecting the centers of the two corresponding crystals, from which the image can be reconstructed using analytical inversion formulas. In spite of its approximate essence, FBP has enjoyed widespread use due to its computational simplicity. Regrettably, FBP amplifies the noise when it is applied to the low-count data of PET imaging (it is still a good method for applications where the number of measured counts is high, as computed tomography, and/or where linearity of the reconstruction is needed). Consequently, there has been a great interest in the development of reconstruction alternatives that provide better resolution and noise characteristics [25–28].

The introduction of the iterative reconstruction methods allowed for the explicit inclusion of the physical factors ignored by analytical schemes. These algorithms achieve improved spatial resolution and signal-to-noise ratio, while maintaining the quantitative nature of the data [29,30]. The price of this added refinement is that the resulting set of equations describing the problem becomes very large, and solving for the tracer distribution by direct inversion of the problem becomes unworkable. In this case, the equations must be solved using iterative methods. All iterative reconstruction methods share a number of common traits. The general model, shown in Fig. 2, involves repeating the process of projecting an image estimate (the so-called forward projection operation), comparing the estimated projections to the measured data to compute some form of error, backprojecting the error (backward projection operation) and using the error to update the image estimate.

Two wide classes of iterative algorithm exist, namely algebraic and statistical. Classical algebraic iterative algorithms, such as ART and MART [31] are based on the Kaczmarz method of solving systems of linear equations [32]. Although not widely used in nuclear medicine, they form an important basis for understanding the statistical algorithms that were developed later and for which iterative methods are mainly used. Statistical algorithms have two basic

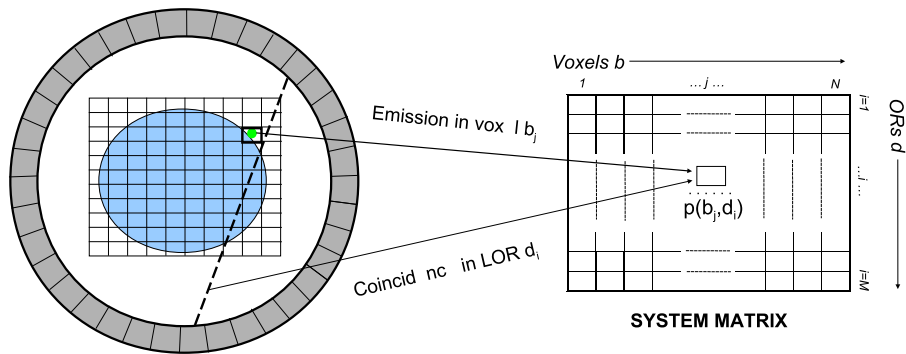
parts: a statistical criterion (the basis for determining which image among the many possible is to be the solution) and a numerical algorithm (the method for finding the solution prescribed by the criterion). The most successful early statistical algorithm, MLEM [33] uses a Poisson model of the produced data, related to the maximum likelihood (ML) statistical criterion, along with the expectation maximization (EM) algorithm as numerical method. A serious disadvantage of MLEM is its slow convergence, which can lead to prohibitive reconstruction times on standard computer platforms [27]. Several acceleration techniques have been proposed for the MLEM algorithm. Lewitt and Muehllehner [34] improve convergence speed by incorporating an over-relaxation parameter. Tanaka [35] uses a frequency amplification method to accelerate the estimation of the higher-frequency components in the image. The ordered subsets expectation maximization (OSEM) method [36] divides the data into subsets, giving a speed-increase factor proportional to the number of subsets chosen. Browne and DePierro [37] proposed the row-action maximum-likelihood algorithm (RAMLA), which uses as many subsets as there are projections. However, these block methods do not necessarily converge, as reported empirically by Byrne [38]. The space-alternating generalized EM (SAGE) [39] improves the convergence rate. Rebinning methods such as single slice rebinning (SSRB) [40] or Fourier rebinning (FORE) [41] can reduce the dimensionality of 3D acquisitions by performing a set of 2D reconstructions to obtain volumetric data, although there is a loss of image quality with respect to the more time consuming fully 3D implementation.

In addition to the slow convergence issue, there is another important drawback related to the MLEM derived methods: at high iteration numbers, images exhibit high-variance behavior [42]. This is usually attributed to either the fact that there is no explicit stopping rule in this kind of iterative reconstruction or to the statistical (noisy) nature of the detection process and reconstruction method [43]. A usual approach to overcome this drawback is either design some kind of practical stopping rule [44–46] or to smooth the images with kernels [47,48] filters [49] or wavelet based methods [50]. As for solutions integrated in the reconstruction algorithm, the maximum a posteriori (MAP) methods [51–54] offer a very convenient way of reinforcing desirable properties in the reconstructed image by incorporating *a priori* information (priors) that models the distribution of activity and noise in the data. The regularizing influence of the prior controls the variance of the reconstruction, and hence the MAP methods do not exhibit the instabilities at higher iterations encountered using MLEM [55]. On another note, dynamic RAMLA (DRAMA) is an iterative algorithm similar to RAMLA, which optimizes noise propagation from each subset to the reconstructed image by introducing and optimizing subset dependent relaxation parameters [56,57].

There exist also statistical algorithms based on a Gaussian model of the produced data, related to the weighted least squares (WLS) criteria, which result in quadratic objective functions. These can efficiently use many established numerical algorithms, such as coordinate gradient [58,59] or coordinate descent [60], and result in faster reconstruction algorithms.

Conventional PET reconstruction algorithms involve a preprocessing step where the data is binned as a function of the LOR spatial orientations, through structures called sinograms. The sinograms are a representation of the detections measured at a given plane through two-dimensional matrices, with the vertical columns representing evenly spaced projection angles and the horizontal rows representing spatial positions within the projections. Alternatively, LOR-based algorithms enable reconstruction of raw histograms (also known as LOR histograms), which use the native projection space as discretized by crystal pair indices, eliminating the interpolation step required to form the uniform sinograms [61,62]. Another option is the storage mode known as list-mode. List-mode





**Fig. 3.** The system matrix relates each basis function ( $b_j, j = 1, 2, \dots, N$ , being  $b_j$  cubic voxels in this figure) in the image space to each LOR ( $d_i, i = 1, 2, \dots, dM$ ) in the projection space. Specifically, each  $p(b_j, d_i)$  term of the system matrix stands for the probability that an emission within voxel  $b_j$  generates a coincidence in LOR  $d_i$ .

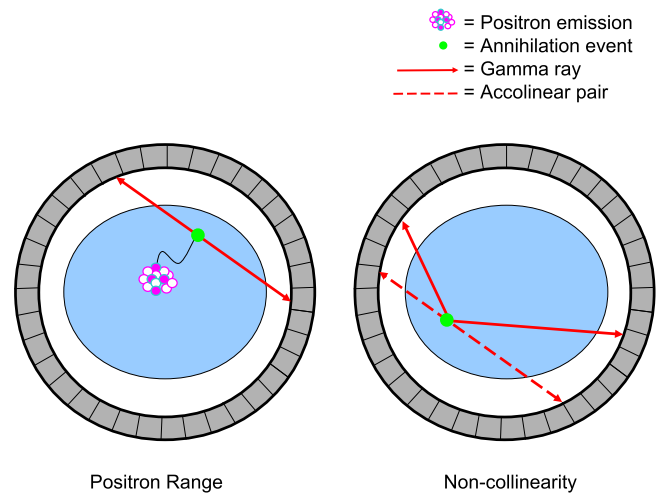
acquisition is achieved by storing information regarding the acquired events as they are detected one-by-one in the form of a list. The list-mode data is not amenable to sinogram-based reconstruction methods without first histogramming the data. However, there exist several versions of the MLEM and OSEM algorithms [63–65] for this type of data.

### 3.2. System matrix for iterative reconstruction

In order to perform the forward and backward projection operations, the iterative reconstruction algorithms require the use of the so-called system matrix (also known as system response matrix or system model). The system matrix describes the relationship between sources and data, taking into account the physical properties of the data acquisition process. Specifically, in a reconstruction in which the continuous radiotracer distribution is approximated as a sum of discrete basis functions  $b_j$  ( $j = 1, 2, \dots, N$ ) (normally cubic voxels, although blob basis functions have shown better noise properties [66]) and in which the scanner contains  $d_i$  ( $i = 1, 2, \dots, M$ ) possible LORs each  $p(b_j, d_i)$  term of the  $N \times M$  system matrix stands for the probability that an event generated in the region defined by the basis function  $b_j$  is detected in a detector pair  $d_i$  (see Fig. 3).

Besides the system matrix, iterative algorithms usually make use of a second matrix, whose  $N$  terms  $p(b_j)$  stand for the probability of detecting a pair of photons arbitrarily emitted from within function  $b_j$  (considering all possible detectors). Since this matrix accounts for the discrepancies in the sensitivities of the different basis functions, it is usually referred to as sensitivity matrix. Although alternative direct approaches exist [63,67–69], the terms of the sensitivity matrix are typically computed as the summation over the probabilities of detection of  $b_j$  in each LOR  $d_i$  given by the system matrix terms (i.e.  $p(b_j) = \sum_{i=1}^M p(b_j, d_i)$ ).

The quality of the images obtained by the iterative algorithms depends strongly on how faithfully the theoretical model incorporated through the system and sensitivity matrices depicts the real PET process. On the other hand, the price of the enhancements achieved by an accurate modeling is that the reconstructions require greater execution times and/or storage resources. This is because iterative algorithms perform repeatedly the forward and backward projection operations that require the knowledge of the whole system matrix, whose size is large and its computation is costly in high resolution 3D PET systems. Reaching a good compromise between image quality and efficiency is thus a key challenge in iterative reconstruction. This is the reason why much effort has been devoted to find alternative implementations of the system matrix that provide high quality images along with feasible reconstruction time and storage conditions. A review and classification of the numerous works on this topic is provided in this article.



**Fig. 4.** Positron physics-related effects. For illustration purposes, the non-collinearity is drawn much larger than it really is (the angle of deviation from  $180^\circ$  is a sample from a Gaussian distribution having FWHM of about half a degree).

### 3.3. Image-degrading effects to be modeled in iterative reconstruction

There are a variety of factors, inherent to the PET data acquisition process, that can contribute to the degradation of the reconstructed images:

- The geometry of the specific scanner.
- The positron physics (see Fig. 4), including the distance the positron travels before annihilating, which is termed positron range and the angular deviation from the ideal antiparallel directions of the emitted gamma rays, which is termed non-collinearity.
- The unwanted coincidences (see Fig. 5), including the coincidences in which the photons detected come from different annihilation pairs, termed random coincidences, and those in which one or both of the photons of the pair scatter in the object being scanned, termed scattered coincidences.
- The attenuation of photons prior to detection due to the absorption within the object being imaged (see Fig. 5).
- The intrinsic detector resolution. This is mainly conditioned by the geometric resolution of the detectors, which is roughly half the detector pixel width, and by two important physical effects: the crystal penetration and the inter-crystal scatter (depicted in Fig. 6). The crystal penetration occurs when a photon enters the crystal at an oblique angle and travels with no interaction through this crystal, before being absorbed in a different crystal. In this

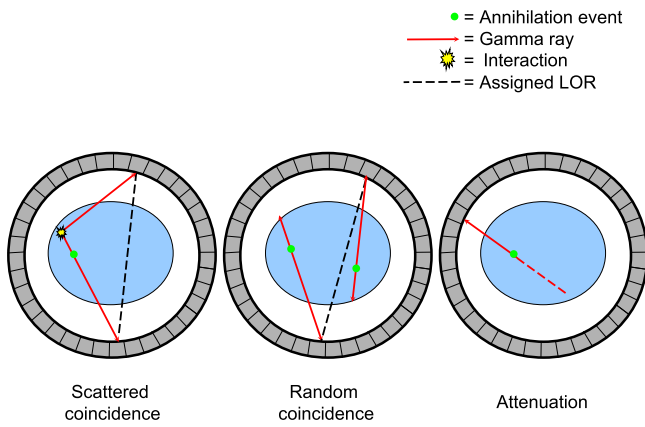


Fig. 5. Unwanted and attenuated events.

case, the oblique event is assigned a wrong LOR, and this source of error is known as parallax error (similar to the well-known parallax error when making a distance measurement using a thick ruler and viewing from an oblique angle). The inter-crystal scatter occurs when a photon undergoes several interactions in the scintillator and the event is assigned a wrong LOR. Inter-crystal scatter can arise for any angle of photon entry (oblique or non-oblique) into the crystal.

- The variations in the detection efficiency between detector pairs, due to physical constraints, like crystal imperfections, or to effects that take time during the data acquisition, such as misalignments in the timing windows, or the dead time the counting system is occupied processing previously detected events.

In order to compensate these effects during the image generation process it is necessary to develop models that describe them properly. The integration of all the models in the system matrix can be a challenging task if the matrix is too large and/or its calculation is too complex. Moreover, the amount of randoms increases at the square of the activity in the subject. In consequence, random coincidences do not fit in the linear formulation that relates the source and data vectors through the system matrix.

An approximate alternative procedure consists of considering certain effects (typically randoms and scatter coincidences on one side and attenuation and differences in detector efficiencies on the other) as errors in the data and treat them in the measurement space. Several methods exist to estimate or measure these

nuisance components, with a high level of accuracy. The most commonly implemented methods for estimating the randoms are the delayed coincidence method, that measures the random coincidences directly and the singles-based method, that estimates them from independent (singles) measurements [70]. Photon scatter correction is a slightly more complicated problem, that has generated a number of advanced approaches and still remains an area of active research. A discussion of existing ideas is beyond the scope of this paper, but several works provide an extensive account of methods to estimate the scatter [71,19,72,73].

Concerning the attenuation correction, one approach involves using a radioactive source of single photons (or photon pairs) that moves in a circular path just inside the detector ring. The probability of survival for each LOR of the scanner can be computed as the ratio of the number of photons (or photon pairs) detected with the patient present (the transmission scan) to the number detected in the absence of the patient (blank scan) [19]. Alternatively, the transmission scan can be used to reconstruct an image of the attenuation values via statistical methods [74–76]. With the advent of multimodality PET/Computed Tomography scanners, attenuation correction using separately acquired anatomical images has become a popular alternative [77,78].

Normalization weights to compensate the varying sensitivities among the different LORs include direct [79] and component-based methods [80–83]. In direct methods a source of activity that illuminates all LORs is scanned. Then, the normalization factors are estimated as the ratio between the ideal number of coincidences and those actually measured. The problem of these direct methods is that they require long scans for sufficient statistical accuracy. The component-based methods represent the normalization weights as a product of factors representing the different effects affecting detection efficiency. These approaches reduce the number of counts required by reducing the degrees of freedom in the normalization model so that the correction factors are computed by averaging over multiple detector pairs.

Once the values for correcting the different degradation factors have been estimated using one of the above mentioned methods, the corrections can be applied in two different ways. A first option is to apply the corrections before the reconstruction step, subtracting or multiplying the estimated values to the raw data. However, precorrected scans are no longer Poisson, which could bias the reconstruction in iterative algorithms. Alternatively, the corrections can be incorporated to the iterative reconstruction process itself, minimizing their impact on signal-to-noise ratio (SNR). In order to account for scatter and randoms in the reconstruction, their respective corrections are commonly added to the forward projection [84]. Rehfeld and Alber [85] prove the robustness of this approach for the scatter correction by comparison to a reconstruction that incorporates the model to the system matrix. An alternative scatter and randoms multiplicative scheme is proposed in [86]. The attenuation and normalization corrections are always multiplicative and can be easily incorporated to the algorithm. Several iterative reconstruction schemes incorporating one or both of these corrections in the forward, backward projection and/or the sensitivity term have been proposed [87–89]. In the following, the modeling of these effects will just be considered in the cases in which it is explicitly part of the system matrix, which, as will be shown later, includes the option of multiplying the elements of the system matrix by the normalization and attenuation correction factors.

#### 4. Approaches for calculating the system matrix

The three following subsections deal with the three basic methodologies that, in descending order of accuracy, can be used

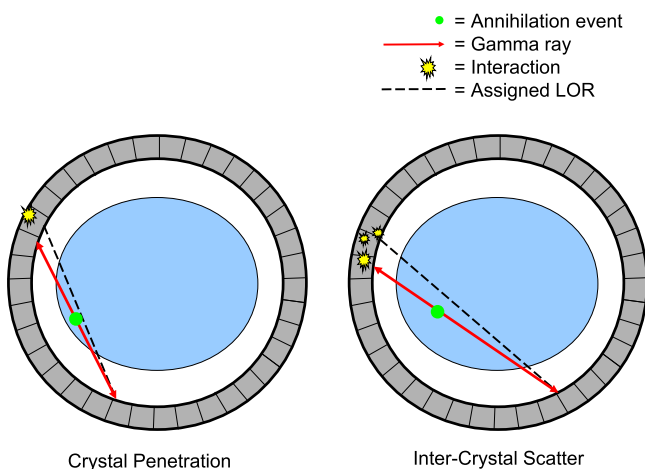
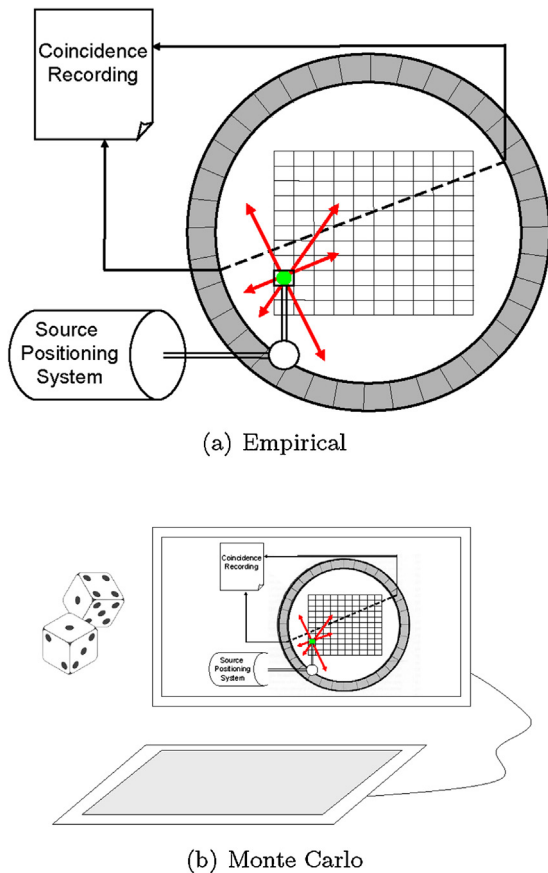


Fig. 6. Detection-related physical effects.



**Fig. 7.** Two of the three fundamental methodologies available for the system matrix estimation. Empirical methodologies (a) place a source at each basis function position in the FOV of the scanner and record the coincidences registered at each LOR. Alternatively, this process can be simulated in a computer, using Monte Carlo stochastic methods (b).

to accomplish the computation of the system matrix. The fourth subsection describes how these methodologies can be combined to form practical hybrid approaches.

#### 4.1. Empirical methodologies

The most accurate way of determining the system matrix would be, under ideal experimental conditions, its direct measurement from the scanner for which data are going to be reconstructed. The elements of the system matrix could be obtained by positioning and scanning a source at every basis function location, and recording the response (see Fig. 7a). Unfortunately, this is also the most challenging approach. On the one hand, for a full 3D matrix, the huge number of scans required makes this approach burdensome. For example, [90] make a temporal estimation of 2.6 years to complete the system matrix acquisition of a clinical scanner, even after assuming scanner symmetries. On the other hand, the experimental setup can be tedious due to the need to precisely position the source, a task that requires complex and expensive equipment (usually sophisticated robots that can only place one source at a time). These obstacles still make a pure empirical methodology presently impractical. However, as will be discussed in detail below, several practical approaches have been proposed that estimate the whole system matrix from a representative set of point or line source measurements.

#### 4.2. Monte Carlo (MC) methodologies

A variety of software packages offer the possibility of simulating the physics of a PET scanner very precisely. These tools rely on Monte Carlo algorithms based on the repeated random sampling of probability models of the transport of radiation through matter. A number of different PET simulators have been developed (PET-SIM [91], SIMSET: Simulation System for Emission Tomography [92], Eidolon [93], GATE, the GEANT4 Application for Tomographic Emission [94], that encapsulates the accurate GEANT libraries for the simulation of the passage of particles through matter [95], GRAY [96], PENELOPE [97], SORTEO, Simulation Of Realistic Tridimensional Emitting Objects, [98]), each of them with different advantages, disadvantages and levels of accuracy. Overviews of these packages can be found in [99,100], and an experimental comparison between two popular PET MC simulators, SIMSET and GEANT, is performed in [101].

An alternative to the direct measurement of the system matrix is the MC simulation of the above mentioned source scanning process (see Fig. 7b). The feasibility and benefits of this approach have been shown by a number of authors [102–107,101,108,109,21,110,22]. Nevertheless, the simulated acquisition of the system matrix has some issues, for which a variety of different solutions have been proposed:

##### 4.2.1. Accuracy

Although current Monte Carlo simulators provide very accurate modeling of the photon transport in the object being imaged and in the detector crystals, there are still effects in the data acquisition process (light propagation effects, electronic noise, etc.) and features of the devices (light guides, magnetic shields, etc.) that are difficult or impossible to reproduce by the simulators and hence a MC approach will be always less precise than an empirical methodology.

##### 4.2.2. Time

MC simulators are fairly slow and a substantial number of simulations (i.e. decays per simulated source) are required to reduce the statistical noise of the estimated system matrix. If the number of elements of the matrix is too high or not enough computational power is available, the MC based computation of the whole system matrix can be unworkable. Several investigators have dealt with the compromise between time and statistical accuracy in the MC calculation of the system matrix. On the one hand, several authors [111,101,110]) have studied the effect of the number of simulated decays on the noise propagation from the matrix to the images. Other works try to accelerate the MC simulation without jeopardizing the accuracy of the system model by considering symmetries [111] or alternative radioactivity distribution models [112]. In [113], different denoising filters are applied to the elements of a short-time simulated system matrix.

##### 4.2.3. Storage

The computational time demand that the statistical accuracy imposes on the simulation of the system matrix often forces MC approaches to precompute the system matrix and store it to be used during reconstruction. The size of the system matrix increases proportionally with the number of basis elements and the number of LORs. In high resolution 3D PET scanners, these numbers can lead to matrix sizes of the order of  $10^{12}$  elements. Under these circumstances, the computation of the system matrix poses a challenge from the storage viewpoint. Several strategies have been proposed in order to save space, which are not strictly restricted to the MC methodologies but are inherent to any approach in which the system matrix values are precalculated offline and stored. These works share two common tactics to reduce

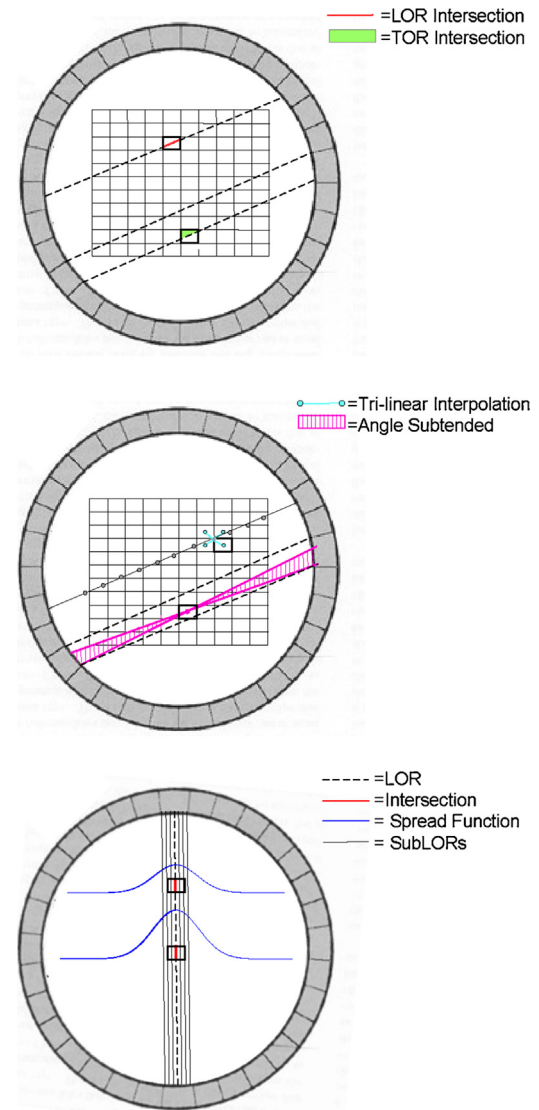
the storage requirements, mainly exploiting the sparsity and the symmetries of the system matrix. Additional compression can be obtained resorting to quasi-symmetries [43] and shift invariances [114].

On the other hand, alternatives to the explicit computation of the matrix that keep the accuracy provided by the simulation tools exist. These works use fast voxel-driven [115] or LOR-driven [116] stochastic sampling protocols to implement the projection operations on-the-fly. But in addition to the pure storage problem, the precomputation once and for all of the system matrix entails other complications:

- It is a rather rigid approach. On the one hand, the configuration of the scanner and the grid of basis functions cannot be modified without running a new set of simulations. From a development perspective (e.g. when various crystal dimensions are to be investigated or when a multiresolution scheme needs to be evaluated) a more flexible approach is desirable. Furthermore, a precalculated system matrix cannot account properly for the medium-related effects (basically positron range and scatter and attenuation in the source), which may vary significantly from one study to another. The estimation of these effects has to be based on general assumptions of the imaged object or alternatively be considered apart, and applied in the form of corrections. In order to overcome this drawback, [110] proposed MC simulations that make use of realistic anatomical maps obtained from computed tomography (CT) images, which allow to include the scatter and attenuation effects in the simulated matrix in a very accurate way.
- Even if the size of the system matrix can be reduced enough to store it as a file, the access to disk for every forward and backward projection operation can considerably slow down the reconstruction. In [21] the elements of the system matrix were stored in an efficient way that allows fast sequential access during the reconstruction.
- The sorting of the simulated data into the matrix is a computation-expensive procedure hindered by the huge size of the matrix. Rafecas et al. [105] employed a database management system to overcome this problem.
- In the case of continuous detectors devices [117–120], the positions  $d$  assigned to the detected photons do not correspond to the centers of discrete crystals, but to continuous locations within scintillator blocks. The resulting  $p(b, d)$  terms are, therefore, continuous functions of  $d$  which should be discretized in order to be stored. The data under reconstruction would have to be rebinned to fit the discretized system matrix terms, losing part of its original precision, which is precisely the key benefit of the continuous scintillators technology.

#### 4.3. Analytical methodologies

Finally, there exist analytical implementations of the system matrix. Unlike the MC approaches, the analytical approaches are noise-free and usually very fast. This normally allows the computation of the elements of the system matrix on-the-fly at reconstruction time, thus avoiding the storage problems mentioned above. However, in its quest to speed up the reconstruction process, these methods overlook important physical effects of the PET acquisition process and use simple geometric system models based on the intersection of the lines [121], or the tubes of response (volumes centered around the LOR, referred to as TORs) [122–125] with voxels or blob basis functions [62,126] (see Fig. 8, top). Certain improved projectors allow to include more basis functions than those strictly involved in the intersections, making use of bilinear and trilinear interpolations [63,127] or the Wu-antialiased line-tracing algorithm [128,129]. An example of



**Fig. 8.** The most basic analytical methodologies (top) approximate the relationship between sources and data with LOR or TOR intersections. More elaborate approaches (middle) calculate the solid angle subtended by the source at the detectors that make up a LOR or use improved projectors. As an example of the latter, the figure illustrates, for the 2D case, a LOR-driven trilinear interpolation projection technique, in which each LOR contributes (in backward projections) and gets contributions (in forward projections) from its four nearest voxels. The basic geometric models can be enhanced with the use of detector responses (bottom). In this example, the simple (red) intersection model would assign identical probabilities of detection in the dotted LOR to the two voxels. This can be solved if the intersections are weighted with the (blue) depth-dependent detector responses. These spread functions can be further exploited, for example, by averaging the weighted intersections of an array of subLORs (black rays joining the detectors). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

trilinear interpolation is shown in Fig. 8 (middle) to illustrate this kind of procedures. [130,131] propose more sophisticated ray-tracers incorporating a linear attenuation model to a basic LOR-basis intersection scheme. The approaches based on the calculation of the solid angles subtended at the detectors by each basis function (Fig. 8, middle) [132,133,29,55,134,135] can take into account the depth dependence (i.e. the distance of the positron annihilation point from the detectors). However, the complexity of the on-the-fly solid angle calculations can compromise the reconstruction time and some authors choose the precomputation of the matrix [135].



#### 4.4. Hybrid approaches

A variety of articles [136,137,87,138–140,114] show how the enhancement of simple analytical models with detector response weights (accounting for the detection-associated effects, mainly intercrystal penetration and scatter) may improve the quality of the results without significantly compromising the efficiency of the calculations. Fig. 8 (bottom) illustrates the principles in which these procedures are based on. All these approaches start with a simple geometrical calculation (line–voxel intersection or distance) to which a spread function approximating the additional effects is imposed. Although all these works propose analytical detector responses, [137] includes as well a MC estimation and [87] states that their kernels should ultimately be based on measured line spread functions. These two authors point to an interesting direction: the benefits that can be obtained exploiting synergies between the basic methodologies for the system matrix computation. A large number of hybrid approaches have been proposed following this philosophy, always seeking for higher precision models at affordable storage and/or time requirements.

For example, a number of works exist in which the system matrix is calculated through a combination of parametrized functions representing the scanner response and the interpolation or extrapolation of a limited number of source acquisitions. Fig. 9 illustrates the main procedures to which these works resort. In some cases, the acquisitions are identified directly with the model [141–143], and the parametrization is performed only in the projection or image space variables, allowing for an efficient size reduction, for example, by grouping LORs by their incident angles. In other cases, the models are parametrized as well, and the acquisitions serve to extract the parameters that define preestablished analytical responses. A great variety of such models has been used, including circular [144] and elliptical [145,146] Gaussians, combination of half Gaussians [147,144,90,148], superposition of symmetrical Gaussian functions [147,90] or convolution of Gaussians with rectangular pulses [149], which sometimes are different for the transaxial and axial components [147,90,144] or for opposite or adjacent detector blocks [149].

The source acquisitions can be simulated [145,149,146,142] or measured [141,147,90,144,148,143]. MC simulations provide the origin of positron emission exactly, whereas in experimental measurements, this information is only available when using point sources. This has the advantage of allowing to perform the whole data collection through a single uniform source simulation study. However, only [142] takes advantage of this issue, and only simulates one cylindrical source. In the rest of the cases, either simulated or measured point sources (or lines in the case of [141]), with representative locations strategically chosen to minimize the number of acquisitions, are used. Furthermore, due to the lack of certain components of the detection process in the simulators, several discrepancies of the simulated responses have been observed with respect to the, always more accurate, measured approaches [150].

Alternatively, the use of parameterized responses can be used to reduce the storage requirements rather than the number of measurements (which in MC estimations are not such a critical factor). In this spirit, the works of [43,151] perform a preliminary MC calculation once, and then approximate the results are by spline profiles whose coefficients fit in RAM memory. Fig. 10 shows the operating principles of these methods. The coefficients can be retrieved during reconstruction for a fast calculation of the system matrix terms using interpolation techniques.

Finally, the combination of empirical measurements and MC simulations is exploited in [152], where a MC obtained system matrix is updated toward an ideal system matrix by training the resulting forward projector with an artificial neural network. The inputs are digital images of a phantom and the desired outputs

are the corresponding projections obtained in a PET scanner, from which the neural network can learn physical effects that might not be described correctly using MC simulations.

### 5. Modeling of individual effects and their combination into the system model

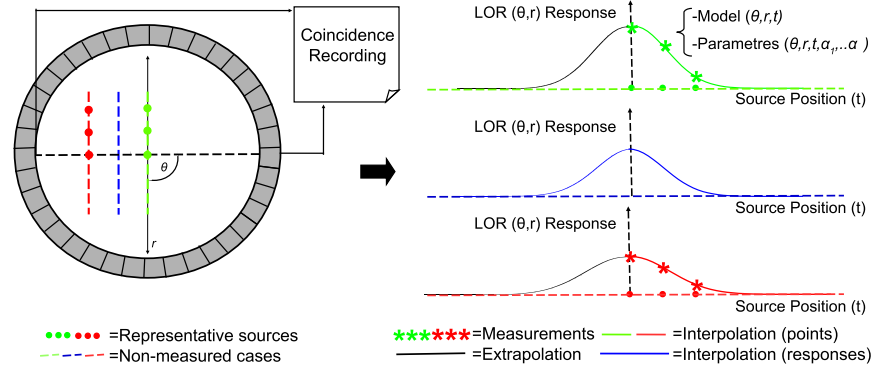
The most general of the hybrid models introduced in the previous section are those models whose calculation involves each of the three basic methodologies (empirical, Monte Carlo and analytical). This is commonly achieved by modeling each effect that occurs in the data stream of a PET scanner separately. In the following, we introduce the common principles of the great variety of approaches that rely on the separation of effects to compute the system matrix (Section 5.1). Then, as shown in Fig. 11, we classify these approaches in two groups, according to the way they group the separated models for the different effects. On the one hand, there are schemes based on the factorization of matrices (Section 5.2). These well-established schemes can in turn be classified in three subgroups (see Sections 5.2.1, 5.2.2 and 5.2.3) according to the emphasis they put in the matrices that account for the resolution “blurring” factors. In Section 5.2.4, we point out how such matrices can be used in reconstruction frameworks different from the factorization of effects. On the other hand, there are approaches that combine the effects through multi integral schemes. As pointed out in Section 5.3, the development of these methods is hindered by the high computational times such integrals impose.

#### 5.1. Effects separation

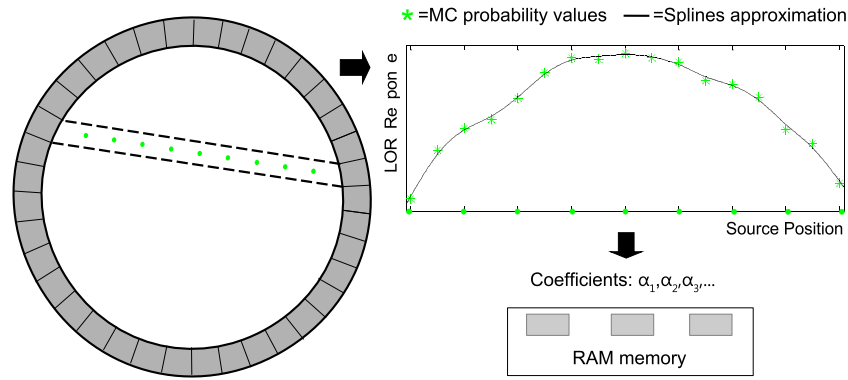
The benefits of the effects separation rely on explicitly considering the diversity of the phenomena occurring during the PET acquisition process, for which an homogeneous treatment may result in non-optimal calculations. By modeling each effect using the methodology that best fits to its requirements, these approaches achieve the best synergy between methodologies. For example, it is not worth to perform extensive experimental measurements or simulations to model the response of effects for which accepted analytical fast noise-free models are available. This is clearly the case of the geometrical factors, but also of other physical effects, such as the photon non-collinearity deviation angle, which is known to have a Gaussian distribution [153,154]. However, for those effects for which a model can hardly be expressed using finite analytical calculations an experimental estimation can be used instead. This can be the case for phenomena such as the photon interaction physics within the scintillator. While in this case an empirical model should provide more accurate results, the complexity of the measurement process can lead to the preference for a MC estimation.

For certain phenomena, the choice of methodology to estimate a model still remains an open problem. This is the case for the positron range, for which, as stated by [155] a considerable variation has been found among the analytical [156,157], simulated [158,159,155] and measured [160,161] models accounting for the complicated path that positrons follow in tissues.

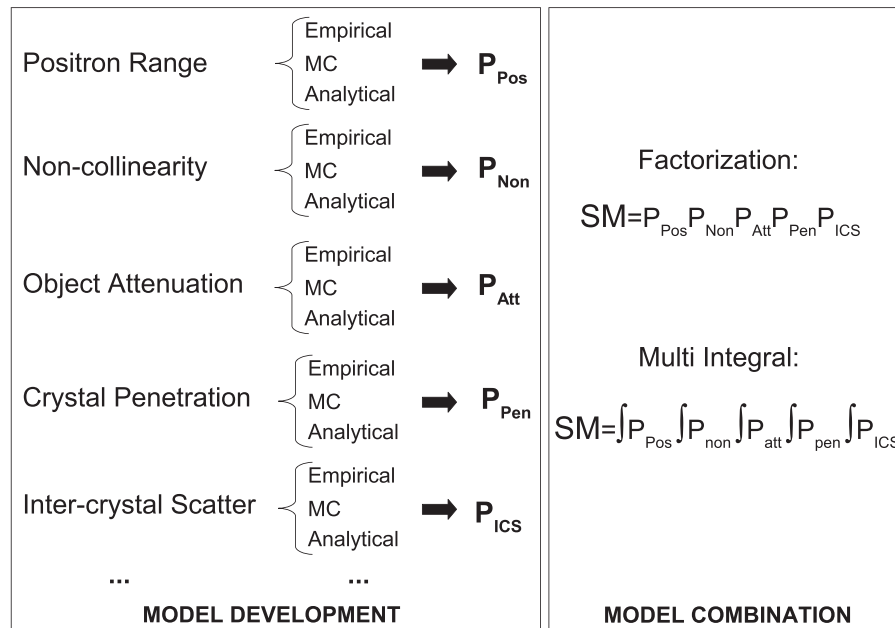
The degree of separation between the different effects present during the acquisition process may vary as well among the different proposals. For example, certain works merge the effects of crystal penetration and inter-crystal scatter in a single detector model, while others account for them through two different models obtained through two different methodologies. In any case, since the model used to describe each effect, or collection of effects, may be changed from one reconstruction to another without altering the rest of the models, the separated-effects approaches enjoy a great flexibility in terms of scanner configurations (geometry, material,



**Fig. 9.** Summary of the strategies used by the class of hybrid approaches that combine parameterized response functions with limited source acquisitions. Measurements are obtained for a collection of representative sources. This information can be stored as the model itself, along with the parameters defining the sources and LOR locations ( $t$  and  $\theta$ ,  $r$  in this case), or fitted to a preestablished analytical model, from which extra parameters ( $\alpha_1 \dots \alpha_N$ ) are extracted. The responses for the non-measured cases can be obtained by means of extrapolation (for example, to a symmetric case, as shown in the figure) or by interpolation, either between points within the responses or between responses themselves. The figure shows interpolations between source positions, but responses can also be interpolated between LORs.



**Fig. 10.** Use of parameterized responses for storage purposes. The MC estimates of probability at points along and/or across (although just the longitudinal case is shown) the region defined by a LOR are approximated by spline profiles, whose coefficients can be stored in RAM.



**Fig. 11.** The hybrid approaches that make use of the three available methodologies (empirical, Monte Carlo and analytical) obtain separated models for the different effects, making use in each case of the best suited approach. Once the separated models have been estimated, they are combined to form the system matrix, either as a product of matrices or through a multi integral scheme.

etc.), experimental conditions (isotope, subject under study) or reconstruction features (type and size of the basis function).

Furthermore, the individual estimation of the different effects allows putting the emphasis of the model on the more degrading factors, according to the PET modality or scanner for which the reconstruction is going to be performed. For example, in a small animal scanner, a greater effort is typically done in the detector response model responsible for the intercrystal penetration and scatter, whereas the human studies are usually more concerned with the attenuation and the scatter in the object. Many authors focus their work on calculating just one or several parts of a well-established scheme of separated effects, while using standard models for the others.

Independently of the effects considered and the methodology used to model each part of the process, a different problem arises when the individually modeled effects have to be combined to form the whole system matrix. Two important different approaches exist to address this issue, namely the factorization and the integral formulation, which are described in the following.

## 5.2. System matrix factorization for resolution modeling

The most popular way of combining models is to express the system matrix as a product of independent matrices, each one standing for one or a collection of effects. This scheme achieves substantial savings in storage, since each matrix component is highly structured and sparse. It is advantageous as well in terms of reconstruction time, since the forward and backward projection can be performed efficiently with a sparse-matrix multiplication operation, although some authors precompute certain matrix products to reduce the computational load, when memory is not a constraint [162].

Although earlier versions of the factorization of effects had been proposed [163,164], and implicit factorizations in the form of the previously mentioned attenuation and normalization multiplicative corrections had been previously applied, nowadays, all the factorized schemes are based on the scheme proposed in the seminal works of [165,55], for 2D and 3D acquisitions respectively:

$$P = P_{\text{det.sens}} P_{\text{blur.proj}} P_{\text{att}} P_{\text{geo}} P_{\text{blur.im}} \quad (1)$$

In a reconstruction in which the continuous radiotracer distribution is approximated as a sum of  $b_j$  ( $j=1, 2 \dots N$ ) discrete basis functions and in which the scanner contains  $d_i$  ( $i=1, 2 \dots M$ ) possible detector pairs,  $P_{\text{det.sens}}$  and  $P_{\text{att}}$  are the  $M \times M$  diagonal Attenuation and Detector Sensitivity matrices. The diagonal elements ( $i, i$ ) of these matrices contain, respectively, the factors that account for the attenuation and detector efficiency in each LOR  $d_i$ . These factors match the attenuation and normalization coefficients used for data correction that have been introduced in Section 3.3 and can be thus calculated using one of the cited methodologies.

$P_{\text{geo}}$  is the  $N \times M$  geometric projection matrix with each element ( $i, j$ ) equal to the probability that a photon pair emitted in a basis function  $b_j$  reaches the front faces of the detector pair  $d_i$  in the absence of attenuation and assuming perfect photon-pair collinearity. In principle, the non-collinearity should be a geometrical effect as it concerns the direction of the gamma rays. However, all the approaches assume this factor is depth-independent and lump it with the blurring matrices that will be described next, based on the fact that it is a phenomenon affecting the resolution of the images. The computation of the geometric projection matrix is normally based on basic intersection or solid angle approaches, with some authors averaging several points within the voxels [55,166] or multiple rays within the LORs [167–169] to improve the accuracy of the model. Other works focus as well on compacting the storage of the

geometric projection matrix [55,166,170] or accelerate its calculation [171], since it can still be large in high-resolution applications.

Although different implementations of the three above described matrices exist, their physical meaning is always the same in all factored schemes. However, for the two remaining matrices, namely the Projection Blurring matrix  $P_{\text{blur.proj}}$  and the image blurring matrix  $P_{\text{blur.im}}$ , the different implementations may imply remarkable conceptual differences. These matrices are intended to account for the factors that limit the spatial resolution, typically the positron range, non-collinearity, crystal size and mispositioning of photons caused mainly by crystal penetration and multiple photon interactions, etc. Each element ( $i, i'$ ) in  $P_{\text{blur.proj}}$  represents the probability that an event which should have been detected along LOR  $d_i$  is actually detected in LOR  $d_{i'}$  (as shown in Fig. 12, left) whereas each element ( $j, j'$ ) in  $P_{\text{blur.im}}$  stands for the probability that a decay originated within a basis function  $b_j$  is placed in another basis function  $b_{j'}$  (as shown in Fig. 12, right). According to their definitions, the  $P_{\text{blur.proj}}$  and  $P_{\text{blur.im}}$  matrices account, respectively, for the overall detection (i.e data-space) and emission (i.e image-space) blurring. Three categories of approaches (differentiated by the use of these matrices) exist, as described in the following subsections.

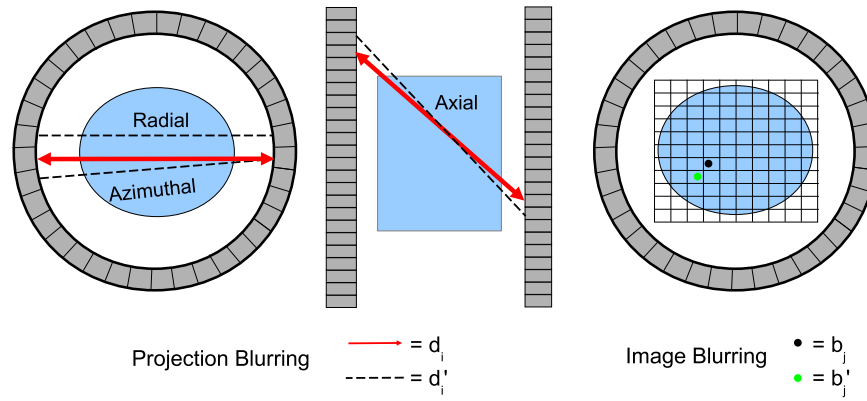
### 5.2.1. Factorization for resolution modeling using both projection and image blurring

Surprisingly, just a small number of recent works apply strictly the definition of the matrices putting the effects in the  $P_{\text{blur.proj}}$  or  $P_{\text{blur.im}}$  matrices, according to the level (data or image) to which they belong. For example, the positron range is best described acting on the image since the detector actually sees the true image as blurred by its effect. In contrast, the non-collinearity, crystal size, penetration and intercrystal scatter are best described as acting on the localization of the LORs. Following these guidelines, the approach of [172] consists of estimating separated analytical models for the non-collinearity, penetration and positron range and a measured model for the inter-crystal scatter, and then placing each of these models in its corresponding Projection or image blurring matrix. The method developed by Zhou and Qi [170] is more general in terms of the effects modeled, as it estimates the two blurring matrices using a limited number of point-source scans. This allows to account not just for certain predefined mentioned effects, but for any other resolution degradation factor that takes place during the source acquisitions.

### 5.2.2. Factorization for resolution modeling with emphasis on projection blurring matrix

A number of authors ignore the image blurring matrix and assume that the resolution spread happens only in the space of the acquired data. The main argument for this approach is related to physics. In applications with  $^{18}\text{F}$  the positron range is sub-millimeter and the majority of blurring effects (mainly parallax and intercrystal scatter) are related to the detection events. Therefore, it seems *a priori* more logical to spend more effort in modeling the distortion of the data. Many successful system matrix implementations are based on this claim, differing just in the way the projection blurring is estimated. In some cases this matrix is estimated from source measurements obtained from MC simulations [165,55,173–176,169] whereas in other cases it is estimated empirically [173,177], which allows to account for all the classical detection blurring effects, and even the emission associated effect of the positron range [162].

The projection-based factorization is a well-established approach that has led to high quality efficient reconstructions, but it has two weak points. First, the correct measurement of the projection blurring would require a collimated source, but this is impractical and uncollimated sources are used instead [177]. Alessio et al. [162] demonstrate with MC simulations that



**Fig. 12.** Definition of the projection (left) and image (right) blurring matrices. The  $(i, i')$  element of the projection blurring matrix contains the blurring contribution from LOR  $d_i$  to LOR  $d_{i'}$ . The projection blurring can be classified into radial, azimuthal and axial blurring, according to the relative position of  $d_{i'}$  with respect to  $d_i$ . The  $(j, j')$  element of the image blurring matrix contains the blurring contribution from basis function  $b_j$  to basis function  $b_{j'}$ .

a non-collimated source is acceptable for modeling the radial blurring (i.e. blurring among LORs in different radial positions), but that it ignores the azimuthal blurring (i.e. blurring among LORs with different angular orientations), since all angles are simultaneously collected during acquisition. Second, and more inconvenient, even if the size saving from factoring the effects is remarkable, the Projection Blurring matrix is still a non-stationary 4D matrix, whose large size makes its storage challenging unless simplifications are introduced. For example, the sparsity of the matrix is increased by ignoring the axial component of the blurring [165,55,173,177], or truncating the kernels (i.e. considering that a LOR  $d_i$  blurs just into a few neighboring elements  $d_{i'}$ ) [165,169].

The dimensionality of the sinogram blurring matrix is reduced as well by assuming certain spatial invariances. Specifically, some approaches assume that the blurring remains the same for all azimuthal angles and only obtain the matrix terms for a single projection angle of  $d_i$  [165,55,173,174,174]. Other approaches assume as well that the blurring terms are identical for all the ring differences (the axial difference between the two crystals that define the LOR  $d_i$ ) and ring planes (i.e. the relative position of the transaxial plane containing the detectors in the scanner). Such approaches apply the projection blurring obtained for the LORs  $d_i$  at the central transaxial plane (i.e. at 0 mm from the axial center of the scanner) to all other rings [165,55,174].

### 5.2.3. Factorization for resolution modeling with emphasis on image blurring matrix

A third alternative arises in response to the inconveniences of the projection-based approaches. The so-called image-space approaches attempt to incorporate all the blurring effects in the image blurring matrix, based on the idea that the reconstructed image may be seen as the blurring of the true image by a point spread function (PSF) [178]. As stated by [168], resolution modeling in image space is more than only incorporating the effect of the positron range, because the procedures used to estimate the PSFs usually account automatically for the detector resolution and other blurring factors. The image-based approaches have not been extensively explored prior to the recent years, but have considerably evolved in their relatively short life. The simplest approaches represent the PSF as a set of shift-invariant Gaussian kernels, with deviations heuristically estimated [179,180,178,181–184]. Added refinements have been subsequently introduced to these basic kernels, like shift variances [185,186,168,187,188], anisotropy [185,186,168,187,188], the use of more complex functions, like mixtures of Gaussians and exponential [186,189] or Pearson distributions [168], and sophisticated procedures for the estimation of the corresponding spread parameters based on measurements

of points [189] or arrays of point sources [188]. In [171] the image blurring matrix is obtained empirically, it is compatible for Time of Flight PET devices and it can correct the sampling artifacts of the geometrical projector. Kotasidis et al. [190] adds the dependency on progressively increasing count rate statistics to the PSF model introduced in [188].

The main benefit of the image approaches is that they are computationally more efficient. Even if the image blurring matrices have been proven to be also non-stationary and asymmetrical, the image space is three dimensional, in contrast to the four dimensional projection domain. Furthermore, the image-space approach is attractive because it can easily be applied to list-mode MLEM reconstruction.

Since the major factors affecting the spatial resolution occur in the projection space, and with the image-based approaches only approximating these effects, the price for the gained efficiency and versatility could be expected to be the outperformance of projection-based approaches over image-based approaches. Surprisingly, the preliminary comparisons [189,188] between both approaches demonstrate very similar performance. As pointed out by Kotasidis et al. [188] the outcome of such comparisons might be strongly dependent on the specific implementations characteristics of the reconstruction. Therefore, it remains to be demonstrated if the image-based approach is a valid alternative to the projection-based approach until a reliable method to compare these two approaches is well established.

### 5.2.4. Alternative uses of the blurring matrices

At this point, and although the topic is out of the scope of this paper, it is interesting to note that the projection and image-based matrices that characterize the blurring processes in the factored approaches can be used in a variety of ways different from their incorporation to the system matrix. Several works illustrate their use to rebin the sinograms [104], deblur the registered data [191,80,192–194], or to deconvolve the images reconstructed without resolution modeling [186,195–197,168]. In other cases, the kernels are applied between iterations, blurring the forward projection with the projection kernel [198,199] or deconvolving the updated images with the image kernel [200,184]. These approaches are not restricted to the iterative algorithms but can be applied as well to FBP [192] or ART [198] to compensate for the blurring processes. In other cases, novel iterative schemes are developed, in which the image updates are alternated with deblurring update steps [201,202].

Although the results in terms of resolution have been shown to be close, in shorter times, to those obtained when the models are included in the system matrix, several authors have observed



amplified levels of the statistical noise in the emission data [192] and the background noise in the images [168]. Among the possible reasons for this issue, Rahmim et al. [127] prove the fact that while system matrix modeling is able to improve both resolution and noise characteristics of the reconstructed images, the deconvolution schemes have their tradeoffs. Spinelli et al. [203] propose the use of an anatomical prior in order to reduce the amount of noise introduced when the forward projection is blurred with a projection space kernel.

### 5.3. Alternative to factorization: multivariate integral formulation

As an alternative to the matrix factorization, several works exist in which the independent models are combined in a scheme of multivariate integrals [204–208,88,69,209]. In these types of schemes, the models for the different effects are approximated by probability distributions. A uniform distribution is normally used for the angular orientation of the emitted photon pairs, which determines the geometrical factor. The crystal penetration is approximated by an exponential distribution representing the probability of interaction of 511 keV photons with the crystals, although [208] demonstrates that a uniform distribution works with certain scintillator materials. For the inter-crystal scatter, some authors use Gaussian distributions [208,69] whereas others perform MC simulations to estimate *ad hoc* distributions [88,172]. A good agreement with fully MC based calculations using less computation has been shown [206–208,88,69] by these approaches.

However, in the integral schemes the separation of effects loses the sparsity of their factored counterparts, which makes the storage of the system matrix an impractical task. Furthermore, the resulting integrals are too complex to give place to closed expressions and have to be solved numerically, which makes the on-the-fly calculation of the matrix a computational challenge. In order to speed up the computation, these schemes typically ignore important effects. In several cases [205–207,209] the effect of inter-crystal scatter is ignored, and none of the integral proposals explicitly accounts for the positron range, non-collinearity, detector sensitivity and object attenuation, although [88,207] include the last two in the form of multiplicative corrections. Furthermore, even after these simplifications, only [207,88,209] get to incorporate their models in a reconstruction algorithm, and their implementations are based on several assumptions and approximations of the formulas, that go beyond the unavoidable numerical integration approximations.

## 6. Topics of current research

As stated for the above-mentioned integral approaches, a point has been reached where there exist such complex schemes for the system matrix computation, that they are not yet ready to be implemented for reconstruction (although can be excellent development tools for new system designs). In these circumstances, the focus of the PET system matrix literature is broadening its scope from the development of new sophisticated calculation methods to other important related issues. In this regard, two recurrent themes can be found nowadays in the field of system models for iterative reconstruction, which are addressed in the following subsections.

### 6.1. Efficient implementations of iterative reconstruction with system modeling

The first current research topic deals with the efficiency of the iterative reconstructions with accurate system models. Concerning the reduction of the long reconstruction times, these works rely mainly on the technological advances that provide increasingly greater computational power. The hardware acceleration of

the iterative reconstruction is not new, and dates back to 1987, when [210] implemented the MLEM algorithm in a Cray-I supercomputer. Afterward, implementations of the reconstructions that can run on several CPUs with lower cost using protocols like the message passing interface (MPI) or OpenMP were preferred [132,211,130,212–215], given that the iterative schemes are naturally suited for parallelization and can benefit from the advantages of distributed computing. But nowadays the most appealing solution to speed up the iterative reconstructions is the graphics processing unit (GPU)-based implementation of the algorithms. Although GPUs were initially developed to provide high-definition graphics for video games in real time (at a very low cost due to their sheer commercial volume), recent improvements in the computing power and programability have favored their use as cost effective high performance coprocessors in time-consuming scientific applications. This includes the acceleration of iterative reconstruction with accurate system matrix for PET.

In the earlier GPU-implementations the representations of the system responses were not the most sophisticated. In ascending order of complexity, these representations include intersections enhanced with linear interpolation [216], voxel-line distances weighted with Gaussian detector responses [217] and shift invariant [218] or variant [219] Gaussian kernels. However, it has been sufficient to demonstrate the GPU possibilities and has encouraged several groups to perform GPU-reconstructions with system models previously exploited in conventional CPUs, showing remarkable speed-ups. For example, [220–222] propose GPU-accelerated implementations of the MOLAR, FIRST and DIRECT reconstruction packages (and the system models therein) proposed respectively in [87,43,223]. [170,224] implement MLEM and MAP with factorized system matrices and [225] does likewise with the integral model of [204] and the OSEM algorithm.

The success in the acceleration of well-established schemes has led to the development of new forms of iterative reconstruction specially suited to GPU implementations, in order to achieve further speed-up. In [226], the MLEM algorithm is reformulated so that it can scale up to many GPU nodes with less frequent inter-node communication, which demonstrates a greater effectiveness when executed on a multi-GPU cluster. In [227], a new system model specially suited for GPU implementation is developed. In this work, detector response functions for each LOR are precalculated in CPUs and modeled by polynomials whose coefficients are stored in the GPU and used to calculate the elements of the matrix on-the-fly. This approximation reduces the number of conditional statements and the required memory size, which matches GPU implementation better. Although the results do not show remarkable speedup with respect with a non-GPU-matched (Gaussian) model, the extent to which the efficiency of acceleration depends on the manner of implementation remains an open question, which will likely give scope for future research.

In addition to being an excellent tool for the reduction of the reconstruction times, the hardware acceleration has been exploited to improve the performance of the MC packages [228,229]. This allows not only to improve the tradeoff between time and statistical noise of the simulated models but has even enabled the development of new MC-based methodologies for the scatter correction [230,231].

The storage reduction is the second battlefront of the system modeling for iterative reconstruction in the context of the efficiency. The high resolution applications impose an increasingly fine discretization of the image and projection spaces, which translates into an enormous number of basis functions and LORs. These numbers lead to prohibitive dimensions of the system matrix, incompatible with its RAM allocation. As has been mentioned, the use of symmetries is a valuable technique to reduce the memory for storing the system matrix, and has been well studied for

cubic voxels in a Cartesian grid, for which the maximum obtainable symmetry is 4-fold in the transaxial plane and 2-fold in the axial dimension. The new strategies to reduce storage requirements point to the exploitation of further symmetries, based on the use of alternatives either to the conventional voxels or to the rectangular assemblies. Concerning the former strategy, the use of polar basis functions is not a new approach [210], but has been recently taken up again by several authors [232–235] in order to explore their potential advantages for fully exploiting the rotational symmetry of cylindrical scanners. As for the latter, in [236], the object grids are rotated by a Gaussian rotator in order to have the same cylindrical symmetry as the scanner, which reduces the system matrix storage by a factor proportional to the number of scanner blocks. Alternatively, [124] propose a new organization of projection data combined with highly rotation-symmetric voxel assemblies, reporting factors of reduction even higher than the transaxial block symmetry of the system.

The system matrix dimensions are even larger for scanners with depth of interaction (DOI) measurement, in which the number of crystal pairs increases in proportion to the square of the number of DOI layers. For this kind of scanners, [237] present a methodology for achieving very high rates of system matrix compression, using the concept of virtual detector ring (which permits to represent multiple layers of detector elements in the radial dimension as a single detector volume).

## 6.2. Performance characteristics of iterative reconstruction with system modeling

The second current research topic is concerned with assessment of the advantages and limitations of the system modeling, and with how they depend on the different metrics and image contexts. In principle, a proper system modeling is assumed to visually improve the images, providing higher resolution for small or narrow structures, as corroborated by the traditional quality metrics. It is questionable, however, if this will make a real difference in the daily performance of PET-based studies, which will be worth the cost of replacing a operating model by a likely more complex alternative. Indeed, some authors argue that these enhancements are deceptive and that they lead to degradation elsewhere. Excellent discussions on the importance of analyzing the statistical properties of the various quantitative metrics, and of properly assessing in which conditions the use of system modeling is beneficial can be found in [238,239]. There are three recurrent topics that these and other studies within this framework address.

First, some papers address the features that are affected by the resolution modeling in a less straightforward way than the improved image resolution. For example, [240] evaluates the impact of resolution modeling on the contrast recovery, an issue of the reconstructions which has been less studied than resolution. And above all, the effect on the noise structure, which has been observed to diminish voxel variance while increasing inter-voxel covariance [189,127]. [241] provides formulations for how changes in variance and covariance are expected to impact various measures, and evaluates the noise properties in different combinations of reconstruction methods and parameters. [238,242] propose more thorough noise propagation models in order to accurately characterize the noise texture of reconstructed images in the presence of resolution modeling.

Second, it is also very important to notice that there are unresolved aspects of the iterative reconstruction with system modeling which remain to be fully understood. A collection of authors have recently investigated the causes and possible corrections of an old problem [47] that becomes much more severe with system modeling, that is, the edge artifacts, reminiscent of the Gibbs ringing overshoot at the edges [243–245,238,242]. Along with the

edge-overshoot, [246] considers other artifacts, like hyper-resolution and reduced resolving power.

Finally, there is a collection of works that focus on evaluating the impact of the model on the measurement of specific features involved in real-application reconstructions. These include, for example the quantification of several biological parameters in a clinical study of the dopamine transporter [247], the standardized uptake value of lump node metastasis in FDG-PET/CT [248] or the glucose metabolic rate, perfusion and cardiac output in dynamic mouse studies [249]. More generally, [238,242] analyze the limitations that system modeling encounters in the context of quantitative images, where increased intervoxel correlations can lead to significant loss of precision in small regions of interest.

## 7. Discussion

The success of iterative reconstruction methods in PET relies to a great extent on the quality of the system matrix that characterizes the data acquisition process. A big effort has been made to improve the accuracy and efficiency with which the statistical relationship between images and data is estimated, arranged in the matrix and used during the reconstructions. Such an effort translates into a great number of works that make the computation of the system matrix an important subgenre of the PET reconstruction research. This article reviews many of these works. Our guiding thread is the methodology employed to compute the system matrix, but the cited articles could have been arranged chronologically, according to their relevance or to any other criteria. For example, an interesting classification of various approaches in terms of how they approximate the factored formulation and how they estimate the individual resolution degrading terms is proposed in [238]. In either case, there are several conclusions that can be drawn from the inspection of the system matrix literature, that are described next. The computation of the system matrix entails two main challenges, which to a greater or lesser extent, are addressed by all cited articles. First, the accuracy with which the superposition of a rather heterogeneous collection of physical effects can be modeled. Second, the efficacy with which the model can be used to carry on the forward and backward projections, in terms either of the memory storage resources needed or of the speed with which these operations are performed during the reconstruction.

The attention that has been paid to each of these two issues has fluctuated throughout the years. The earliest works on the topic were founded on simple but efficient LOR-voxel intersection models, which can be seen now as basic proofs of concept that aimed to show the benefits of the iterative algorithms over the analytical approaches. Once the better performance of the iterative reconstruction was demonstrated, the next generation of works focused on improving the accuracy of the models, with efficiency receiving less attention. The finding that small improvements of the geometrical models, like solid angle calculations or the addition of simple invariant detector responses, translated into greatly improved images, probably encouraged the PET community to pursue increasingly complex models by resorting to the capabilities of the Monte Carlo simulators and measurement mechanisms. But as the quality of the models was improved, the extension of calculation time and storage resources started to be significant.

Even if multiple strategies (that have been mentioned throughout the text) have been developed to achieve a balance between performance and practicality, the requirements for computer hardware have certainly imposed a limit in the quest for the exactness of the system response models. Such limit could be in what we have classified as “Multivariate Integral formulation” approaches in Section 5.3. The philosophy of such schemes is to be extremely rigorous in its mathematical formulation. However, as has been

already mentioned, the resulting formulas turn out to be unworkable and their implementations are based on assumptions that in the end distort the rigorous principles. In this regard, the widespread use of other schemes, like the factored approach, which achieve a better compromise between effort and quality, are the proof of what can be considered a golden principle in the system modeling for iterative reconstruction: more benefits can be achieved with fast and straightforward implementations of fairly accurate schemes than with more advanced modeling that cannot be faithfully implemented due to the restrictive cost of computing, and possibly storing, the system matrix. Therefore, the point has been reached where the computational constraints are hindering the progress toward more accurate system matrices.

From this point, the evolution of the statistical modeling for iterative reconstruction cannot be predicted. However, it should be no surprise that the ideal system response model is yet to be developed and that this problem will still be an area of very active research. The direction of this research will likely be conditioned by the outcome of the two issues addressed in Section 6. They will determine the degree of complexity of the models that is compatible with the feasibility of the implementations and has real significance in the performance of the PET studies, which is the ultimate goal of reconstructions. Along with the capacity of accommodation to the new features of the novel devices (e.g. higher resolution inserts that give place to irregular geometries [250,251]), these constraints will likely dictate the future development of the statistical modeling for iterative reconstruction.

## 8. Summary

One of the strengths of iterative algorithms is their ability to model the physical phenomena underlying the data acquisition process (positron range, noncollinearity, attenuation, crystal penetration, inter-crystal scatter, etc.) in the so-called system matrix. The system matrix is a probability matrix that describes the relationship between the object and detection spaces. The number of probability coefficients is large in high resolution PET scanners and their computation is complicated, due to the variety and complexity of the phenomena involved (a description of these phenomena has been provided in Section 5). As a result, the accurate calculation of the system matrix elements without incurring excessive computational cost is a very active subject of research.

In this article we have presented a review of the different approaches used to estimate the system matrix. In declining order of accuracy (and, correspondingly, of complexity) there are three available methodologies to obtain the system matrix: empirical measurements, Monte Carlo simulations and analytical calculations. Section 4 has dealt with the benefits and drawbacks of each of these methodologies.

Many investigators have proposed combinations of the basic methodologies in order to achieve different levels of compromise between the quality of the models and the computational cost of their estimation. The hybrid strategies, which have been described in Section 4.4, include geometrical models weighted with detector response functions, parametrized models estimated from a limited number of source measurements, MC models that use sophisticated analytical strategies of storage or MC models enhanced with neural networks that can learn physical effects from empirical measurements.

Within the group of hybrid approaches, those schemes in which each physical effect is individually modeled with the most suited of the three available methodologies are particularly appealing, as has been discussed in Section 5.1. There are two families of works following this philosophy, namely factorized and integral

formulation schemes, which differ in the way they merge the separately obtained models.

Section 5.2 has dealt with the factorized approaches, in which the system matrix is expressed as an efficient product of matrices, each one containing the model for one or a collection of related effects. The numerous approaches that rely on the factorization of the system matrix can be, in turn, classified into three different categories, according to the way they model the “blurring” (i.e. resolution degrading) effects. Although traditionally the weight of the blurring has been put in the projection domain, in the last decade several authors have claimed the benefits of modeling the blurring in the image-space domain. More recently, authors have tried to distribute the blurring factors between both domains, according to their image or projection-degrading nature. In addition to the importance they are given in the system matrix calculation, the matrices modeling the blurring effects have alternative uses in the sinogram rebinning, data deblurring or image deconvolving processes.

The alternative to the factorization of the models is their combination through schemes of, in principle more rigorous, multivariate integrals, which have been introduced in Section 5.3. However, the costly numerical implementation of the integrals has limited the practical application of these schemes. In a sense, this is representative of the current situation in the field of statistical modeling for iterative reconstruction, i.e. the benefits obtained by improving the accuracy of the current system models may not compensate for the additional computation cost.

Under these circumstances, two topics of research (which have been discussed in Section 6) are becoming increasingly important: improving the efficiency of the reconstructions and properly evaluating the real benefits obtained by using accurate system models. As pointed out in our discussion, the outcome of these two issues will probably determine the evolution of system modeling for PET reconstruction.

## References

- [1] Torizuka T, Tamaki N, Inokuma T, Magata Y, Sasayama S, Yonekura Y, et al. In vivo assessment of glucose metabolism in hepatocellular carcinoma with FDG-PET. *J Nucl Med* 1995;36(10):1811–7.
- [2] Tjuvajev JG, Doubrovina M, Akhurst T, Cai S, Balatoni J, Alauddin MM, et al. Comparison of radiolabeled nucleoside probes (FIAU, FHBG, and FHPG) for PET imaging of HSV1-tk gene expression. *J Nucl Med* 2002;43(8):1072–83.
- [3] Kubota K. From tumor biology to clinical PET: a review of positron emission tomography (PET) in oncology. *Ann Nucl Med* 2001;15(6):471–86.
- [4] Belhocine T, Spaepen K, Dusart M, Castaigne C, Muylle K, Bourgeois P, et al. <sup>18</sup>F-FDG PET in oncology: the best and the worst (Review). *Int J Oncol* 2006;28(5):1249.
- [5] Charnley N, Jones T, Price P. In: Shreve P, Townsend DW, editors. *Principles of PET in cancer treatment for the assessment of chemotherapy and radiotherapy response and for radiotherapy treatment planning*. Springer; 2010.
- [6] Kalnicki S. PET/CT in radiation oncology. *Semin Nucl Med* 2012;42:281–352.
- [7] Love C. Non-Oncology PET (Part I). *Semin Nucl Med* 2013;43:331–402.
- [8] Love C. Non-oncology PET (Part II). *Semin Nucl Med* 2013;43:403–74.
- [9] Nekolla SG, Martinez-Moeller A, Saraste A. PET and MRI in cardiac imaging: from validation studies to integrated applications. *Eur J Nucl Med Mol Imaging* 2009;36(1):121–30.
- [10] Cherry SR, Gambhir SS. Use of positron emission tomography in animal research. *ILAR J* 2001;42(3):219–32.
- [11] Cabeza R, Nyberg L. Imaging cognition: an empirical review of PET studies with normal subjects. *J Cogn Neurosci* 1997;9(1):1–26.
- [12] Cabeza R, Nyberg L. Imaging cognition II: an empirical review of 275 PET and fMRI studies. *J Cogn Neurosci* 2000;12(1):1–47.
- [13] Silverman DH. Brain 18F-FDG PET in the diagnosis of neurodegenerative dementias: comparison with perfusion SPECT and with clinical evaluations lacking nuclear imaging. *J Nucl Med* 2004;45(4):594–607.
- [14] Chatziioannou AF. PET scanners dedicated to molecular imaging of small animal models. *Mol Imaging Biol* 2002;4(1):47–63.
- [15] Koo V, Hamilton P, Williamson K. Non-invasive in vivo imaging in small animal research. *Anal Cell Pathol* 2006;28(4):127–39.
- [16] Gu Z, Taschereau R, Vu N, Wang H, Prout D, Silverman R, et al. NEMA NU-4 performance evaluation of PETbox4, a high sensitivity dedicated PET preclinical tomograph. *Phys Med Biol* 2013;58(11):3791.



- [17] Moses WW. Fundamental limits of spatial resolution in PET. *Nucl Instrum Methods Phys Res A* 2011;648:S236–40.
- [18] Carson RE. Tracer kinetic modeling in PET. In: *Positron Emission Tomography*. Springer; 2005. p. 127–59.
- [19] Lewellen T, Karp J. PET systems. In: Wernick MN, Aarsvold JN, editors. *Emission Tomography. The Fundamentals of PET and SPECT*. Academic Press; 2004. p. 172–9.
- [20] Stickel JR, Cherry SR. High-resolution PET detector design: modelling components of intrinsic spatial resolution. *Phys Med Biol* 2005;50(2):179–95.
- [21] Ortuño JE, Kontaxakis G, Rubio JL, Guerra P, Santos A. Efficient methodologies for system matrix modelling in iterative image reconstruction for rotating high-resolution PET. *Phys Med Biol* 2010;55(7):1833–61.
- [22] Lee NY, Choi Y. Simulation studies on depth of interaction effect correction using a Monte Carlo computed system matrix for brain positron emission tomography. *Comput Methods Prog Biomed* 2012;108(2):820–31.
- [23] Cherry SR, Chatzioannou AF. In: Wernick MN, Aarsvold JN, editors. *Emission Tomography First Edition*. Academic Press; 2004.
- [24] Shepp L, Logan B. The Fourier reconstruction of a head section. *IEEE Trans Nucl Sci* 1974;21(3):21–33.
- [25] Kinahan PE, Rogers JG. Analytic 3-D image-reconstruction using all detected events. *IEEE Trans Nucl Sci* 1989;36(1):964–8.
- [26] Ollinger JM, Fessler JA. Positron-emission tomography. *IEEE Signal Process Mag* 1997;14(1):43–55.
- [27] Lewitt RM, Matej S. Overview of methods for image reconstruction from projections in emission computed tomography. *Proc IEEE* 2003;91(10):1588–611.
- [28] Qi JY, Leahy RM. Iterative reconstruction techniques in emission computed tomography. *Phys Med Biol* 2006;51(15):R541–78.
- [29] Johnson CA, Seidel J, Carson RE, Gandler WR, Sofer A, Green MV, et al. Evaluation of 3D reconstruction algorithms for a small animal PET camera. *IEEE Trans Nucl Sci* 1997;44(3):1303–8.
- [30] Chatzioannou A, Qi J, Moore A, Annala A, Nguyen K, Leahy R, et al. Comparison of 3-D maximum a posteriori and filtered backprojection algorithms for high-resolution animal imaging with microPET. *IEEE Trans Med Imaging* 2000;19(5):507–12.
- [31] Gordon R, Bender R, Herman GT. Algebraic reconstruction techniques (ART) for 3-dimensional electron microscopy and X-ray photography. *J Theor Biol* 1970;29(3):471.
- [32] Kaczmarz S. Angenherzte auflösung von systemen linearer gleichungen. *Bull Int Acad Polon Sci* 1937;29:471–81.
- [33] Vardi Y, Shepp LA, Kaufman L. A statistical-model for positron emission tomography. *J Am Stat Assoc* 1985;80(389):8–20.
- [34] Lewitt RM, Muehllehner G. Accelerated iterative reconstruction for positron emission tomography based on the EM algorithm for maximum likelihood estimation. *IEEE Trans Med Imaging* 1986;5(1):16–22.
- [35] Tanaka E. A fast reconstruction algorithm for stationary positron emission tomography based on a modified EM algorithm. *IEEE Trans Med Imaging* 1987;6(2):98–105.
- [36] Hudson HM, Larkin RS. Accelerated image-reconstruction using ordered subsets of projection data. *IEEE Trans Med Imaging* 1994;13(4):601–9.
- [37] Browne J, DePierro AR. A row-action alternative to the EM algorithm for maximizing likelihoods in emission tomography. *IEEE Trans Med Imaging* 1996;15(5):687–99.
- [38] Byrne CL. Convergent block-iterative algorithms for image reconstruction from inconsistent data. *IEEE Trans Image Process* 1997;6(9):1296–304.
- [39] Fessler JA, Hero AO. Space-alternating generalized expectation-maximization algorithm. *IEEE Trans Image Process* 1994;42(10):2664–77.
- [40] Daube-Witherspoon ME, Muehllehner G. Treatment of axial data in three-dimensional PET. *J Nucl Med* 1987;28(11):1717–24.
- [41] Defrise M, Kinahan PE, Townsend DW, Michel C, Sibomana M, Newport DF. Exact and approximate rebinning algorithms for 3-D PET data. *IEEE Trans Med Imaging* 1997;16(2):145–58.
- [42] Leahy RM, Qi J. Statistical approaches in quantitative positron emission tomography. *Stat Comput* 2000;10(2):147–65.
- [43] Herraiz JL, España S, Vaquero JJ, Desco M, Udías JM. FIRST: fast iterative reconstruction software for (PET) tomography. *Phys Med Biol* 2006;51(18):4547–65.
- [44] Veklerov E, Llacer J. Stopping rule for the MLE algorithm based on statistical hypothesis testing. *IEEE Trans Med Imaging* 1987;6(4):313–9.
- [45] Johnson V. A note on stopping rules in EM-ML reconstructions of ECT images. *IEEE Trans Med Imaging* 1994;13(3):569–71.
- [46] Coakley KJ. A cross-validation procedure for stopping the EM algorithm and deconvolution of neutron depth profiling spectra. *IEEE Trans Nucl Sci* 1991;38(1):9–15.
- [47] Snyder DL, Miller M, Thomas LJ, Polite DG. Noise and edge artifacts in maximum-likelihood reconstructions for emission tomography. *IEEE Trans Med Imaging* 1987;6(3):228–38.
- [48] Liow JS, Strother SC. Practical tradeoffs between noise, quantitation, and number of iterations for maximum likelihood-based reconstructions. *IEEE Trans Med Imaging* 1991;10(4):563–71.
- [49] Slijpen E, Beekman F. Comparison of post-filtering and filtering between iterations for SPECT reconstruction. *IEEE Trans Nucl Sci* 1999;46(6):2233–8.
- [50] Mair BA, Carroll RB, Anderson JMM. Filter banks and the EM algorithm. In: *Nuclear science symposium*. 1996. p. 1747–51.
- [51] Green PJ. Bayesian reconstructions from emission tomography data using a modified EM algorithm. *IEEE Trans Med Imaging* 1990;9(1):84–93.
- [52] Wang G, Qi J. Penalized likelihood PET image reconstruction using patch-based edge-preserving regularization. *IEEE Trans Med Imaging* 2012;31(12):2194–204.
- [53] Rapisarda E, Presotto L, De Bernardi E, Gilardi MC, Bettinardi V. Optimized Bayes variational regularization prior for 3D PET images. *Comput Med Imaging Graph* 2014;38(6):445–57.
- [54] Yang L, Zhou J, Ferrero A, Badawi RD, Qi J. Regularization design in penalized maximum-likelihood image reconstruction for lesion detection in 3D PET. *Phys Med Biol* 2014;59(2):403.
- [55] Qi J, Leahy RM, Cherry SR, Chatzioannou A, Farquhar TH. High-resolution 3D Bayesian image reconstruction using the microPET small-animal scanner. *Phys Med Biol* 1998;43(4):1001–13.
- [56] Tanaka E, Kudo H. Subset-dependent relaxation in block-iterative algorithms for image reconstruction in emission tomography. *Phys Med Biol* 2003;48(10):1405.
- [57] Tanaka E, Kudo H. Optimal relaxation parameters of DRAMA (dynamic RAMLA) aiming at one-pass image reconstruction for 3D-PET. *Phys Med Biol* 2010;55(10):2917.
- [58] Tsui BMW, Zhao XD, Frey EC, Gullberg GT. Comparison between ML-EM and WLS-CG algorithms for SPECT image-reconstruction. *IEEE Trans Nucl Sci* 1991;38(6):1766–72.
- [59] Kaufman L. Maximum-likelihood, least-squares, and penalized least-squares for PET. *IEEE Trans Med Imaging* 1993;12(2):200–14.
- [60] Fessler JA. Penalized weighted least-squares image-reconstruction for positron emission tomography. *IEEE Trans Med Imaging* 1994;13(2):290–300.
- [61] Kadmaras DJ. LOR-OSEM: statistical PET reconstruction from raw line-of-response histograms. *Phys Med Biol* 2004;49(20):4731.
- [62] Hu Z, Wang W, Gualtieri E, Hsieh Y, Karp J, Matej S, et al. An LOR-based fully-3D PET image reconstruction using a blob-basis function. In: *Nuclear science symposium*. 2007. p. 4415–8.
- [63] Reader AJ, Erlandsson K, Flower MA, Ott RJ. Fast accurate iterative reconstruction for low-statistics positron volume imaging. *Phys Med Biol* 1998;43(4):835–46.
- [64] Reader AJ, Manavaki R, Zhao S, Julyan PJ, Hastings DL, Zweit J. Accelerated list-mode EM algorithm. *IEEE Trans Nucl Sci* 2002;49(1):42–9.
- [65] Rahmim A, Lenox M, Reader AJ, Michel C, Burbar Z, Ruth TJ, et al. Statistical list-mode image reconstruction for the high resolution research tomograph. *Phys Med Biol* 2004;49(18):4239–58.
- [66] Matej S, Lewitt RM. Practical considerations for 3-D image reconstruction using spherically symmetric volume elements. *IEEE Trans Med Imaging* 1996;15(1):68–78.
- [67] Levkowitz R, Falikman D, Zibulevsky M, Ben-Tal A, Nemirovski A. The design and implementation of COSEM, an iterative algorithm for fully 3-D listmode data. *IEEE Trans Med Imaging* 2001;20(7):633–42.
- [68] Soares EJ, Germino KW, Glick SJ, Stodilka RZ. Determination of three-dimensional voxel sensitivity for two- and three-header coincidence imaging. *IEEE Trans Nucl Sci* 2003;50(3):405–12.
- [69] Iriarte A, Sorzano COS, Carazo JM, Rubio JL, Marabini R. A theoretical model for EM-ML reconstruction algorithms applied to rotating PET scanners. *Phys Med Biol* 2009;54(7):1909–34.
- [70] Brasse D, Kinahan PE, Lartizien C, Comtat C, Casey M, Michel C. Correction methods for random coincidences in fully 3D whole-body PET: impact on data and image quality. *J Nucl Med* 2005;46(5):859–67.
- [71] Zaidi H. Scatter modelling and correction strategies in fully 3-D PET. *Nucl Med Commun* 2001;22(11):1181–4.
- [72] Watson CC, Casey ME, Michel C, Bendriem B. Advances in scatter correction for 3D PET/CT. In: *Nuclear science symposium*. 2004. p. 3008–12.
- [73] Zaidi H, Koral KF. Scatter modelling and compensation in emission tomography. *Eur J Nucl Med Mol Imaging* 2004;31(5):761–82.
- [74] Lange K, Bahn M, Little R. A theoretical study of some maximum likelihood algorithms for emission and transmission tomography. *IEEE Trans Med Imaging* 1987;6(2):106–14.
- [75] Mumcuoglu EU, Leahy R, Cherry SR, Zhou Z. Fast gradient-based methods for Bayesian reconstruction of transmission and emission PET images. *IEEE Trans Med Imaging* 1994;13(4):687–701.
- [76] Fessler JA, Ficaro EP, Clinthorne NH, Lange K. Grouped-coordinate ascent algorithms for penalized-likelihood transmission image reconstruction. *IEEE Trans Med Imaging* 1997;16(2):166–75.
- [77] Kinahan P, Townsend D, Beyer T, Sashin D. Attenuation correction for a combined 3D PET/CT scanner. *Med Phys* 1998;25(10):2046–53.
- [78] Chow PL, Rannou FR, Chatzioannou AF. Attenuation correction for small animal PET tomographs. *Phys Med Biol* 2005;50(8):1837.
- [79] Defrise M, Townsend D, Bailey D, Geissbuhler A, Jones T. A normalization technique for 3D PET data. *Phys Med Biol* 1991;36(7):939.
- [80] Hoffman EJ, Guerrero TM, Germano G, Digby W, Dahlbom M. PET system calibrations and corrections for quantitative and spatially accurate images. *IEEE Trans Nucl Sci* 1989;36(1):1108–12.
- [81] Casey ME, Gadagkar H, Newport D. A component based method for normalization in volume PET. In: Grangeat P, Amans JL, editors. *Three-dimensional image reconstruction in radiology and nuclear medicine*. Kluwer Academic; 1996. p. 66–71.
- [82] Badawi RD, Marsden PK. Developments in component-based normalization for 3D PET. *Phys Med Biol* 1999;44(2):571–94.



- [83] Bai B, Li Q, Holdsworth CH, Asma E, Tai YC, Chatzioannou A, et al. Model-based normalization for iterative 3D PET image reconstruction. *Phys Med Biol* 2002;47(15):2773–84.
- [84] Lange K, Carson R. EM reconstruction algorithms for emission and transmission tomography. *J Comput Assist Tomogr* 1984;8(2):306–16.
- [85] Rehfeld N, Alber M. A parallelizable compression scheme for Monte Carlo scatter system matrices in PET image reconstruction. *Phys Med Biol* 2007;52(12):3421.
- [86] Cheng JCK, Laforest R, OSullivan JA. A scatter and randoms weighted (SRW) iterative PET reconstruction. *Med Phys* 2011;38(6):3186–92.
- [87] Carson RE, Barker WC, Liow JS, Johnson CA. Design of a motion-compensation OSEM list-mode algorithm for resolution-recovery reconstruction for the HRRT. In: Nuclear science symposium conference. 2003. p. 3281–5.
- [88] Moehrs S, Defrise M, Belcari N, Guerra AD, Bartoli A, Fabbri S, et al. Multi-ray-based system matrix generation for 3D PET reconstruction. *Phys Med Biol* 2008;53(23):6925–45.
- [89] Rapisarda E, Bettinardi V, Thielemans K, Gilardi MC. Evaluation of a new regularization prior for 3-D PET reconstruction including PSF modeling. *IEEE Trans Nucl Sci* 2012;59(1):88–101.
- [90] Panin VY, Kehren F, Michel C, Casey M. Fully 3-D PET reconstruction with system matrix derived from point source measurements. *IEEE Trans Med Imaging* 2006;25(7):907–21.
- [91] Thompson C, Moreno-Cantu J, Picard Y. PETSIM: Monte Carlo simulation of all sensitivity and resolution parameters of cylindrical positron imaging systems. *Phys Med Biol* 1992;37(3):731.
- [92] Harrison R, Vannoy S, Haynor D, Gillispie S, Kaplan M, Lewellen T. Preliminary experience with the photon history generator module of a public-domain simulation system for emission tomography. In: Nuclear science symposium and medical imaging conference. 1993. p. 1154–8.
- [93] Zaidi H, Morel C. Eidolon: a Monte Carlo simulator for multi ring 3D PET scanners. Switzerland: PARAPET, Geneva University Hospital Division of Nuclear Medicine; 1999.
- [94] Jan S, Santin G, Strul D, Staelens S, Assie K, Autret D, et al. GATE: a simulation toolkit for PET and SPECT. *Phys Med Biol* 2004;49(19):4543–61.
- [95] Agostinelli S, Allison J, Amako K, Apostolakis J, Araujo H, Arce P, et al. Geant4: a simulation toolkit. *Nucl Instrum Methods Phys Res* 2003;506(3):250–303.
- [96] Olcott PD, Buss SR, Levin CS, Prax G, Sramek C. GRAY: high energy photon ray tracer for PET applications. In: Nuclear science symposium. 2006. p. 2011–5.
- [97] Espina S, Herraiz JL, Vicente E, Vaquero JJ, Desco M, Udias JM. PenelopePET, a Monte Carlo PET simulation tool based on PENELOPE: features and validation. *Phys Med Biol* 2009;54(6):1723–42.
- [98] McLennan A, Reilhac A, Brady M. SORTEO: Monte Carlo-based simulator with list-mode capabilities. In: Engineering in Medicine and Biology Society. 2009. p. 3751–4.
- [99] Zaidi H. Relevance of accurate Monte Carlo modeling in nuclear medical imaging. *Med Phys* 1999;26(4):574–608.
- [100] Buvat I, Castiglioni I, Feuardent J, Gilardi MC. Unified description and validation of Monte Carlo simulators in PET. *Phys Med Biol* 2005;50(2):329–46.
- [101] Shokouhi S, Vaska P, Southekal S, Schlyer D, Purschke M, Dzordzhadze V, et al. Statistical 3D image reconstruction for the RatCAP PET tomograph using a physically accurate, Monte Carlo based system matrix. In: Nuclear science symposium. 2004. p. 3901–5.
- [102] Keller NA, Lupton LR. PET detector ring aperture function calculations using Monte Carlo techniques. *IEEE Trans Nucl Sci* 1983;30(1):676–80.
- [103] Veklerov E, Llacer J, Hoffman EJ. MLE reconstruction of a brain phantom using a Monte Carlo transition matrix and a statistical stopping rule. *IEEE Trans Nucl Sci* 1988;35(1):603–7.
- [104] Boning G, Pichler BJ, Rafecas M, Lorenz E, Schwaiger M, Ziegler SI. Implementation of Monte Carlo coincident aperture functions in image generation of a high-resolution animal positron tomograph. *IEEE Trans Nucl Sci* 2001;48(3):805–10.
- [105] Rafecas M, Mosler B, Dietz M, Pogl M, Stamatakis A, McElroy DP, et al. Use of a Monte Carlo-based probability matrix for 3-D iterative reconstruction of MADPET-II data. *IEEE Trans Nucl Sci* 2004;51(5):2597–605.
- [106] Ortuño JE, Kontaxakis G, Guerra P, Santos A. 3D-OSEM transition matrix for high resolution PET imaging with modeling of the gamma-event detection. In: Nuclear science symposium. 2004. p. 3046–50.
- [107] Rannou FR, Chatzioannou AF. Fully 3D system model estimation of OPET by Monte Carlo simulation. In: Nuclear science symposium. 2004. p. 3433–6.
- [108] Ortuño JE, Guerra-Gutiérrez P, Rubio JL, Kontaxakis G, Santos A. 3D-OSEM iterative image reconstruction for high-resolution PET using precalculated system matrix. *Nucl Instrum Methods Phys Res Sect A* 2006;569(2):440–4.
- [109] Vandenberghe S, Staelens S, Byrne CL, Soares EJ, Lemahieu I, Glick SJ. Reconstruction of 2D PET data with Monte Carlo generated system matrix for generalized natural pixels. *Phys Med Biol* 2006;51(12):3105–25.
- [110] Southekal S, Purschke ML, Schlyer DJ, Vaska P. Quantitative PET imaging using a comprehensive Monte Carlo system model. *IEEE Trans Nucl Sci* 2011;58(5):2286–95.
- [111] Rafecas M, Boning G, Pichler BJ, Lorenz E, Schwaiger M, Ziegler SI. Effect of noise in the probability matrix used for statistical reconstruction of PET data. *IEEE Trans Nucl Sci* 2004;51(1):149–56.
- [112] Nerina Gimenez E, Nacher E, Gimenez M, Beniloch J, Rafecas M. Comparison of different approaches based on Monte Carlo methods to calculate the system matrix for small animal PET. *Nucl Instrum Methods Phys Res Sect A* 2006;569(2):346–9.
- [113] Cabello J, Rafecas M. System response matrix denoising in measurement space in positron emission tomography. In: Nuclear science symposium and medical imaging conference. 2011. p. 4214–8.
- [114] Yamaya T, Yoshida E, Obi T, Ito H, Yoshikawa K, Murayama H. First human brain imaging by the jPET-D4 prototype with a pre-computed system matrix. *IEEE Trans Nucl Sci* 2008;55(5):2482–92.
- [115] Kudrolli H, Worstell W, Zavarzin V. SS3D-Fast fully 3D PET iterative reconstruction using stochastic sampling. *IEEE Trans Nucl Sci* 2002;49(1):124–30.
- [116] Gillam J, Solevi P, Oliver J, Rafecas M. Simulated one-pass list-mode: an approach to on-the-fly system matrix calculation. *Phys Med Biol* 2013;58(7):2377.
- [117] Hun JJ, Miyaoka RS, Lewellen TK. cMiCE: a high resolution animal PET using continuous LSO with a statistics based positioning scheme. *Nucl Instrum Methods Phys Res A* 2002;489(1):584–98.
- [118] Tavernier S, Bruyndonckx P, Leonard S, Devroede O. A high-resolution PET detector based on continuous scintillators. *Nucl Instrum Methods Phys Res A* 2005;537(12):321–5.
- [119] Beniloch J, Carrilero V, Catret J, Colom R, Correcher C, Gadea R, et al. Design and calibration of a small animal PET scanner based on continuous LYSO Crystals and PSPMTs. In: Nuclear science symposium. 2006. p. 2328–32.
- [120] Balcerzyk M, Kontaxakis G, Delgado M, Garcia-Garcia L, Correcher C, Gonzalez AJ, et al. Initial performance evaluation of a high resolution Albira small animal positron emission tomography scanner with monolithic crystals and depth-of-interaction encoding from a user's perspective. *Meas Sci Technol* 2009;20:104011.
- [121] Siddon RL. Fast calculation of the exact radiological path for a three-dimensional CT array. *Med Phys* 1985;12(2):252–5.
- [122] Ollinger JM, Goggin AS. Maximum likelihood reconstruction in fully 3D PET via the SAGE algorithm. In: Nuclear science symposium. 1996. p. 1594–8.
- [123] Scheins JJ, Bosch F, Herzog H. Analytical calculation of volumes-of-intersection for iterative, fully 3-D PET reconstruction. *IEEE Trans Med Imaging* 2006;25(10):1363–9.
- [124] Scheins J, Herzog H, Shah NJ. Fully-3D PET image reconstruction using scanner-independent, adaptive projection data and highly rotation-symmetric voxel assemblies. *IEEE Trans Med Imaging* 2011;30(3):879–92.
- [125] Lougovski A, Hofheinz F, Maus J, Schramm G, Will E, van den Hoff J. A volume of intersection approach for on-the-fly system matrix calculation in 3D PET image reconstruction. *Phys Med Biol* 2014;59(3):561.
- [126] Lougovski A, Hofheinz F, Maus J, Schramm G, van den Hoff J. On the relation between Kaiser-Bessel blob and tube of response based modelling of the system matrix in iterative PET image reconstruction. *Phys Med Biol* 2015;60(10):4209.
- [127] Rahmim A, Cheng JC, Blinder S, Cambarde ML, Sossi V. Statistical dynamic image reconstruction in state-of-the-art high-resolution PET. *Phys Med Biol* 2005;50(20):4887.
- [128] Schretter C. A fast tube of response ray-tracer. *Med Phys* 2006;33(12):4744–8.
- [129] Aguiar P, Rafecas M, No JEO, Kontaxakis G, Santos A, Pava J, et al. Geometrical and Monte Carlo projectors in 3D PET reconstruction. *Med Phys* 2010;37(11):5691–702.
- [130] Musmann P, Pietrzyk U, Schramm N, Weber S. Parallel list-mode reconstruction and calculation of the system matrix for the high-resolution ClearPET Neuro. In: Nuclear science symposium. 2006. p. 2783–6.
- [131] Weber S, Morel C, Simon L, Krieguer M, Rey M, Gundlich B, et al. Image reconstruction for the ClearPET Neuro. *Nucl Instrum Methods Phys Res Sect A* 2006;569(2):381–5.
- [132] Chen CM, Lee SY, Cho ZH. Parallelization of the EM algorithm for 3-D PET image reconstruction. *IEEE Trans Med Imaging* 1991;10(4):513–22.
- [133] Terstege A, Weber S, Herzog H, Muller-Gartner HW, Halling H. High resolution and better quantification by tube of response modelling in 3D PET reconstruction. In: Nuclear science symposium. 1996. p. 1603–7.
- [134] De la Prieta R, Hernandez J, Schiavi E, Malpica N. Analytical geometric model for photon coincidence detection in 3D PET. In: Nuclear science symposium. 2006. p. 2229–32.
- [135] Yun MK, Liu SQ, Shan BG, Wei L. An advanced fully 3D OSEM reconstruction for positron emission tomography. *Chin Phys C* 2010;34(2):231–6.
- [136] Johnson CA, Yan Y, Carson RE, Martino RL, Daube-Witherspoon ME. A system for the 3D reconstruction of retracted-septa PET data using the EM algorithm. In: Nuclear science symposium and medical imaging conference. 1994. p. 1325–9.
- [137] Selivanov VV, Picard Y, Cadorette J, Rodrigue S, Lecomte R. Detector response models for statistical iterative image reconstruction in high resolution PET. *IEEE Trans Nucl Sci* 2000;47(3):1168–75.
- [138] Yamaya T, Hagiwara N, Obi T, Yamaguchi M, Ohyama N, Kitamura K, et al. Transaxial system models for jPET-D4 image reconstruction. *Phys Med Biol* 2005;50(22):5339.
- [139] Yamaya T, Yoshida E, Kitamura K, Obi T, Tanimoto K, Yoshikawa K, et al. First Human brain images of the jPET-D4 using 3D OS-EM with a pre-computed system matrix. In: Nuclear science symposium. 2006. p. 3384–7.
- [140] Takahashi H, Yamaya T, Kobayashi T, Kitamura K, Hasegawa T, Murayama H, et al. System modeling of small bore DOI-PET scanners for fast and accurate 3D image reconstruction. In: Nuclear science symposium. 2007. p. 3478–81.
- [141] Frese T, Rouze NC, Bouman CA, Sauer K, Hutchins GD. Quantitative comparison of FBP, EM, and Bayesian reconstruction algorithms for the IndyPET scanner. *IEEE Trans Med Imaging* 2003;22(2):258–76.

- [142] Yao R, Ramachandra RM, Mahajan N, Rathod V, Gunasekar N, Panse A, et al. Assessment of a three-dimensional line-of-response probability density function system matrix for PET. *Phys Med Biol* 2012;57(21):6827.
- [143] Fin L, Bailly P, Daouk J, Meyer ME. A practical, semi-experimental system matrix for 2-D PET image reconstruction: comparison with a geometrical model. In: Nuclear science symposium. 2008. p. 5053–9.
- [144] De Bernardi E, Mazzoli M, Zito F, Baselli G. Resolution recovery in PET during AWOSEM reconstruction: a performance evaluation study. *IEEE Trans Nucl Sci* 2007;54(5):1626–38.
- [145] Ortu no JE, Sportelli G, Guerra P, Santos A. Projector model for efficient list-mode reconstruction in PET scanners with parallel planar detectors. In: Nuclear science symposium and medical imaging conference. 2011. p. 4180–3.
- [146] Sportelli G, Ortu no J, Vaquero J, Desco M, Santos A. Massively parallelizable list-mode reconstruction using a Monte Carlo-based elliptical Gaussian model. *Med Phys* 2012;40(1):012504.
- [147] Panin VY, Kehlen F, Rothfuss H, Hu D, Michel C, Casey ME. PET reconstruction with system matrix derived from point source measurements. *IEEE Trans Nucl Sci* 2006;53(1):152–9.
- [148] Wiant DB, Gersh JA, Bennett MC, Bourland JD. PET image reconstruction using LOR-OSEM with a 3D spatially variant system matrix. In: Nuclear science symposium and medical imaging conference. 2009. p. 3752–8.
- [149] Taschereau R, Rannou F, Chatzioannou A. A modeled point spread function for a noise-free system matrix. In: Nuclear science symposium and medical imaging conference. 2011. p. 4102–5.
- [150] Jian Y, Yao R, Mulnix T, Carson R. Validation of the spatially variant probability density functions for the HRRT. In: Nuclear science symposium and medical imaging conference. 2011. p. 2443–6.
- [151] Shakirin G, Crespo P, Enghardt W. A method for system matrix construction and processing for reconstruction of in-beam PET data. *IEEE Trans Nucl Sci* 2007;54(5):1710–6.
- [152] Su KH, Wu LC, Lee JS, Liu RS, Chen JC. A novel method to improve image quality for 2-D small animal PET reconstruction by correcting a Monte Carlo-simulated system matrix using an artificial neural network. *IEEE Trans Nucl Sci* 2009;56(3):704–14.
- [153] DeBenedetti S, Cowan CE, Konneker WR, Primakoff H. On the angular distribution of two-photon annihilation radiation. *Phys Rev* 1950;77(2):205–12.
- [154] Shibuya K, Yoshida E, Nishikido F, Suzuki T, Tsuda T, Inadama N, et al. Annihilation photon acollinearity in PET: volunteer and phantom FDG studies. *Phys Med Biol* 2007;52(17):5249.
- [155] Cal-Gonzalez J, Herraiz JL, España S, Desco M, Vaquero JJ, Udiás JM. Positron range effects in high resolution 3D PET imaging. In: Nuclear science symposium and medical imaging conference. 2009. p. 2788–91.
- [156] Palmer MR, Brownell GL. Annihilation density distribution calculations for medically important positron emitters. *IEEE Trans Med Imaging* 1992;11(3):373–8.
- [157] Palmer MR, Zhu X, Parker JA. Modeling and simulation of positron range effects for high resolution PET imaging. *IEEE Trans Nucl Sci* 2005;52(5):1391–5.
- [158] Levin CS, Hoffman EJ. Calculation of positron range and its effect on the fundamental limit of positron emission tomography system spatial resolution. *Phys Med Biol* 1999;44(3):781–99.
- [159] Champion C, Loirec CL. Positron follow-up in liquid water: II. Spatial and energetic study for the most important radioisotopes used in PET. *Phys Med Biol* 2007;52(22):6605–25.
- [160] Cho ZH, Chan JK, Erickson L, Singh M, Graham S, MacDonald NS, et al. Positron ranges obtained from biomedically important positron-emitting radionuclides. *J Nucl Med* 1975;16(12):1174–6.
- [161] Phelps ME, Hoffman EJ, Huang SC, Ter-Pogossian MM. Effect of positron range on spatial resolution. *J Nucl Med* 1975;16(7):649–52.
- [162] Alessio AM, Stearns CW, Tong S, Ross SG, Kohlmyer S, Ganin A, et al. Application and evaluation of a measured spatially variant system model for PET image reconstruction. *IEEE Trans Med Imaging* 2010;29(3):938–49.
- [163] Baker JR. Spatially variant tomographic imaging: estimation, identification, and optimization. Report. USA: Lawrence Berkeley Lab; 1991.
- [164] Kearfott K. Practical considerations, comment in “A statistical model for positron emission tomography”. *J Am Stat Assoc* 1985;80:26–8.
- [165] Mumcuoglu EU, Leahy RM, Cherry SR, Hoffman E. Accurate geometric and physical response modelling for statistical image reconstruction in high resolution PET. In: Nuclear science symposium. 1996. p. 1569–73.
- [166] Worstell WA, Adler SS, Kudrolli HA, Nevin JD, Romanov LV. Compact representation of PET 3D system response matrices. In: International symposium on biomedical imaging: nano to macro. 2004. p. 756–9.
- [167] Markiewicz P, Reader A, Tamal M, Julian P, Hastings D. An advanced analytic method incorporating the geometrical properties of scanner and radiation emissions into the system model for the true component of 3D PET data. In: Nuclear science symposium. 2005. p. 2310–4.
- [168] Cloquet C, Sureau F, Defrise M, Van Smaey G, Trotta N, Goldman S. Non-Gaussian space-variant resolution modelling for list-mode reconstruction. *Phys Med Biol* 2010;55(17):5045.
- [169] Cecchetti M, Moehrs S, Belcari N, Del Guerra A. Accurate and efficient modeling of the detector response in small animal multi-head PET systems. *Phys Med Biol* 2013;58(19):6713.
- [170] Zhou J, Qi J. Fast and efficient fully 3D PET image reconstruction using sparse system matrix factorization with GPU acceleration. *Phys Med Biol* 2011;56(20):6739.
- [171] Zhou J, Qi J. Efficient fully 3D list-mode TOF PET image reconstruction using a factorized system matrix with an image domain resolution model. *Phys Med Biol* 2014;59(3):541.
- [172] Rahmim A, Tang J, Lodge MA, Lashkari S, Ay MR, Lautamki R, et al. Analytic system matrix resolution modeling in PET: an application to Rb-82 cardiac imaging. *Phys Med Biol* 2008;53(21):5947–65.
- [173] Lee K, Kinahan PE, Fessler JA, Miyaoka RS, Janes M, Lewellen TK. Pragmatic fully 3D image reconstruction for the MiCES mouse imaging PET scanner. *Phys Med Biol* 2004;49(19):4563–78.
- [174] Alessio AM, Kinahan PE, Harrison RL, Lewellen TK. Measured spatially variant system response for PET image reconstruction. In: Nuclear science symposium. 2005. p. 5.
- [175] Alessio AM, Kinahan PE, Lewellen TK. Modeling and incorporation of system response functions in 3-D whole body PET. *IEEE Trans Med Imaging* 2006;25(7):828–37.
- [176] Bao Q, Sanghee C, Quanzheng L, Newport D, Mu C, Leahy R, et al. Monte Carlo based estimation of detector response in a large solid angle preclinical PET imaging system. In: Nuclear science symposium. 2008. p. 5010–3.
- [177] Tohme MS, Qi J. Iterative image reconstruction for positron emission tomography based on a detector response function estimated from point source measurements. *Phys Med Biol* 2009;54(12):3709–25.
- [178] Reader AJ, Julian PJ, Williams H, Hastings DL, Zweit J. EM algorithm system modeling by image-space techniques for PET reconstruction. *IEEE Trans Nucl Sci* 2003;50(5):1392–7.
- [179] Reader AJ, Zweit J. Developments in whole-body molecular imaging of live subjects. *Trends Pharmacol Sci* 2001;22(12):604–7.
- [180] Reader AJ, Ally S, Bakatselos F, Manavaki R, Walledge RJ, Jeavons AP, et al. One-pass list-mode EM algorithm for high-resolution 3-D PET image reconstruction into large arrays. *IEEE Trans Nucl Sci* 2002;49(3):693–9.
- [181] Fazendeiro L, Ferreira NC, Blanco A, Fonte P. EM reconstruction algorithm with resolution modeling applied to an RPC-PET prototype. In: Nuclear science symposium. 2004. p. 2511–5.
- [182] Varrone A, Sjöholm N, Eriksson L, Gulyás B, Halldin C, Farde L. Advancement in PET quantification using 3D-OP-OSEM point spread function reconstruction with the HRRT. *Eur J Nucl Med Mol Imaging* 2009;36(10):1639–50.
- [183] Mourik JE, Lubberink M, Van Velden FH, Kloet RW, Van Berckel BN, Lamertsmas AA, et al. In vivo validation of reconstruction-based resolution recovery for human brain studies. *J Cereb Blood Flow Metab* 2010;30(2):381–9.
- [184] Hoetjes NJ, van Velden FH, Hoekstra OS, Hoekstra CJ, Krak NC, Lammertsma AA, et al. Partial volume correction strategies for quantitative FDG PET in oncology. *Eur J Nucl Med Mol Imaging* 2010;37(9):1679–87.
- [185] Rahmim A, Lenox M, Michel C, Reader AJ, Sossi V. Space-variant and anisotropic resolution modeling in list-mode EM reconstruction. In: Nuclear science symposium. 2003. p. 3074–7.
- [186] Antich P, Parkey R, Seliounine S, Slavine N, Tsyganov E, Zinchenko A. Application of expectation maximization algorithms for image resolution improvement in a small animal PET system. *IEEE Trans Nucl Sci* 2005;52(3):684–90.
- [187] Rapisarda E, Bettinardi V, Thielemans K, Gilardi M. Image-based point spread function implementation in a fully 3D OSEM reconstruction algorithm for PET. *Phys Med Biol* 2010;55(14):4131.
- [188] Kotasidis F, Matthews J, Angelis G, Noonan P, Jackson A, Price P, et al. Single scan parameterization of space-variant point spread functions in image space via a printed array: the impact for two PET/CT scanners. *Phys Med Biol* 2011;56(10):2917.
- [189] Sureau FC, Reader AJ, Comtat C, Leroy C, Ribeiro MJ, Buvat I, et al. Impact of image-space resolution modeling for studies with the high-resolution research tomograph. *J Nucl Med* 2008;49(6):1000–8.
- [190] Kotasidis FA, Angelis GI, Anton-Rodriguez J, Markiewicz P, Lionheart WR, Reader AJ, et al. Image-based spatially variant and count rate dependent point spread function on the HRRT. *IEEE Trans Nucl Sci* 2014;61(3):1192–202.
- [191] Derenzo SE. Mathematical removal of positron range blurring in high resolution tomography. *IEEE Trans Nucl Sci* 1986;33(1):565–9.
- [192] Huesman RH, Salmeron EM, Baker JR. Compensation for crystal penetration in high resolution positron tomography. *IEEE Trans Nucl Sci* 1989;36(1):1100–7.
- [193] Liang Z. Detector response restoration in image reconstruction of high resolution positron emission tomography. *IEEE Trans Med Imaging* 1994;13(2):314–21.
- [194] Orero A, Correcher C, Gonzalez A, Soriano A, Carles M, Moliner L, et al. Promising results on PSF correction applied in the reconstruction process of a small animal PET image. In: Nuclear science symposium and medical imaging conference. 2011. p. 2870–3.
- [195] Teo BK, Seo Y, Bacharach SL, Carrasquillo JA, Libutti SK, Shukla H, et al. Partial-volume correction in PET: validation of an iterative postreconstruction method with phantom and patient data. *J Nucl Med* 2007;48(5):802–10.
- [196] Kirov A, Piao J, Schmidlein C. Partial volume effect correction in PET using regularized iterative deconvolution with variance control based on local topology. *Phys Med Biol* 2008;53(10):2577.
- [197] Barbee DL, Flynn RT, Holden JE, Nickles RJ, Jeraj R. A method for partial volume correction of PET-imaged tumor heterogeneity using expectation maximization with a spatially varying point spread function. *Phys Med Biol* 2010;55(1):221.
- [198] Brix G, Doll J, Bellemann ME, Trojan H, Haberkorn U, Schmidlin P, et al. Use of scanner characteristics in iterative image reconstruction for high-resolution

- positron emission tomography studies of small animals. *Eur J Nucl Med Mol Imaging* 1997;24(7):779–86.
- [199] D'Ambrosio D, Fiacchi G, Cilibizzi P, Lamberti C, Baldazzi G, Boschi S, et al. Partial volume correction of small animal PET cardiac dynamic images using iterative reconstruction: effects on glucose metabolic rate measurement. In: *Computers in cardiology*. 2008. p. 1093–6.
- [200] Rizzo G, Castiglioni I, Russo G, Tana M, Dell'Acqua F, Gilardi M, et al. Using deconvolution to improve PET spatial resolution in OSEM iterative reconstruction. *Method Inf Med* 2007;46(2):231.
- [201] Li H, Mawlawi OR, Zhu RX, Zheng Y. Blind deblurring reconstruction technique with applications in PET imaging. *Int J Biomed Imaging* 2009;2009:3.
- [202] Agbeko N, Cheng JC, Laforest R, O'Sullivan A. Positron range correction in PET using an alternating EM algorithm. In: *Nuclear science symposium and medical imaging conference*. 2010. p. 2875–8.
- [203] Spinelli AE, Fiacchi G, D'Ambrosio D, Boschi S, Franchi R, Marengo M. Iterative EM reconstruction of cardiac small animal PET images using system point spread function modelling and MAP with anatomical priors. In: *Nuclear science symposium*. 2008. p. 5147–52.
- [204] Lecomte R, Schmitt D, Lamoureux G. Geometry study of a high resolution PET detection system using small detectors. *IEEE Trans Nucl Sci* 1984;31(1):556–61.
- [205] Schmitt D, Karuta B, Carrier C, Lecomte R. Fast point spread function computation from aperture functions in high-resolution positron emission tomography. *IEEE Trans Med Imaging* 1988;7(1):2–12.
- [206] Strul D, Slates RB, Dahlblom M, Cherry SR, Marsden PK. An improved analytical detector response function model for multilayer small-diameter PET scanners. *Phys Med Biol* 2003;48(8):979–94.
- [207] Yamaya T, Hagiwara N, Obi T, Yamaguchi M, Kita K, Ohya N, et al. DOI-PET image reconstruction with accurate system modeling that reduces redundancy of the imaging system. *IEEE Trans Nucl Sci* 2003;50(5):1404–9.
- [208] Staelens S, D'Asseler Y, Vandenbergh S, Koole M, Lemahieu I, de Walle RV. A three-dimensional theoretical model incorporating spatial detection uncertainty in continuous detector PET. *Phys Med Biol* 2004;49(11):2337–50.
- [209] Zhang X, Stortz G, Sossi V, Thompson CJ, Retière F, Kozlowski P, et al. Development and evaluation of a LOR-based image reconstruction with 3D system response modeling for a PET insert with dual-layer offset crystal design. *Phys Med Biol* 2013;58(23):8379.
- [210] Kaufman L. Implementing and accelerating the EM algorithm for positron emission tomography. *IEEE Trans Med Imaging* 1987;6(1):37–51.
- [211] Jones M, Yao R, Bhole C. Hybrid MPI-OpenMP programming for parallel OSEM PET reconstruction. *IEEE Trans Nucl Sci* 2006;53(5):2752–8.
- [212] Valastyán I, Imrek J, Molnár J, Novák D, Balkay L, Emri M, et al. Full 3-D cluster-based iterative image reconstruction tool for a small animal PET camera. *Nucl Instrum Methods Phys Res Sect A* 2007;571(1):219–22.
- [213] Beisel T, Lietsch S, Thielemans K. A method for OSEM PET reconstruction on parallel architectures using STIR. In: *Nuclear science symposium conference*. 2008. p. 4161–8.
- [214] Lam CF, Yamaya T, Obi T, Yoshida E, Inadama N, Shibuya K, et al. Parallel implementation of 3-D iterative reconstruction with intra-thread update for the jPET-D4. *IEEE Trans Nucl Sci* 2009;56(1):129–35.
- [215] Torres F, de la Fuente F, Rannou FR. Shared-memory alternatives for parallel image reconstruction. In: *Nuclear science symposium and medical imaging conference*. 2011. p. 2541–4.
- [216] Bai B, Smith AM. Fast 3D iterative reconstruction of PET images using PC graphics hardware. In: *Nuclear science symposium*. 2006. p. 2787–90.
- [217] Cui Jy, Pratz G, Prevhal S, Levin CS. Fully 3D list-mode time-of-flight PET image reconstruction on GPUs using CUDA. *Med Phys* 2011;38(12):6775–86.
- [218] Pratz G, Chinn G, Olcott PD, Levin CS. Fast, accurate and shift-varying line projections for iterative reconstruction using the GPU. *IEEE Trans Med Imaging* 2009;28(3):435–45.
- [219] Cui J, Pratz G, Prevhal S, Zhang B, Shao L, Levin CS. Measurement-based spatially-varying point spread function for list-mode PET reconstruction on GPU. In: *Nuclear science symposium and medical imaging conference*. 2011. p. 2593–6.
- [220] Barker WC, Thada S, Dieckmann W. A GPU-accelerated implementation of the MOLAR PET reconstruction package. In: *Nuclear science symposium*. 2009. p. 4114–9.
- [221] Herraiz JL, Na SE, Cabido R, Montemayor AS, Desco M, Vaquero JJ, et al. GPU-based fast iterative reconstruction of fully 3-D PET sinograms. *IEEE Trans Nucl Sci* 2011;58(5):2257–63.
- [222] Ha S, Matej S, Ispiryan M, Mueller K. GPU-accelerated forward and back-projections with spatially varying kernels for 3D DIRECT TOF PET reconstruction. *IEEE Trans Nucl Sci* 2013;60(1):166.
- [223] Matej S, Surti S, Jayanthi S, Daube-Witherspoon ME, Lewitt RM, Karp JS. Efficient 3-D TOF PET reconstruction using view-grouped histo-images: Direct-Direct image reconstruction for TOF. *IEEE Trans Med Imaging* 2009;28(5):739–51.
- [224] Lin Y, Li Q, Leahy RM. Fast GPU-based time-of-flight MAP reconstruction with a factored system matrix. In: *Nuclear science symposium and medical imaging conference*. 2010. p. 2889–93.
- [225] Pratz G, Levin C. Online detector response calculations for high-resolution PET image reconstruction. *Phys Med Biol* 2011;56(13):4023–40.
- [226] Cui J, Pratz G, Meng B, Levin CS. Distributed MLEM: an iterative tomographic image reconstruction algorithm for distributed memory architectures. *IEEE Trans Med Imaging* 2013;32(5):957–67.
- [227] Kinouchi S, Yamaya T, Yoshida E, Tashima H, Kudo H, Haneishi H, et al. GPU-based PET image reconstruction using an accurate geometrical system model. *IEEE Trans Nucl Sci* 2012;59(5):1977–83.
- [228] Vega-Acevedo N, Lanás I, Rannou FR. GATE simulations on multicomputer architectures. In: *Nuclear science symposium and medical imaging conference*. 2009. p. 550–3.
- [229] Wang Z, Olcott PD, Levin CS. Acceleration of PET Monte Carlo simulation using the graphics hardware ray-tracing engine. In: *Nuclear science symposium and medical imaging conference*. 2010. p. 1848–55.
- [230] Magdics M, Szirmay-Kalos L, Tóth B, Légrády D, Cserkaszy A, Balkay L, et al. Performance evaluation of scatter modeling of the GPU-based Tera-Tomo 3D PET reconstruction. In: *Nuclear science symposium and medical imaging conference*. 2011. p. 4086–8.
- [231] Kim K, Son Y, Cho Z, Ra J, Ye J. Ultra-fast hybrid CPU-GPU multiple scatter simulation for 3D PET. *IEEE J Biomed Health Inf* 2013;18(1):148–56.
- [232] Mora C, Rafecas M. Polar pixels for high resolution small animal PET. In: *Nuclear science symposium*. 2006. p. 2812–7.
- [233] Ansorge R. List mode 3D PET reconstruction using an exact system matrix and polar voxels. In: *Nuclear science symposium*. 2007. p. 3454–7.
- [234] Leroux JD, Thibaudeau C, Lecomte R, Fontaine R. Fast, accurate and versatile Monte Carlo method for computing system matrix. In: *Nuclear science symposium*. 2007. p. 3644–8.
- [235] Cabello J, Rafecas M. Comparison of basis functions for 3D PET reconstruction using a Monte Carlo system matrix. *Phys Med Biol* 2012;57(7):1759.
- [236] Zhang L, Staelens S, Van Hoen R, De Beenhouwer J, Verhaeghe J, Kawrakow I, et al. Fast and memory-efficient Monte Carlo-based image reconstruction for whole-body PET. *Med Phys* 2010;37(7):3667–76.
- [237] Li K, Safavi-Naeini M, Franklin D, Han Z, Rosenfeld AB, Hutton B, et al. A new virtual ring-based system matrix generator for iterative image reconstruction in high resolution small volume PET systems. *Phys Med Biol* 2015;60(17):6949.
- [238] Rahmim A, Qi J, Sossi V. Resolution modeling in PET imaging: theory, practice, benefits, and pitfalls. *Med Phys* 2013;40(6):064301.
- [239] Alessio AM, Rahmim A, Orton CG. Resolution modeling enhances PET imaging. *Med Phys* 2013;40(12):120601.
- [240] Daube-Witherspoon ME, Matej S, Werner ME, Surti S, Karp JS. Impact of resolution modeling on accuracy and precision of lesion contrast measurements. In: *Nuclear science symposium and medical imaging conference*. 2011. p. 4373–7.
- [241] Tong S, Alessio A, Kinahan P. Noise and signal properties in PSF-based fully 3D PET image reconstruction: an experimental evaluation. *Phys Med Biol* 2010;55(5):1453.
- [242] Rahmim A, Tang J. Noise propagation in resolution modeled PET imaging and its impact on detectability. *Phys Med Biol* 2013;58(19):6945.
- [243] Ahn S, Asma E, Thielemans K, Deller T, Ross S, Stearns C. Trade-off between contrast recovery, image noise and edge artifacts in PET image reconstruction using detector blurring models. In: *Nuclear science symposium and medical imaging conference*. 2011. p. 4110–4.
- [244] Stute S, Comtat C. Practical considerations for image-based PSF and blobs reconstruction in PET. *Phys Med Biol* 2013;58(11):3849.
- [245] Tong S, Alessio AM, Thielemans K, Stearns C, Ross S, Kinahan PE. Properties and mitigation of edge artifacts in PSF-based PET reconstruction. *IEEE Trans Nucl Sci* 2011;58(5):2264–75.
- [246] Watson CC. Estimating effective model kernel widths for PSF reconstruction in PET. In: *Nuclear science symposium and medical imaging conference*. 2011. p. 2368–74.
- [247] Sureau F, Comtat C, Reader A, Leroy C, Santiago-Ribeiro M, Buvat I, et al. Improved clinical parametric imaging using list-mode reconstruction via resolution modeling. In: *Nuclear science symposium*. 2006. p. 3507–10.
- [248] Akamatsu G, Mitsumoto K, Taniguchi T, Tsutsui Y, Baba S, Sasaki M. Influences of point-spread function and time-of-flight reconstructions on standardized uptake value of lymph node metastases in FDG-PET. *Eur J Radiol* 2014;83(1):226–30.
- [249] Spinelli AE, D'Ambrosio D, Fiacchi G, Boschi S, Franchi R, Marengo M. Pixel-based partial volume correction of small animal PET images using point spread function system characterization: evaluation of effects on cardiac output, perfusion and metabolic rate using parametric images. In: *Nuclear science symposium*. 2008. p. 4260–5.
- [250] Pal D, Keesing D, O'Sullivan A, Komarov S, Tai YC. Image reconstruction algorithm for a special geometry of the PET-insert system. In: *Nuclear science symposium*. 2007. p. 2791–8.
- [251] Keesing DB, Pal D, O'Sullivan A, Komarov S, Wu H, Tai YC. System modeling of a DOI-capable half-ring PET insert device for breast cancer imaging. In: *Nuclear science symposium*. 2008. p. 4218–22.