Submitted to: *Phys. Med. Biol.* ***51*** *(2006) 2919-2937*

PET Energy-based Scatter Estimation and Image Reconstruction with Energy-dependent Corrections

Lucreţiu M. Popescu, Robert M. Lewitt, Samuel Matej and Joel S. Karp
University of Pennsylvania, Department of Radiology
423 Guardian Drive, 4-th floor Blockley Hall, Philadelphia, PA 19104-6021, USA
E-mail: `popescu@mipg.upenn.edu`

Abstract. In this paper we propose a comprehensive energy based scatter correction approach for positron emission tomography (PET). We take advantage of the marked difference between the energy spectra of the unscattered and scattered photons, and use the detailed energy information that comes with the list-mode data for the estimation of the scattered events distribution in the data space. Also, inside the maximum-likelihood expectation maximization (ML-EM) image reconstruction algorithm, we introduce energy dependent factors that individualize the correction terms for each event, given its position and energy information. The central piece of our approach is the two-dimensional detector energy response model represented as a linear combination of four components, each one representing a particular state a PET event can be found in: both photons unscattered, the second scattered while the first not, first photon scattered while the second not, and both photons scattered. For a set of events collected in the vicinity of a point in the projection space, the coefficient of each component is determined by applying a statistical estimator. As result we obtain the number of scattered events that are in the given set. The model gives us also the variation of scatter fraction with the photon pair energies for that particular position in the data space. A simulation study that demonstrates the proposed methods is presented.

1. Introduction

The detector systems of current positron emission tomography (PET) scanners are capable of estimating the energies (in addition to the positions) of the two photons that are detected in time coincidence. Although the detectors do not have perfect energy resolution, it would be desirable for the measured energies of the two photons to be taken into account in the subsequent processing of the coincidence data to form the reconstructed image. This adds two new parameters to the other four to seven parameters usually measured and used in the reconstruction algorithms. More specifically, the usual parameters are the direction (ϕ, θ), the position in the projection plane (u, v), the position of the scanner relative to the patient, the time stamp t used in dynamic reconstruction, and in certain cases the expected emission position along the projection line l estimated from the time of flight (TOF) difference of the two photons. While the usual parameters carry information about the spatio-temporal position of an emission event, the energy parameters bring information about the history of the detected photons. More specifically, the energy parameters indicate the likelihood that the history of the detected particles actually corresponds to the model used in the reconstruction and not to other competing mechanisms, which are, in our case, coincidences involving at least one scattered photon.

In the current general practice the energy information is only marginally used in the reconstruction algorithms, most often by simply accepting for reconstruction only the events with measured energies in an energy window designed to include most of the unscattered photons (Ollinger 1996a), and not making further use of the specific energy values of the individual events. These accepted events include not only the events that really correspond to the reconstruction model (trues), but also include, with variable fractions, coincidences due to scattered photons (scatters) and random coincidences (randoms) both representing competing mechanisms. In order to produce unbiased estimation of the activity distribution one has to estimate these fractions and use them as correction terms or factors in the reconstruction algorithm. The amount of randoms is usually estimated from additional measurements (counts in a delayed coincidence window, or count rates of single detected photons "singles") from which the distribution of a single scalar over the sinogram space is determined. For scatter estimation the current state of the art methods involve image based simulations in which the distribution of a single scalar value (the scatter fraction) in the sinogram space is determined from an initial estimate of the activity distribution, together with a map of the attenuation distribution. The usual used image based methods are either single scatter simulations (with multiple scatter neglected) (Ollinger 1996b, Watson et al. 1996, Watson 2000, Werling et al. 2002, Accorsi et al. 2004) or Monte Carlo simulations (Zaidi 1999, Holdsworth et al. 2002, Qi & Huesman 2002). Other scatter correction methods were recently reviewed in (Meikle & Badawi 2003, Zaidi & Koral 2004). Of special interest for us are the energy based approaches.

The energy spectrum analysis methods were first introduced in single photon emission computed tomography (SPECT) as the dual energy window (DEW) method (Jaszczak et al. 1984, King et al. 1992). Refined approaches using triple or multiple energy windows (Koral et al. 1988, Ogawa et al. 1991, Wang & Koral 1992, Haynor et al. 1995, El Fakhri et al. 2003) were also developed. Judging from the ability to accurately estimate the scatter fraction the energy methods in SPECT give acceptable results for detectors with good energy resolution (Farncombe et al. 2004), when their assumptions are in good agreement with the theory of photon Compton scatter (see the discussion in 2.1). Today energy based scatter correction methods in SPECT are accepted in clinical practice (Narayanan et al. 2003).

Inspired from SPECT the dual energy window method was empirically extended to PET in (Bendriem et al. 1993, Grootoonk et al. 1996) and subsequently several improved variants were proposed (Shao et al. 1994, Adam et al. 2000). Compared to SPECT the energy based methods in PET had a rather limited success (Harrison et al. 1991, Cutler & Hoffman 1993) due to a combination of factors. Probably the most important factor is the more complicated relation between the photons' energies and the scatter status of a PET event compared to the case of SPECT, as we will show in the next section. Another reason was the limited two-dimensional spectral information acquired in only two or three energy windows for both photons combined. A significant factor was also the poor energy resolution of the earlier PET scanners (the DEW method was often tested on BGO scanners that had quite poor energy resolution).

A particular case of energy spectral analysis approach was proposed in (Bentourkia et al. 1995a, Bentourkia et al. 1995b). Detailed two-dimensional energy spectra were collected in 16×16 energy windows for each projection line. Scattered and non-scattered energy dependent spatial components in the projection space (mono-exponential functions) were identified through a fit. The hardware complexity as well its high statistics demand limited the adoption of this approach.

The interest in energy based scatter correction methods has been revived with the advent of list-mode data storage, where the energy information can be stored for each photon as precisely as the detector can measure. In (Levkovitz et al. 2001) a simple method assigns to each event a different weight depending on whether both photon energies are above a certain threshold, or one is below, or both energies are below the threshold. The proper threshold and the weights are empirically determined. An improved method that uses energy dependent weights for scatter correction was proposed recently (Chen et al. 2003, Chen 2005). The weights, which should represent the probability an event is a "true" given the values of its two energies, are obtained from a precomputed representative trues fraction table obtained by averaging Monte Carlo simulation results for a variety of activity and attenuation configurations.

In this paper we propose a comprehensive approach in which we use the detailed energy information that comes with the list-mode data and take advantage of the marked difference between the energy spectra of unscattered and scattered photons. The energy information is used in two ways (which may be considered separately or combined together):

- For the estimation of the scattered events distributions.
- In the reconstruction algorithm, for individualizing the correction terms (scatter and randoms) for each event, by taking into account the probability of each event to be a true, a scatter or a random given its position and energy information.

The main aim of this work is to present the principles of the proposed methods. It is focused less on the implementation details, or the search for the optimal parameter values needed for specific application conditions. However, we will give examples of practical applications demonstrating the methods, and the possible limitations and the tradeoffs are discussed. The paper is organized as follows. The theoretical aspects are presented in section 2, where we start by describing the detector energy response model for the single photon case, and the coincidence detected photons case, respectively. In the following subsection 2.2 (and in more detail in the appendices) we present the ways in which the proposed energy model can be used for statistical scatter estimation, followed by a discussion of the conditions for a robust application. Then we introduce the energy model in the maximum-likelihood image estimation formalism obtaining the equations for a maximum-likelihood expectation maximization (ML-EM) energy dependent image reconstruction algorithm. In section 3 we present a simulation study that shows how the proposed methods can be applied, and in section 4 we present the scatter estimation and image reconstruction results. The final section presents our discussion and conclusions.

2. Theory

The central piece of our approach is the two-dimensional detector energy response model represented as a linear combination of four components, each one representing a particular state a PET event can be found in: both photons unscattered, the second scattered while the first not, first photon scattered while the second not, and both photons scattered. For a set of events collected in the vicinity of a given data point (line of response — LOR), we can estimate the coefficient of each component by applying a statistical estimator. As a result we obtain the number of scattered events that are in the given set: this is the first manner in which we use the energy information. The model also gives us the variation of scatter fraction with the energies of the photons for that particular position in the data space: introducing this information in the reconstruction algorithm is the second manner in which we take advantage of the energy information.

2.1. Detector energy response model

The unscattered incident photons hit the detector with their original emission energy (511 keV for PET) and give an energy response (spectrum) typically described by the curve $A(E)$ in Figure 1. For a given detection system (or distinct regions within a detector) this shape depends mostly on the characteristics of the detector, and it can be experimentally determined.

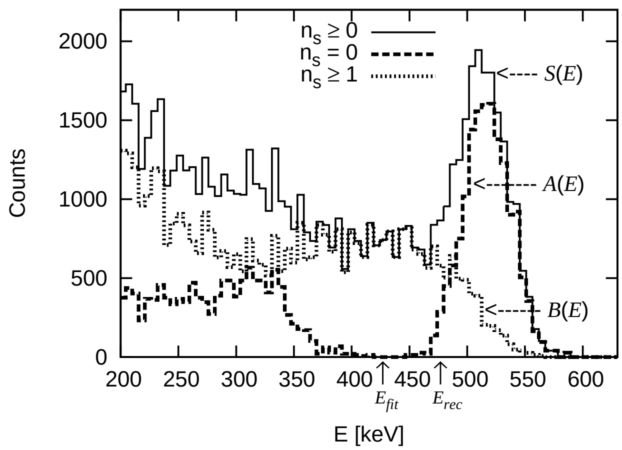

Figure 1. Typical plots of the detected photons energy spectra as a function of the number of scatters (n_s) suffered by the photon. The curve $A(E)$ represents the spectrum of unscattered photons, $B(E)$ the spectrum of the photons that suffered at least one scatter and $S(E)$ is the total spectrum. Monte Carlo simulation realized for a NaI 2.5 cm thick block detector with $\approx 10\%$ FWHM relative energy resolution. Note that the plots show the absolute count levels for $A(E)$, $B(E)$, and $S(E)$ whereas in the text the $A(E)$ and $B(E)$ are normalized.

The incident photons that suffered at least one scatter in the attenuation medium (the patient body and the bed) and lost part of their energy have a different spectrum, typically represented by the curve $B(E)$ in Figure 1. The shape of this curve depends on the detector position and intrinsic characteristic, as well as the spatial distributions of the activity and the attenuation medium. However, the energy interval we are interested in, lies above a certain limit E_{fit} and if this value is high enough the variation of the $B(E)$ curve in this region is limited. We will discuss this point in more detail in the next section (see also Figure 3). As a first approximation we will consider a unique curve $B(E)$.

With the unscattered energy component $A(E)$ determined from detector measurements, and the average scattered component $B(E)$ determined from measurements and simulations performed for typical phantom/patient sizes, we can express the energy distribution of the detected photons as

$$S(E) = tA(E) + sB(E)\,. \tag{1}$$

In this equation, s, t are the number of scattered, respectively true (not scattered) photons detected, with $A(E)$ and $B(E)$ normalized so that $\int_{\mathcal{E}} A(E)\mathrm{d}E = \int_{\mathcal{E}} B(E)\mathrm{d}E = 1$ over the energy domain of interest $\mathcal{E}$.

In the case of a PET photon pair each photon can fall into one of the two categories. Therefore, in any collection of coincidence detected photon pairs we can distinguish four groups: a fraction σ_0 of true coincidences (both photons unscattered), a fraction σ_1 of events with the second photon scattered while the first not; a fraction σ_2 of events with the first photon scattered while the second not, and another fraction σ_3 of events with both photons scattered. The total fraction of events with scattered photons is $s = \sigma_1 + \sigma_2 + \sigma_3$. The joint energy distribution for the coincidence events is given by the bilinear form

$$C(E_1, E_2) = \sum_{l=0}^{3} \sigma_l e_l(E_1, E_2) \tag{2}$$

where $e_0(E_1, E_2) = A(E_1)A(E_2)$, $e_1(E_1, E_2) = A(E_1)B(E_2)$, $e_2(E_1, E_2) = B(E_1)A(E_2)$ and $e_3(E_1, E_2) = B(E_1)B(E_2)$.

2.2. Energy based scatter estimation

In the single photon case, for a given set of detected photons with measured energies $\{E_1, \dots, E_N\}$, one can estimate s through a statistical fit of the linear expression (1). In Appendix A several methods are presented. In the case of photon pairs, similarly, by fitting the energies of the events collected in the vicinity of a given data point (LOR) with the two dimensional surface (2) we can estimate the local values of the σ coefficients. As a result we obtain the number of scattered events that are in the given set. Also, for that particular position in the data space, we obtain the variation of the scatter amount with the energies of the two photons.

For performing the fit several methods are available. The simplest is the extension of the dual energy window method for single photons to the case of photon pairs, which in fact is a 2×2 energy window method (two energy windows for each photon) as shown in Appendix B.1. More versatile are the methods that take advantage of each photon's individual energy information, such as the moments method or the maximum likelihood method. More details about the application of these methods are given in Appendix B.

The goodness of the fit depends on how different the two shapes A and B are on the energy interval being considered, as well as on the number of events available for the fit. Given the fact that the scatter (as well as the randoms) are smooth slowly varying distributions, in order to obtain good statistics for the fit we can use a coarse partitioning of the data space and wide overlapping interpolation (smoothing) kernels.

In order to increase the difference between the two shapes A and B a larger energy interval down to a lower limit E_{fit} can be used, while for the image reconstruction itself only the events above a higher setting of the minimum energy level E_{rec}, just below the photopeak, need to be considered (see Figure 1). In this case the estimates obtained for the larger energy interval (for fitting) must be renormalized for the smaller energy interval (for reconstruction). The scatter coefficients are adjusted as follows

$$\sigma_l^{\text{rec}} = \sigma_l^{\text{fit}} \int_{E_{\text{rec}}}^{E_{\max}} \int_{E_{\text{rec}}}^{E_{\max}} e_l(E_1, E_2) \mathrm{d}E_1 \mathrm{d}E_2 \,, \quad l = 0, 1, 2, 3 \tag{3}$$

We assume that initially the $e_l(E_1, E_2)$ have been normalized to unity for the fitting energy interval $(E_{\text{fit}}, E_{\max})$, with $E_{\max}$ the upper energy limit.

The robustness of the approach depends also on how stable is the shape B for different positions in the data space and with the variation of the activity and the attenuation distributions. Therefore the low energy limit cannot be set too low where the multiple scatter components induce a high variability and the detector internal Compton scatters become manifest. For the detectors with good energy resolution the $E_{\rm fit}$ can be taken high enough so the B curve will have a stable shape over the data space as shown in Figure 3. Since the A and B curves are normalized we are interested only in their relative shapes and not in their absolute values. Refinements using parametrized position dependent shapes can be also considered.

An important feature of the energy-based scatter estimation approach is that, with a proper normalization of A and B spectra components, the values obtained are absolute, thus avoiding the scaling difficulties often encountered with image based scatter estimations.

2.3. Incorporation of the energy model into the image reconstruction algorithm

We denote a point in the data space as $\mathbf{y} \equiv (\mathbf{y}^g, \mathbf{y}^e) \in \mathcal{D} \times \mathcal{E}^2$ where $\mathbf{y}^g$ stands for the geometrical coordinates covering the projection data space $\mathcal{D}$, and $\mathbf{y}^e \equiv (E_1, E_2)$ are the energies of the coincidence detected photon pair. With the activity and the scattered events distributions generically represented by f and respectively σ, the count rate density g in data point $\mathbf{y}$ can be expressed as

$$g(\mathbf{y}|f,\sigma) = g_t(\mathbf{y}|f) + g_s(\mathbf{y}|\sigma)\,, \quad \mathbf{y} = (\mathbf{y}^g, \mathbf{y}^e) \tag{4}$$

where $g_t(\mathbf{y}|f)$ is the distribution of the true (unscattered) events and $g_s(\mathbf{y}|\sigma)$ is the distribution of the scattered photons. The true events distribution depends on the activity distribution f discretely represented by $\{f_i\}_{i=\overline{1,I}}$ image elements (voxels or image basis function coefficients). With $a_i(\mathbf{y}^g)$ the probability that an emission from voxel i be detected in the data space point $\mathbf{y}^g$ (line of response — LOR) and taking into account the energy dependence of the distribution we have

$$g_t(\mathbf{y}|f) = e_0(\mathbf{y}^e) \sum_{i=1}^{I} a_i(\mathbf{y}^g) f_i \tag{5}$$

While in reality the scatter contribution depends on the image, in our model the scattered events are treated as background counts and their distribution $g_s(\mathbf{y}|\sigma)$ does not depend directly on the image. The spatial and energy distribution of the scattered events is modeled using a finite set of coefficients $\sigma \equiv \{\sigma_{1k}, \sigma_{2k}, \sigma_{3k}\}_{k=\overline{1,K}}$ following the K points of a coarse grid that covers uniformly the geometrical data space (sinogram). We use a continuous model for the spatial variation of the $\sigma_l(\mathbf{y}^g)$ coefficients, as a superposition of the neighboring σ_{lk} values

$$\sigma_l(\mathbf{y}^g) = \sum_{k=1}^{K} b_k(\mathbf{y}^g)\sigma_{lk}\,, \quad l = 1, 2, 3 \tag{6}$$

with $b_k(\mathbf{y}^g) = b(\mathbf{y}^g - \mathbf{y}^g_k)$ and $\int b(\mathbf{y}^g)\mathrm{d}\mathbf{y}^g = 1$. Keeping from the equation (2) only the terms corresponding to scattered events we have

$$g_s(\mathbf{y}|\sigma) = \sum_{l=1}^{3} e_l(\mathbf{y}^e) \sum_{k=1}^{K} b_k(\mathbf{y}^g)\sigma_{lk}\,. \tag{7}$$

For a detected event the probability density of being reported with the attributes $\mathbf{y}$ is

$$p(\mathbf{y}|f,\sigma) = \frac{1}{\Lambda}\left[g_t(\mathbf{y}|f) + g_s(\mathbf{y}|\sigma)\right] \,, \tag{8}$$

where Λ is the expected total number of events (trues and scattered) $\Lambda = \int_{\mathcal{D}} g(\mathbf{y}|f,\sigma)\mathrm{d}\mathbf{y} = \sum_{i=1}^{I} Q_i f_i + \sum_{k=1}^{K}\sum_{l=1}^{3}\sigma_{lk}$. $Q_i = \int_{\mathcal{D}} a_i(\mathbf{y}^g)\mathrm{d}\mathbf{y}^g$ is the detection sensitivity for image element i.

For a set of N events $Y \equiv \{\mathbf{y}_n\}_{n=\overline{1,N}}$. The likelihood function (the probability density) of Y given f and σ is

$$\mathcal{L}(Y|f,\sigma) = P(N|f,\sigma)\prod_{n=1}^{N} p(\mathbf{y}_n|f,\sigma) \tag{9}$$

where $P(N|f,\sigma) = \frac{\Lambda^N}{N!}\mathrm{e}^{-\Lambda}$ is the probability of detecting N events following a Poisson distribution.

The log-likelihood function is

$$\ln\mathcal{L}(Y|f,\sigma) = \sum_{n=1}^{N}\ln\left(e_0(\mathbf{y}_n^e)\sum_{i=1}^{I} a_i(\mathbf{y}_n^g)f_i + \sum_{k=1}^{K} b_k(\mathbf{y}_n^g)\sum_{l=0}^{3} e_l(\mathbf{y}_n^e)\sigma_{lk}\right) - \sum_{i=1}^{I} Q_i f_i - \sum_{k=1}^{K}\sum_{l=1}^{3}\sigma_{lk} - \ln(N!) \tag{10}$$

and has the following partial derivatives

$$\frac{\partial \ln\mathcal{L}(Y|f,\sigma)}{\partial f_i} = \sum_{n=1}^{N}\frac{a_i(\mathbf{y}_n^g)e_0(\mathbf{y}_n^e)}{g_t(\mathbf{y}_n|f) + g_s(\mathbf{y}_n|\sigma)} - Q_i \tag{11}$$

$$\frac{\partial \ln\mathcal{L}(Y|f,\sigma)}{\partial \sigma_{lk}} = \sum_{n=1}^{N}\frac{b_k(\mathbf{y}_n^g)e_l(\mathbf{y}_n^e)}{g_t(\mathbf{y}_n|f) + g_s(\mathbf{y}_n|\sigma)} - 1 \tag{12}$$

Maximum likelihood is reached for $\frac{\partial \ln\mathcal{L}(Y|f,\sigma)}{\partial f_i} = 0$ and $\frac{\partial \ln\mathcal{L}(Y|f,\sigma)}{\partial \sigma_{lk}} = 0$. The second order partial derivatives are all non-positive, which (together with the complete coverage of the object) assures the convexity of the log-likelihood function, hence the uniqueness of the solution. The activity distribution f, and that of scattered photons σ, most likely to fit the data can be obtained by using the maximum likelihood expectation maximization algorithm (ML-EM) with the following iterative equations

$$f_i^{(m+1)} = f_i^{(m)}\frac{1}{Q_i}\sum_{n=1}^{N}\frac{a_i(\mathbf{y}_n^g)e_0(\mathbf{y}_n^e)}{g_t(\mathbf{y}_n|f^{(m)}) + g_s(\mathbf{y}_n|\sigma^{(m)})} \tag{13}$$

$$\sigma_{lk}^{(m+1)} = \sigma_{lk}^{(m)}\sum_{n=1}^{N}\frac{b_k(\mathbf{y}_n^g)e_l(\mathbf{y}_n^e)}{g_t(\mathbf{y}_n|f^{(m)}) + g_s(\mathbf{y}_n|\sigma^{(m)})}, \quad l = 1,2,3 \tag{14}$$

These equations, especially (13), are similar to the ML-EM equations for list-mode data (Snyder & Politte 1983, Barrett et al. 1997, Reader et al. 1998, Qi & Huesman 2002) with the difference that we have included the energy-dependent factors, since $\mathbf{y}_n \equiv (\mathbf{y}_n^g, \mathbf{y}_n^e)$ (note the energy-dependent terms in the numerators and denominators of these equations). One can use the equations (13) and (14) together to simultaneously estimate the activity and scatter distributions, using the energy pairs as discriminating variables. However, in the numerical experiments that will be discussed in this paper we have limited ourselves to the approach described in the next subsection.

2.4. Reconstruction using energy dependent scatter correction

If equation (13) is used alone with the scatter coefficients $(\sigma_1, \sigma_2, \sigma_3)$ estimated prior to the image reconstruction through an external procedure, then we have a list mode ML-EM image reconstruction algorithm with energy dependent scatter correction for each individual event.

The coefficients $\{(\sigma_{1k}, \sigma_{2k}, \sigma_{3k})\}$ can be determined separately through an image based model such as Monte Carlo simulation. It is also possible to determine σ_{1k} and σ_{2k} by adapting the current single scatter simulation procedures to calculate separately the scatter occurring for each photon and then neglect or empirically estimate the double scatter term as $\sigma_{3k} \propto \sigma_{1k}\sigma_{2k}$. The image based simulations in addition can return LOR dependent $B(E)$ spectra.

If only the equations (14) are used, with the term $g_t(\mathbf{y}|f)$ replaced with the zero-th order scatter component $b_k(\mathbf{y}^g)e_0(\mathbf{y}^e)\sigma_{0k}$ and the corresponding equation for the component $l = 0$ included, we obtain the ML-EM equations for energy based estimation of the scatter components $(\sigma_{0k}, \sigma_{1k}, \sigma_{2k}, \sigma_{3k})$.

2.5. Consideration of random coincidences

In addition to trues and scatter coincidences the detector records also random coincidences. Therefore in the equation (4) an additional term $g_r(\mathbf{y}|\rho)$ must be considered (and as a consequence added also to the denominator in equations (13) and (14)). As in the case of trues and scatters, for a given point $\mathbf{y}^g$ in the data space the two dimensional energy distribution of random coincidence photons can be expressed as a combination $g_r(\mathbf{y}^e|\rho) = \sum_{l=0}^{3} \rho_l e_l(\mathbf{y}^e)$ with $r = \rho_0 + \rho_1 + \rho_2 + \rho_3$ total random correction. Because in the case of random coincidences the two photons are uncorrelated, unlike in the case of scatter, the estimation of the ρ_l coefficients does not require a two dimensional fit. We can express the energy distribution as the product of single scatter energy distributions of the two photons $g_r(E_1, E_2) = rR_1(E_1)R_2(E_2)$ with $R_j(E_j) = (1 - s_{rj})A(E_j) + s_{rj}B(E_j)$ for $j = 1, 2$. We have $\rho_0 = r(1 - s_{r1})(1 - s_{r2})$, $\rho_1 = r(1 - s_{r1})s_{r2}$, $\rho_2 = rs_{r1}(1 - s_{r2})$ and $\rho_3 = rs_{r1}s_{r2}$. One can estimate r, s_{r1} and s_{r2} by fitting the energies either from single photon or delayed coincidence measurements.

3. Simulation study

We considered a cylindrical PET scanner with 84 cm diameter and 24 cm length. For studying the energy spectra of scattered and unscattered photons we modeled an inhomogeneous

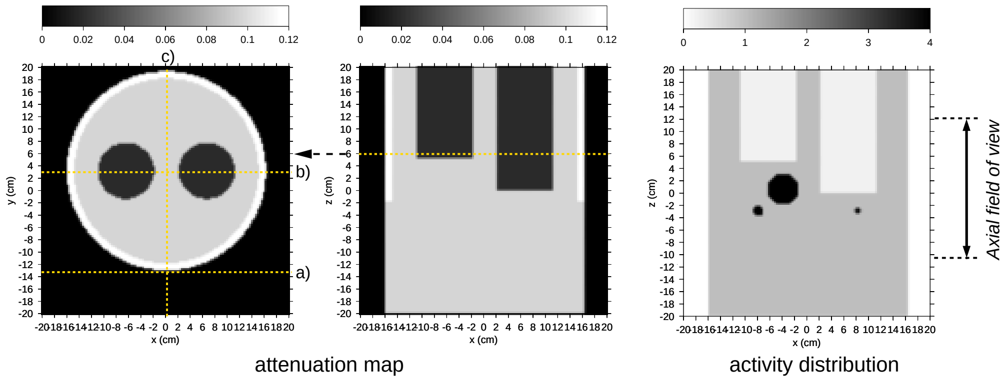

Figure 2. Sections of the inhomogeneous phantom used. The background is water equivalent, while the two cylindrical cold regions have the attenuation coefficient 20% of that of water. Several profiles are marked with dotted lines and labeled for future reference.

phantom with shape similar to that of the Utah phantom. The physical dimensions of the phantom can be seen in Figure 2 where sections through the attenuation map and activity distribution are shown. The activity distribution follows that of the attenuation coefficient with the values in the two cylindrical cold regions 0.2 times background activity. The phantom's attenuation and activity distributions are extended axially outside the scanner field of view. A large hot sphere with radius 2.5 cm and contrast 4:1 is placed off center in the middle region. In order to reduce further the system symmetry we shifted the phantom 3 cm along the y axis.

The data were generated using the Monte Carlo simulation program described in (Popescu & Lewitt 2003, Popescu 2000, Popescu 2003) by considering an analog simulation of the photon transport through the phantom and a simplified model for the photon detection using Gaussian distributions for the the energy and time of flight (TOF) resolutions. Detailed information about each detected photon pair, including the energies, the TOF data, and flags indicating the scattered photons were saved in a list using a compact format described in (Popescu et al. 2004). In this study we have not considered random coincidences.

Here we present results for two scatter estimation cases performed with the energy based moments method as described in Appendix B.2 and using the following four (m, n) pairs of moments $\{(0,0),(1,2),(2,1),(2,2)\}$.

In the first test we considered a non-TOF case with the energy resolution $\varepsilon_E = 6.4\%$ relative standard deviation (≈ 15% relative FWHM). We have selected this energy resolution value after performing several preliminary tests with various energy resolutions. The 15% FWHM energy resolution resulted as roughly the limit beyond which the energy-based scatter estimation methods performed less reliably. However, we caution against any direct extrapolation to the real scanner data since specific hardware and energy calibration aspects may have an influence that has to be determined in further investigations.

In the second case we considered a TOF case with a positioning precision along LOR of $\sigma_{\rm tof} = 3.0$ cm (≈ 470 ps FWHM timing resolution) and energy resolution $\varepsilon_E = 3.4\%$ (≈

8% FWHM). These values have been chosen as rather conservative estimates of the average operating parameters of a new TOF scanner prototype based on $LaBr_3$ scintillator crystals (Karp et al. 2005).

The data space is represented using the (ϕ, θ, u, v, l) coordinates where ϕ and θ are the usual polar angles designating a direction in space, u and v are position coordinates in the projection plane and l represents the estimated position along a LOR estimated from TOF data. Further in the text when not specified otherwise the numerical values for angles will be given in degrees and for the lengths in centimeters; directions with $\theta = 90°$ are perpendicular to the scanner axis. The scatter data was calculated for a grid using a spacing of $\Delta\phi$ and $\Delta\theta$ of $5°$, Δu and Δv of 2 cm, and in the TOF case $\Delta l = 2\sigma_{\mathrm{tof}}$. In order to reduce the statistical fluctuations and obtain smoother scatter estimations, the counts were spread among nearby u, v and l bins using weights given by bell shaped kernels $b(x) = [1-(x/R)^2]^2$ along each direction with a maximum radius of $R = r \times \Delta$ where Δ is the grid spacing along a given dimension; we used $r \approx 2$ for u, v and l dimensions. In both cases the data lists contained about 15 million true events. We set an energy threshold E_{fit} of 440 keV for the $\varepsilon_E = 6.4\%$ case, and 450 keV for the $\varepsilon_E = 3.4\%$ case.

4. Results

In Figure 3 are shown the energy spectra for non-scattered (A) and scattered photons (B) for several positions in the data space, and the overall average, determined from Monte Carlo data for the non-TOF and 15% FWHM energy resolution case. One can notice that while the low energy region of the B spectra shape varies significantly with the position, the high energy portion is more stable, hence for a high enough energy threshold E_{fit}, a unique B shape will represent a robust approximation.

The plots in Figure 4 (non-TOF) and Figure 5 (TOF) show a good agreement between the estimated scatter components and the Monte Carlo data. In the TOF case one can notice the significant variation of the scatter with the TOF position.

In Figure 6 and Figure 7 are presented comparisons between the dual energy window (DEW), moments (Mt) and the maximum likelihood methods (ML). One can notice the good performance of the Mt and ML methods compared with the DEW method. This is not surprising given the simplicity of the DEW method and its very rough consideration of the spectral shapes. The DEW method was applied according to the description in Appendix B.1 with energy windows $\mathcal{E}_0 \equiv [511\,\mathrm{keV}, E_{\mathrm{max}}]$ and $\mathcal{E}_1 \equiv [E_{\mathrm{fit}}, 480\,\mathrm{keV}]$.

In order to further test our concept we adapted a recently developed TOF list-mode image reconstruction algorithm in order to accept energy dependent correction terms and coefficients. Aspects of the reconstruction algorithm used here were recently described in (Popescu & Lewitt 2004, Popescu et al. 2004). The reconstructed images were represented using 4 mm voxels. We used iterations over 12 subsets made from consecutive events in the list. Further in the text a complete pass through the data will be referred to as an iteration.

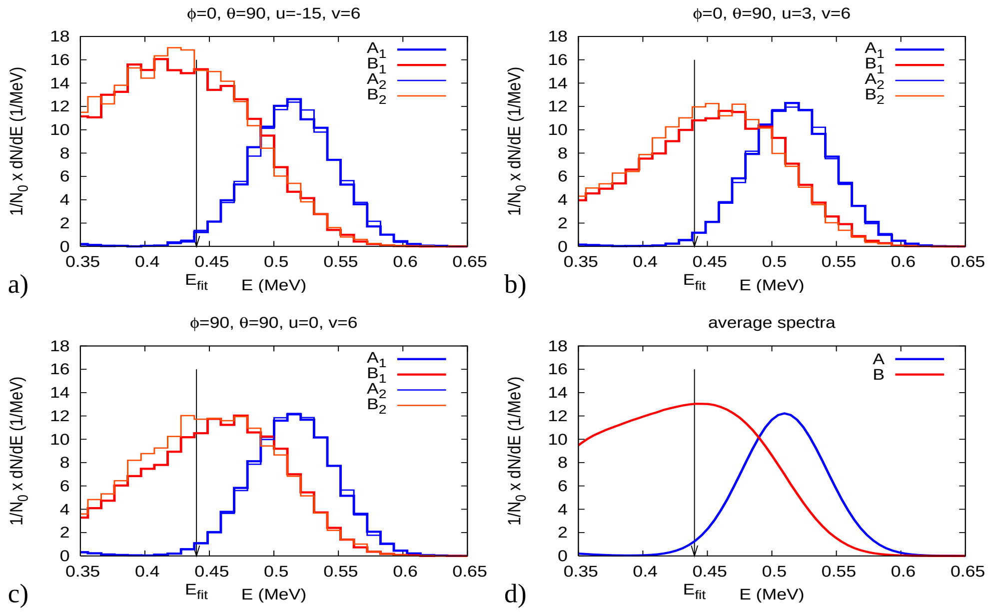

Figure 3. The energy spectra $A(E)$ for not scattered, and $B(E)$ for scattered PET photons at several positions in the data space as determined from Monte Carlo data for the non-TOF and 15% relative FWHM energy resolution case. The spectra are normalized to unity for $E > E_{\rm fit}$ region. The a), b) and c) cases correspond to the projections along the lines indicated by a), b), c) respectively in the left image in Figure 2. In the d) case are shown the averages over all of the data domain that are used for the scatter estimation.

We performed a multi-realization study using a cylindrical phantom of 32 cm diameter symmetrically placed along the axis of the scanner. The phantom features were placed in two slices each at 2.5 cm distance from the scanner central slice. In each slice we have considered 10 identical hot spheres uniformly placed on the circle of radius 8 cm. In this arrangement the distance between hot spheres is greater than 5 cm, therefore the correlation between spheres in the same image is small enough to enable us to consider each reconstructed sphere as a independent realization. By this means in each reconstructed volume we will obtain 20 realizations of each type of feature and with a few reconstructed images (3–5) we obtain a statistically significant number of feature realizations.

In the case of small features with size comparable with the image voxel size (or blob grid size), for a particular position the possible realizations depend on the particular approximation of the feature at that specific grid position. Instead of having multiple realizations of a unique feature, we have multiple realizations of multiple distinct approximations of that feature. This fact can lead to noticeable inconsistencies in statistical results. In order to avoid this problem we have introduced small fluctuations of the sphere positions as well as small random shifts of the reconstructed image grid.

Four image reconstruction cases were compared: no scatter correction (sn), scatter events removed (sr), energy dependent scatter correction (sr), and static (not energy dependent) scatter correction (ss). In the (sn) case all events with energies above the reconstruction

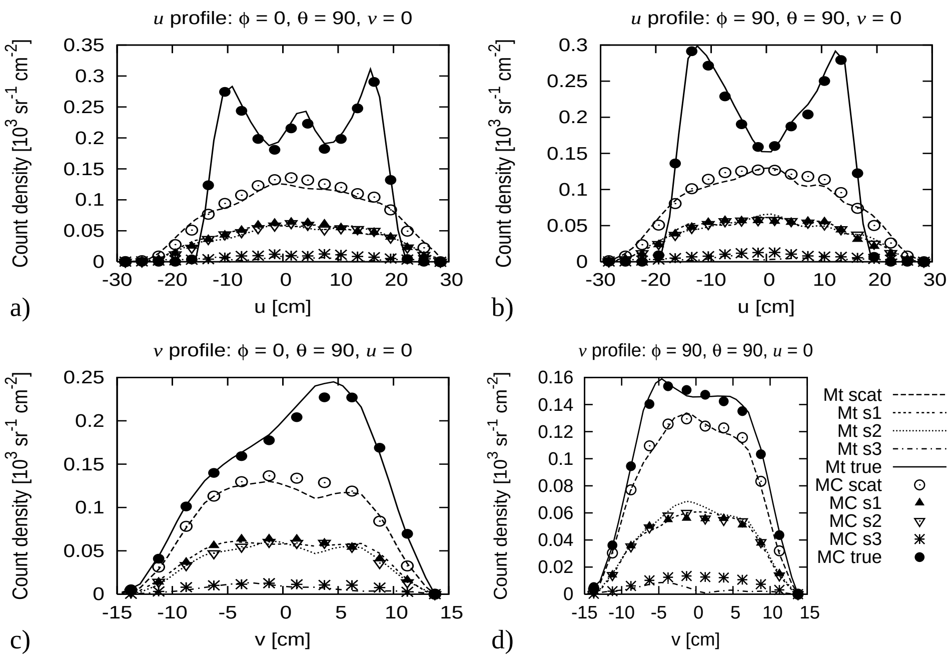

Figure 4. Comparison between the scatter estimates obtained with the energy based moments method (Mt) and the Monte Carlo data (MC), for a non-TOF case and energy resolution $\varepsilon_E = 6.4\%$ (15% relative FWHM). The cases a) and b) represent two orthogonal projections of the central slice, while c) and d) represent two axial profiles. In the legend s_1, s_2 are the partial scatter count densities in the case when the first, respectively the second, PET photon is scattered while the other not, and s_3 corresponds to the case with both photons scattered.

threshold were used and no scatter correction was applied. In the (sr) case the events marked as scattered were not used. In the other two cases we used scatter corrections estimated with the energy based approach applying the moments method in similar conditions as in the tests presented above, with $E_{\text{fit}} = E_{\text{rec}} = 450\,\text{keV}$. In the (se) case we used the equation (13) with the prior estimated scatter coefficients $\{\sigma_{lk}\}$. In the (ss) case we applied the scatter correction by considering only the estimated total scatter distribution in each point $g_s(\mathbf{y}^g) = \sum_{k=1}^{K} b_k(\mathbf{y}^g) \sum_{l=1}^{3} \sigma_{lk}$, obtained by eliminating from equation (13) all the $e_l(\mathbf{y}^e)$ factors; this is the more traditional way of applying the scatter correction.

Multiple image realizations were obtained by using distinct data sets containing 12 million true coincidences, with an energy resolution of $\varepsilon_E = 3.4\%$ (8% relative FWHM) for non-TOF and TOF cases with $\sigma_{\text{tof}} = 3\,\text{cm}$ (≈ 470 ps FWHM timing resolution). Hot spheres with contrast 4:1 and 0.5 cm radius were considered.

For each realization of a sphere of radius $R = 0.5\,\text{cm}$ we measured the signal S as the image average inside the sphere and the background B as the image average outside the sphere and within a radius $2R + 0.5\,\text{cm}$ from the center of the sphere. The contrast is defined as $C = \frac{|S-B|}{B}$ and the contrast recovery coefficient is $\text{CRC} = \frac{C}{C_0} \times 100$, where C_0 is the original phantom contrast. For each case we computed the average CRC ($\overline{c_1}$) over all N_1

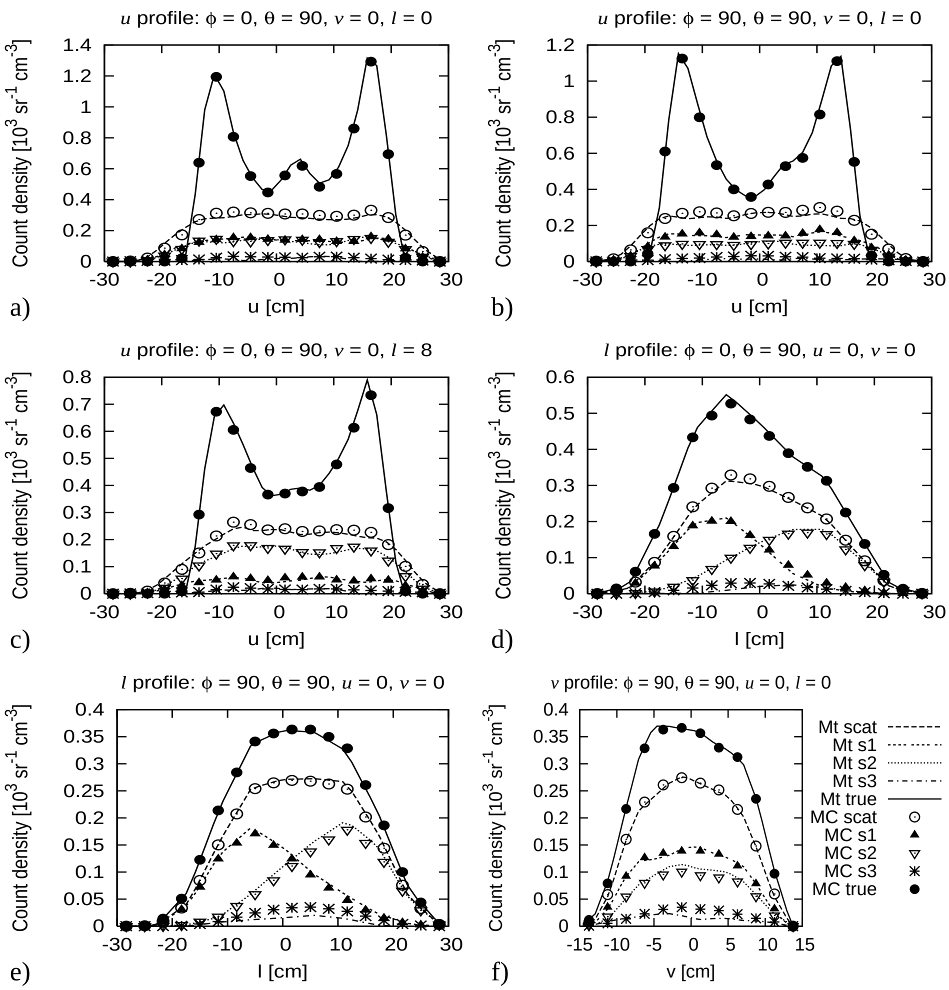

Figure 5. Comparison between the scatter estimates obtained with the energy based moments method (Mt) and the Monte Carlo data (MC), as in Figure 4 but this time a TOF case with $\sigma_{\rm tof} = 3.0\,{\rm cm}$ ($\approx$ 470 ps FWHM timing resolution) and energy resolution $\varepsilon_E = 3.4\%$ (8% relative FWHM). The cases a) and b) represent two orthogonal projections of the central slice and central TOF position, case c) is the same as a) only for a different TOF position $l = 8$, the cases d) and e) correspond to two longitudinal profiles, along TOF data (l), for two perpendicular LORs, while the case f) represents an axial profile. The values shown correspond to $E \geq E_{\rm fit} = 450\,{\rm keV}$.

realizations obtained, the standard deviation σ_{c_1}, the standard error $\sigma_{\overline{c_1}} = \sigma_{c_1}/\sqrt{N_1}$ and the uncertainty of the standard deviation estimation $\Delta\sigma_{c_1} \approx \sigma_{c_1}/\sqrt{2N_1-2}$.

For the noise estimation we computed the same quantities ($\overline{c_0}$, σ_{c_0}, $\sigma_{\overline{c_0}}$, and $\Delta\sigma_{c_0}$) for similar spherical regions at equivalent fixed position but with the signal absent (background only regions). As in the case of non pre-whitening matched filter observer (NPWMF) the noise was evaluated as $n = \sqrt{\frac{1}{2}(\sigma_{c_1}^2 + \sigma_{c_0}^2)}$, with the uncertainty estimate $\Delta n = \sqrt{(\sigma_{c_1}\Delta\sigma_{c_1})^2 + (\sigma_{c_0}\Delta\sigma_{c_0})^2}/\sqrt{\frac{1}{2}(\sigma_{c_1}^2 + \sigma_{c_0}^2)}$.

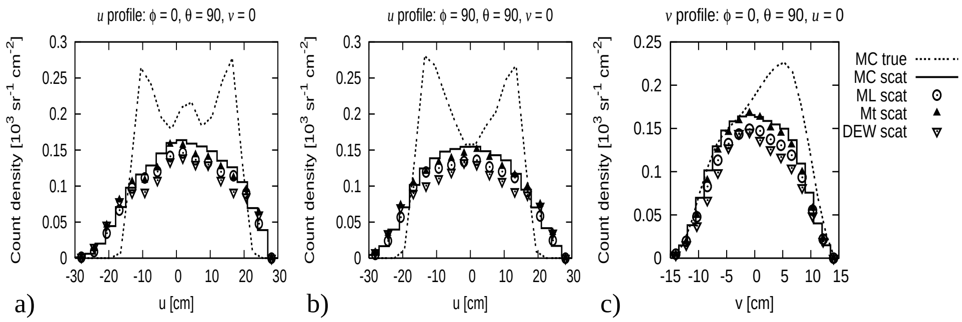

Figure 6. Comparison between the energy based scatter estimation using dual energy window (DEW), moments (Mt), and maximum likelihood (ML) methods for the same no-TOF cases a, b and c as in Figure 4.

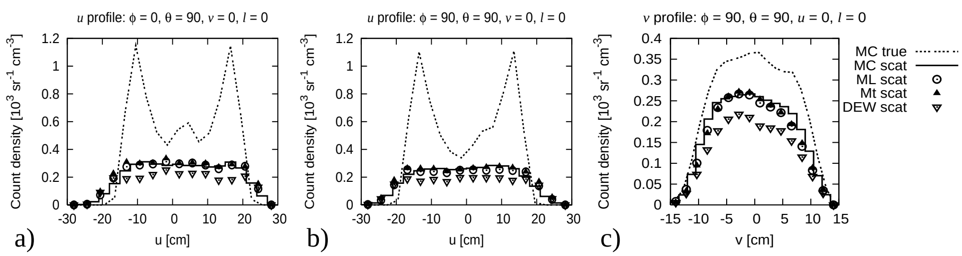

Figure 7. Comparison between the energy based scatter estimation using dual energy window (DEW), moments (Mt), and maximum likelihood (ML) methods for the same TOF cases a, b and f as in Figure 5.

The contrast as function of iteration number and contrast versus noise plots presented in Figure 8 show that the scatter removed (sr) ideal method systematically performs better than the others, as expected, while the no scatter correction (sn) case is relatively largely biased if compared with any other method. The energy dependent correction (se) represents an improvement over the static scatter correction (ss) both in the contrast versus iteration and in the noise versus contrast plots.

5. Discussion and conclusions

The energy dependent reconstruction method compared with the simple (not energy dependent) scatter correction approach, produces images with improved overall quality, having higher contrast and at the same time lower noise. The scatter, as well as other physical effects that represent departures from the idealized tomographic model, are often seen only as bias terms that can be compensated somehow even after an image was already reconstructed. We emphasize here that the scatter is a problem that affects the overall quality of the activity distribution estimation, both the bias and the variance, and therefore it should be, as much as computationally feasible, fully treated in the model used by the reconstruction algorithm.

Given the theoretical generality of the proposed approach, the main purpose of this

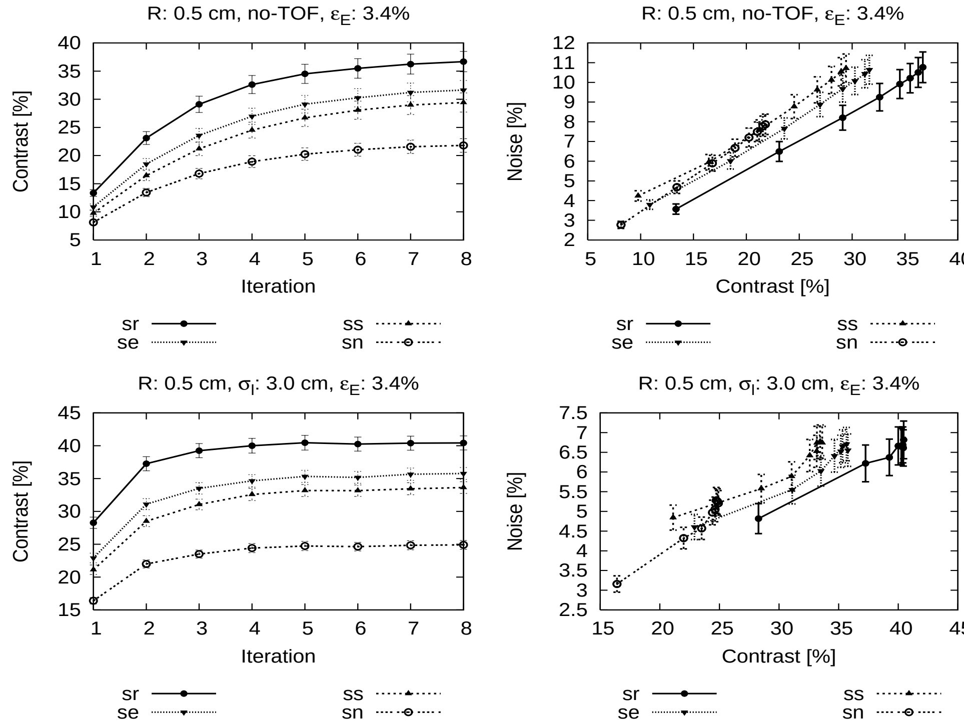

Figure 8. Comparisons between contrast as function of iteration number and contrast versus noise plots for no scatter correction (sn) case, scatter events removed (sr) case, static scatter correction (ss) case, and the energy dependent scatter correction (se) method. Top plots show a non-TOF case, bottom plots show a TOF case.

paper was to present and demonstrate the principles of the methods without attempting to find the optimal parameterizations and application conditions for specific practical situations. Therefore, for simplicity, we have neglected a number of real data factors.

We have not considered the effect of the dispersion of the energy gains, therefore the conclusions relative to the lowest quality of the energy resolution for which the proposed methods are reliable should not be directly extrapolated to the real situations. At the same time, in general the reported energy resolutions of the PET scanners are the system averages that include the system wide variability, with values significantly worse than the energy resolution measured for a single block on the laboratory bench. The energy resolution values that we used in our simulation study can be seen as examples of these system wide values. Also, because we estimate the scatter for a coarse grid and use overlapping kernels, for one set of σ_{lk} coefficients we have contributions from a large number of LORs. Therefore, it is not unreasonable to expect that the average energy resolution will be a good representation of the local one. In this context it should be noted that the methods we propose allow for the use of position specific energy spectra components.

From the numerical experiments we can conclude that the energy based scatter estimation method is feasible in the cases when the energy resolution is good enough so that the photopeak shape A does not spread over the less stable portion of the shape B. Better energy resolution leads to better estimations.

The main difference between the energy based scatter estimation part of our approach and the previous dual energy window (DEW) type approaches applied in PET, is found in the four energy component model we are proposing, which leads to the estimation of four independent coefficients. The previous DEW methods in PET attempted to empirically estimate the scatter fraction by establishing a proportionality relation of some sort between the number of counts collected in the energy windows used, overlooking the fact that the total scatter fraction in reality is the sum of at least three independent parameters. The need for multiple components is clearly revealed when symmetries are broken, as best shown in the TOF depth profiles d) and e) in Figure 5. When used with the moments method or the maximum likelihood method, our approach benefits also from the more detailed energy information that comes with list-mode data. Another benefit of working with an underlying energy model arises from the fact that if the results are obtained for one energy range they can be easily renormalized for a different energy range, as shown in equation (3).

The energy-based image reconstruction approach proposed here uses scatter corrections that vary both with the LOR position (due to the variation with position of the σ_{lk} coefficients) and the energy (due to consideration of the energy dependent factors), while in the methods using energy dependent weights discussed in the introduction (Levkovitz et al. 2001, Chen et al. 2003, Chen 2005) the variation with the LOR position is neglected. It must be noted that using a position independent weight table leads to decreased contrast in the hot regions where the local trues fraction is above the average, an undesirable effect when the goal is to detect these same small hot spots (small lesions). In our approach we incorporate the scatter correction as an additive term that naturally integrates with the randoms correction term and avoids the general drawback of the correction by weights approach that leads to an alteration of the true probabilities.

Compared with the current most widely used image based simulation method for scatter estimation (Ollinger 1996b, Watson et al. 1996, Watson 2000, Werling et al. 2002, Accorsi et al. 2004) the energy based approach inherently considers the scatter from outside the field of view. Another advantage is that it estimates the scatter distributions in absolute values, eliminating the need for the error prone tail fitting scaling. The proposed methods also directly provide the position and energy dependent scatter corrections necessary for the energy-dependent image reconstruction. The single scatter simulation method can be modified in order to provide two of the components (single scatter fractions for each photon in a PET pair). With the third (double scatter fraction) neglected or empirically estimated from the first two, one can still use the model we propose to perform an energy dependent image reconstruction. However, the energy based methods operate by performing only a simple statistical analysis of the population of events recorded, and therefore they are considerably more convenient and flexible.

Acknowledgments

This work was supported in part by the National Institutes of Health under grant R33-EB001684.

Appendix A. Energy based scatter estimation in the single photon case

In these appendices, we present derivations of the dual energy window method, both for SPECT and PET cases based on the proposed energy models (1) and (2). We also present the application of another two standard statistical estimation methods, the moments method and the maximum likelihood expectation maximization method.

Appendix A.1. The dual energy window method for single photons

The dual energy window applied in SPECT is just a particular case of energy spectrum fitting applied to events binned in a two intervals histogram. Let $\mathcal{E}_0$, $\mathcal{E}_1$ be two energy windows (or more generally two energy domains). We define $\alpha_0 = \int_{\mathcal{E}_0} A(E)\mathrm{d}E$, $\alpha_1 = \int_{\mathcal{E}_1} A(E)\mathrm{d}E$, and $\beta_0 = \int_{\mathcal{E}_0} B(E)\mathrm{d}E$, $\beta_1 = \int_{\mathcal{E}_1} B(E)\mathrm{d}E$. If we detect N_0 photons with energy $E \in \mathcal{E}_0$ and N_1 with energy $E \in \mathcal{E}_1$, then in order to have equation (1) fit the data we obtain the equation system

$$\left\{ \begin{array}{rcl} t\alpha_0 + s\beta_0 & = & N_0 \\ t\alpha_1 + s\beta_1 & = & N_1 \end{array} \right. \tag{A.1}$$

from which results the total number of the scattered photons

$$s = \frac{\alpha_0 N_1 - \alpha_1 N_0}{\alpha_0 \beta_1 - \alpha_1 \beta_0} \tag{A.2}$$

Appendix A.2. The moments method

Instead of tallying the number of events in separate energy windows we can estimate moments of the energy distributions $\hat{q}_m = \sum_{n=1}^{N} E_n^m$. This quantity should be equal to the theoretical value

$$q_m = \int_{\mathcal{E}} E^m S(E)\mathrm{d}E = ta_m + sb_m \tag{A.3}$$

where $a_m = \int_{\mathcal{E}} E^m A(E)\mathrm{d}E$ and $b_m = \int_{\mathcal{E}} E^m B(E)\mathrm{d}E$. For two distinct moments m and n, proceeding as in the dual energy window case, we obtain

$$s = \frac{a_m \hat{q}_n - a_n \hat{q}_m}{a_m b_n - a_n b_m} \tag{A.4}$$

We can choose among various moment orders m and n or linear combinations of them in order to have more stable estimates and assure the best discrimination.

Appendix A.3. The ML-EM method

For a given region in the data space the events energy distribution is $\frac{tA(E)+sB(E)}{\Lambda}$, where $\Lambda = t + s$ is the expected total number of counts. For a given data set $\{E_1, \ldots, E_N\}$ the likelihood function is

$$\mathcal{L}(E_1, \ldots, E_N | t, s) = P(N|\Lambda) \prod_{n=1}^{N} \frac{tA(E_n) + sB(E_n)}{\Lambda} \tag{A.5}$$

where $P(N|\Lambda) = \frac{\Lambda^N}{N!} \mathrm{e}^{-\Lambda}$ is the probability of detecting N events following a Poisson distribution. The maximum likelihood point satisfies $\frac{\partial \ln \mathcal{L}}{\partial t} = \frac{\partial \ln \mathcal{L}}{\partial s} = 0$, and can be reached using the following maximum likelihood expectation maximization (ML-EM) iterative update equations:

$$t^{(m+1)} = t^{(m)} \sum_{n=1}^{N} \frac{A(E_n)}{t^{(m)} A(E_n) + s^{(m)} B(E_n)} \tag{A.6}$$

$$s^{(m+1)} = s^{(m)} \sum_{n=1}^{N} \frac{B(E_n)}{t^{(m)} A(E_n) + s^{(m)} B(E_n)} \tag{A.7}$$

Appendix B. Energy based scatter estimation for PET

Appendix B.1. The 2×2 energy windows method

The extension of the dual energy window method to the case of photon pairs gives us a 2×2 energy window method (two energy windows for each photon). Let be N_0 the observed number of photon pairs with both energies in the window $\mathcal{E}_0$, N_1 the number of events with $(E_1, E_2) \in \mathcal{E}_0 \times \mathcal{E}_1$, N_2 the number events with $(E_1, E_2) \in \mathcal{E}_1 \times \mathcal{E}_0$ and N_3 the number of events with $(E_1, E_2) \in \mathcal{E}_1 \times \mathcal{E}_1$. The number of true and scattered photons are obtained by solving the following system of equations

$$\begin{cases} \sigma_0 \alpha_0^2 & + & \sigma_1 \alpha_0 \beta_0 & + & \sigma_2 \beta_0 \alpha_0 & + & \sigma_3 \beta_0^2 & = & N_0 \\ \sigma_0 \alpha_0 \alpha_1 & + & \sigma_1 \alpha_0 \beta_1 & + & \sigma_2 \beta_0 \alpha_1 & + & \sigma_3 \beta_0 \beta_1 & = & N_1 \\ \sigma_0 \alpha_1 \alpha_0 & + & \sigma_1 \alpha_1 \beta_0 & + & \sigma_2 \beta_1 \alpha_0 & + & \sigma_3 \beta_1 \beta_0 & = & N_2 \\ \sigma_0 \alpha_1^2 & + & \sigma_1 \alpha_1 \beta_1 & + & \sigma_2 \beta_1 \alpha_1 & + & \sigma_3 \beta_1^2 & = & N_3 \end{cases} . \tag{B.1}$$

In order to minimize the estimation errors one should choose an energy windows arrangement that maximizes the determinant of the equation system. As in the single photon case, this is achieved when we have a large difference between the contributions of the two distributions in the two energy windows, e.g. $\alpha_0 \beta_1 \gg \alpha_1 \beta_0$.

Appendix B.2. The moments method

Similar to the single photon case, we can estimate the moment of order (m, n) of the energy distribution of the photon pairs by computing the quantity $\hat{q}_{mn} = \sum_{i=1}^{N} E_{1,i}^m E_{2,i}^n$, which should match the theoretical value

$$
\begin{aligned}
q_{mn} &= \int_{\mathcal{E}}\int_{\mathcal{E}} E_1^m E_2^n C(E_1, E_2)\mathrm{d}E_1\mathrm{d}E_2 \\
&= \sigma_0 a_m a_n + \sigma_1 a_m b_n + \sigma_2 b_m a_n + \sigma_3 b_m b_n
\end{aligned}
\tag{B.2}
$$

By conveniently selecting four combinations (m, n) we can obtain an equation system as in the case of the dual energy window method (B.1).

Appendix B.3. The ML-EM method

Using notations similar to those used in section 2.3, for a given region in the data space the expected number of counts is $\Lambda = \sum_{l=0}^{3} \sigma_l$. The energy distribution is $\frac{1}{\Lambda}\sum_{l=0}^{3} \sigma_l e_l(\mathbf{y}^e)$. For a data set $Y_E \equiv \{\mathbf{y}_1^e, \ldots, \mathbf{y}_N^e\}$ the likelihood function is

$$
\mathcal{L}(Y_E|\sigma) = \frac{\Lambda^N}{N!}\mathrm{e}^{-\Lambda}\prod_{n=1}^{N}\frac{1}{\Lambda}\sum_{l=0}^{3}\sigma_l e_l(\mathbf{y}_n^e)
\tag{B.3}
$$

Maximum likelihood corresponds to the point for which $\frac{\partial \ln \mathcal{L}(Y_E|\sigma)}{\partial \sigma_l} = 0$, and can be reached using the following ML-EM iterative update equation

$$
\sigma_l^{(m+1)} = \sigma_l^{(m)} \sum_{n=1}^{N} \frac{e_l(\mathbf{y}_n^e)}{\sum_{l'=0}^{3} \sigma_{l'}^{(m)} e_{l'}(\mathbf{y}_n^e)}\,, \quad l = 0, 1, 2, 3
\tag{B.4}
$$

Appendix B.4. Considerations about the practical implementation

For simplicity we have presented all the above methods for the case of an individual data space position (a LOR vicinity) where are processed only the events from the list that are relevant for that position. In the practical situations we need to determine the entire ensemble of scatter coefficients $\{\sigma_{lk}\}$, with k covering the whole data space, as defined in section 2.3.

In all methods, for each k position index is necessary to evaluate an additional set of four variables. For example in the case of 2×2 energy windows method we have the measured $(N_{0k}, N_{1k}, N_{2k}, N_{3k})$ set, in the case of moments method we have $(q_{0k}, q_{1k}, q_{2k}, q_{3k})$ set with the components defined for four different (m, n) moments pairs. And, in the case of ML-EM method we have $(c_{0k}, c_{1k}, c_{2k}, c_{3k})$ sets, with $c_{lk} = \sum_{n=1}^{N} \frac{e_l(\mathbf{y}_n^e)}{\sum_{l'=0}^{3} \sigma_{l'k}^{(m)} e_{l'}(\mathbf{y}_n^e)}$, $l = 0, 1, 2, 3$ that are used for each iteration update (B.4). The ensemble of these additional 4-variable sets can be evaluated together in an event driven manner through one list pass. It is not necessary to bin or rearrange the events in the list. In the case of 2×2 energy window and moments methods the calculations are completed by solving for each k index a 4-equations system to obtain the scatter coefficients. In the case of ML-EM several iterations are necessary. The size of the scatter coefficients arrays are of easily manageable proportions, due to the coarse sampling. For the examples presented in section 3 the size was about 1 MB for the non-TOF case, and about 12 MB for the TOF case.

In order to obtain good statistics for the fit performed at each k position index, we can split an event between the nearby grid positions by assigning weights computed using distance dependent kernels, as it is exemplified in the simulation study in section 3. Also, in the case when the continuous representation in the form of equation (6) is used, for the ML-EM method we have the equation (14) with the modifications discussed in subsection 2.4.

References

Accorsi, R., Adam, L.-E., Werner, M. E. & Karp, J. S. (2004). Optimization of a fully 3D single scatter simulation algorithm for 3D PET, *Phys. Med. Biol.* **49**(12): 2577–2598.

Adam, L.-E., Karp, J. S. & Freifelder, R. (2000). Energy-based scatter correction for 3-D PET scanner using NaI(Tl) detectors, *IEEE Trans. Med. Imag.* **19**(5): 513–521.

Barrett, H. H., White, T. & Parra, L. C. (1997). List-mode likelihood, *J. Opt. Soc. Am. A* **14**(11): 2914 – 2923.

Bendriem, B., Trébossen, R., Frouin, V. & Syrota, A. (1993). A PET scatter correction using simultaneous acquisition with low and high lower energy thresholds, *IEEE Med. Img. Conf. Rec.*, San Francisco, pp. 1779–1783.

Bentourkia, M., Msaki, P., Cadorette, J. & Lecomte, R. (1995a). Energy dependence of scatter components in multispectral PET imaging, *IEEE Trans. Med. Imag.* **14**(1): 138–145.

Bentourkia, M., Msaki, P., Cadorette, J. & Lecomte, R. (1995b). Object and detector scatter function dependence on energy and position in high resolution PET, *IEEE Trans. Nucl. Sci.* **42**(4): 1162–1167.

Chen, H.-T. (2005). *An Energy-Based Method for PET Scatter Correction*, PhD thesis, University of Chicago, Chicago, Illinois.

Chen, H.-T., Kao, C.-M. & Chen, C.-T. (2003). A fast, energy-dependent scatter reduction method for 3D PET imaging, *IEEE Nucl. Sci. Symp. Conf. Rec.*, Vol. 4, pp. 2630–2634.

Cutler, P. D. & Hoffman, E. J. (1993). Limitations of the dual-energy window scatter corrections for positron emission tomography, *J. Nucl. Med.* **34**(5): 136.

El Fakhri, G., Kijewski, M. F., Maksud, P. & Moore, S. C. (2003). The effects of compensation for scatter, lead X-rays, and high-energy contamination on tumor detectability and activity estimation in Ga-67 imaging, *IEEE Trans. Nucl. Sci.* **50**(3): 439–445.

Farncombe, T. H., Gifford, H. C., Narayanan, M. V., Pretorius, P. H., Frey, E. C. & King, M. A. (2004). Assessment of scatter compensation strategies for ^{67}Ga SPECT using numerical observers and human LROC studies, *J. Nucl. Med.* **45**(5): 802–812.

Grootoonk, S., Spinks, T. J., Sashin, D., Spyrou, N. M. & Jones, T. (1996). Correction for scatter in 3D PET using dual energy window method, *Phys. Med. Biol.* **41**: 2757–2774.

Harrison, R. L., Haynor, D. R. & Lewellen, T. K. (1991). Dual energy window scatter corrections for positron emission tomography, *IEEE Med. Img. Conf. Rec.*, pp. 1684–1688.

Haynor, D. R., Kaplan, M. S., Miyaoka, R. S. & Lewellen, T. K. (1995). Multiwindow scatter correction techniques in single-photon imaging, *Med. Phys.* **22**(12): 2015–2024.

Holdsworth, C. H., Levin, C. S., Janecek, M., Dahlbom, M. & Hoffman, E. J. (2002). Performance analysis of an improved 3-D PET Monte Carlo simulation and scatter correction, *IEEE Trans. Nucl. Sci.* **49**(1): 83–89.

Jaszczak, R. J., Greer, K. L., Floyd, C. E., Harris, C. C. & Coleman, R. E. (1984). Improved SPECT quantification using compensation for scattered photons, *J. Nucl. Med.* **25**: 893–900.

Karp, J. S., Kuhn, A., Perkins, A. E., Surti, S., Werner, M. E., Daube-Witherspoon, M. E., Popescu, L. M., Vandenberghe, S. & Muehllehner, G. (2005). Characterization of TOF PET scanner based on lanthanum bromide, *IEEE Nucl. Sci. Symp. Conf. Rec.* M04-8.

King, M. A., Hademenos, G. J. & Glick, S. J. (1992). A dual-photopeak window method for scatter correction, *J. Nucl. Med.* **33**: 605–612.

Koral, K. F., Wang, X. Q., Rogers, W. L., Clinthorne, N. H. & Wang, X. H. (1988). SPECT Compton-scattering correction by analysis of energy spectra, *J. Nucl. Med.* **29**(2): 195–202.

Levkovitz, R., Falikman, D., Zibulevsky, M., Ben-Tal, A. & Nemirovski, A. (2001). The design and implementation of COSEM, an iterative algorithm for fully 3-D listmode data, *IEEE Trans. Med. Imag.* **20**(7): 633 – 642.

Meikle, S. R. & Badawi, R. D. (2003). Quantitative techniques in positron emission tomography, *in* P. E. Valk, D. L. Bailey & D. W. Townsend (eds), *Positron emission tomography: basic science and clinical practice*, Springer.

Narayanan, M. V., Pretorius, P. H., Dahlberg, S. T., Leppo, J. A., Botkin, N., Krasnow, J., W. Berndt, W., Frey, E. C. & King, M. A. (2003). Evaluation of scatter compensation strategies and their impact on human detection performance Tc-99m myocardial perfusion imaging, *IEEE Trans. Nucl. Sci.* **50**(5): 1522–1527.

Ogawa, K., Harata, Y., Ichihara, T., Kubo, A. & Hashimoto, S. (1991). A practical method for position-dependent compton-scatter correction in single-photon emission ct, *IEEE Trans. Med. Imag.* **10**: 408–412.

Ollinger, J. M. (1996a). The effect of energy threshold on image variance in fully 3D PET, *in* P. Grangeat & J.-L. Amans (eds), *Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, Kluwer, pp. 269–276.

Ollinger, J. M. (1996b). Model-based scatter correction for fully 3D PET, *Phys. Med. Biol.* **41**: 153–176.

Popescu, L. M. (2000). A computer code package for Monte Carlo electron-photon transport simulation. Comparisons with experimental benchmarks, *Nucl. Instr. Meth. Phys. Res. B* **161-163**: 318–322.

Popescu, L. M. (2003). A geometry modelling system for ray tracing or particle transport Monte Carlo simulation, *Comp. Phys. Comm.* **150**(1): 21–30.

Popescu, L. M. & Lewitt, R. M. (2003). A versatile approach for Monte Carlo simulation of tomographic systems, *IEEE Nucl. Sci. Symp. Conf. Rec.*, Vol. 4, pp. 2785 – 2788.

Popescu, L. M. & Lewitt, R. M. (2004). Ray tracing through a grid of blobs, *IEEE Nucl. Sci. Symp. Conf. Rec.*, Vol. 6, pp. 3983 – 3986.

Popescu, L. M., Matej, S. & Lewitt, R. M. (2004). Iterative image reconstruction using geometrically ordered subsets with list-mode data, *IEEE Nucl. Sci. Symp. Conf. Rec.*, Vol. 6, pp. 3536 – 3540.

Qi, J. & Huesman, R. H. (2002). Scatter correction for positron emission mammography, *Phys. Med. Biol.* **47**: 2759–2771.

Reader, A. J., Erlandsson, K., Flower, M. A. & Ott, R. J. (1998). Fast accurate iterative reconstruction for low-statistics positron volume imaging, *Phys. Med. Biol.* **43**: 835 – 846.

Shao, L., Freifelder, R. & Karp, J. S. (1994). Triple energy window scatter correction technique in PET, *IEEE Trans. Med. Imag.* **13**(4): 641–648.

Snyder, D. L. & Politte, D. G. (1983). Image-reconstruction from list-mode data in an emission tomography system having time-of-flight measurements, *IEEE Trans. Nucl. Sci.* **30**(3): 1843–1849.

Wang, X. & Koral, K. F. . (1992). A regularized deconvolution-fitting method for Compton-scatter correction in SPECT, *IEEE Trans. Med. Imag.* **11**: 351–360.

Watson, C. C. (2000). New, faster, image-based scatter correction for 3D PET, *IEEE Trans. Nucl. Sci.* **47**(4): 1587–1594.

Watson, C. C., Newport, D. & Casey, M. E. (1996). A single scatter simulation technique for scatter correction in 3D PET, *in* P. Grangeat & J.-L. Amans (eds), *Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, Kluwer, pp. 255–268.

Werling, A., Bublitz, O., Doll, J., Adam, L.-E. & Brix, G. (2002). Fast implementation of the single scatter simulation algorithm and its use in iterative image reconstruction of PET data, *Phys. Med. Biol.* **47**: 2947–2960.

Zaidi, H. (1999). Relevance of accurate Monte Carlo modeling in nuclear medical imaging, *Med. Phys.* **26**(4): 574–608.

Zaidi, H. & Koral, K. F. (2004). Scatter modelling and compensation in emission tomography, *Eur. J. Nucl.*

Med. **31**(5): 761–782.