ML estimation of the scatter scaling in TOF PET

Michel Defrise[1], Koen Salvo[1], Ahmadreza Rezaei[2], Johan Nuyts[2], Vladimir Panin[3], Michael Casey[3]

Abstract. **In positron emission tomography a scatter estimate is calculated prior to image reconstruction and is then added at each MLEM iteration to the estimated true counts to yield the estimated prompts. The absolute scale of the scatter is usually estimated by matching the scatter estimate to the data for those lines of response that do not intersect the patient. The goal of this work is to explore alternatives to this tail fitting method. Pursuing the method introduced by V. Panin at the 2012 IEEE Medical Imaging Conference for the case with known attenuation, we estimate the scatter scaling factor as an additional voxel, which is updated at each iteration at the same time as the activity. This method is extended for applications with unknown attenuation. This preliminary study uses a single-scatter estimate and simulated 2D data.**

I. Introduction

In positron emission tomography the scatter background may represent over 50 % of the measured (prompt) data in whole-body studies. Because it cannot easily be modeled in the system matrix, the scatter background is usually estimated prior to image reconstruction. It is then considered as a fixed term, which is added at each MLEM iteration to the estimated true counts to yield the estimated prompts.

Three types of methods have been explored to estimate the scatter background [1], [2]. The first one approximates the scatter by convolving the measured data with an empirical kernel. The second approach exploits measurements in additional energy windows [3], [4] and is similar to the standard technique used in SPECT. An attractive implementation measures data in an energy window with a very high lower energy threshold to provide a little biased, though noisy, estimate of the true data, which after smoothing allows estimating the scatter contribution to the photo-peak data [5]. The third approach estimates the scatter background based on the Klein-Nishina cross-section, on a CT attenuation map and on an initial estimate of the tracer uptake [6]–[8]. Current implementations of this model-based method only calculate the single-scatter component, this can be done efficiently by combining sparse sampling and optimized programming [7], [9], [10]. Energy- and model-based methods can be combined as proposed in [11].

All these methods require a multiplicative scaling factor to match the amount of scatter in the photo-peak data. Despite a few attempts [7], obtaining this factor a priori is difficult, since it depends on several effects, including scatter caused by activity outside the field-of-view, multiple scatters, the energy dependence of the detector sensitivity and also the uncertainty in the lower energy-level discriminator (LLD). The standard method to estimate the scaling factor, referred to as the "tail fitting" method, consists in matching the scatter estimate to the data for those lines of response (LORs) that do not intersect the patient. These "out of patient" LORs contain only scatter and random events (see e.g. [9]), and allow direct estimation of the scatter scaling factor if the mean random background is known. The tail fitting method has limitations in some applications where the out of patient LORs are noisy (e.g. for dynamic studies [12]) or scarce (e.g. for large patients scanned with their arms within the field-of-view).

The goal of this work is to explore alternatives to tail fitting both for the case with known attenuation (CT-PET) as for the case where the activity and attenuation are estimated simultaneously (CT-less TOF PET). When the energy of the photons is measured, the ultimate approach consists in modeling the physics of scattering within the system matrix instead of considering the scatter as a fixed background. This was demonstrated by [13], [14] for the case with known attenuation. For CT-less PET, recent works explore the use of the Compton scattering cross-section to estimate the attenuation map [15] from list-mode data. These methods exploit the information that scatter events contain on the tissue density and activity, but are numerically complex. We consider here simpler methods. For known attenuation, Thielemans et al [16] estimate the scatter scaling factor by least-squares fitting using the whole data instead of only out of patient LORs. This requires a prior estimate of the activity, and the method therefore alternates this step with a standard ML reconstruction of the activity at fixed scatter scale. In this paper we pursue the method introduced in [17], which is based on the observation that the scatter scaling factor can be handled very simply as an additional image voxel, which is updated at each iteration at the same time as the activity. We investigate this technique for applications with known and unknown attenuation. This preliminary study uses a standard single-scatter estimate and simulated 2D data.

II. Algorithms

Consider TOF PET data histogrammed in N LORs $i = 1, \ldots, N$ and T TOF bins $t = 1, \ldots, T$. The expectation of the data bin $y_{i,t}$ is modeled as

$$< y_{i,t} >= n_i a_i p_{i,t} + n_i \alpha s_{i,t} + r_i \qquad (1)$$

where a_i is the attenuation factor and α the scatter scale factor. We assume that the sensitivity $n_i \geq 0$, the scatter estimate

[1]Dept. of Nuclear Medicine, Vrije Universiteit Brussel, (Belgium), [2]Dept. of Nuclear Medicine, KULeuven, Leuven (Belgium), [3]Siemens Healthcare, Molecular Imaging, Knoxville (TN). This work was supported in part by K.U. Leuven under Grant GOA/11/006, in part by FWO Project G027514N, and in part by Siemens Healthcare.

$s_{i,t} \geq 0$, and the mean random $r_i \geq 0$ are all given. The expectation value of the unattenuated trues is

$$p_{i,t} = \sum_{j=1}^{M} c_{i,t,j} \lambda_j \tag{2}$$

with given matrix elements $c_{i,t,j}$ and λ_j is the tracer uptake in voxel j. Given the measured TOF-PET data $y_{i,t}$, we seek to maximize the log-likelihood $L(y, a, \lambda, \alpha) = \sum_{i,t}(- <y_{i,t}> + y_{i,t} \log(< y_{i,t} >)$ with respect to λ, α when attenuation is known, or with respect to a, λ, α for CT-less PET.

The data model (1) is linear both in the activity λ and in the scatter scale α, and α can therefore be viewed as an additional voxel $\lambda_0 = \alpha$, with as corresponding column of the system matrix $c_{i,t,0} = s_{i,t}/a_i$. For known attenuation, a straightforward extension of the ML-EM algorithm yields the following update at iteration k [17]:

$$\begin{aligned} \lambda_j^{(k+1)} &= \frac{\lambda_j^{(k)}}{\sum_{l,t'} n_l a_l c_{l,t',j}} \sum_{i,t} \frac{n_i a_i c_{i,t,j} y_{i,t}}{\hat{y}_{i,t}^{(k)}} \\ \alpha^{(k+1)} &= \frac{\alpha^{(k)}}{\sum_{l,t'} n_l s_{l,t'}} \sum_{i,t} \frac{n_i s_{i,t} y_{i,t}}{\hat{y}_{i,t}^{(k)}} \end{aligned} \tag{3}$$

with the estimated prompts

$$\hat{y}_{i,t}^{(k)} = n_i a_i p_{i,t}^{(k)} + n_i \alpha^{(k)} s_{i,t} + r_i \tag{4}$$

and $p_{i,t}^{(k)} = \sum_j c_{i,t,j} \lambda_j^{(k)}$.

For the CT-less case, the same idea leads to straightforward extensions of the MLAA [18], [19] and MLACF [20] algorithms, which simultaneously estimate the activity and, respectively, the attenuation image μ_j and the attenuation factors a_i. This paper presents a preliminary study of the MLACF extension, which alternates an update (see eqn (3,4) in [20]) of the activity and scatter scale,

$$\begin{aligned} \lambda_j^{(k+1)} &= \frac{C^{(k)} \lambda_j^{(k)}}{\sum_{l,t'} n_l a_l^{(k)} c_{l,t',j}} \sum_{i,t} \frac{n_i a_i^{(k)} c_{i,t,j} y_{i,t}}{\hat{y}_{i,t}^{(k)}} \\ \alpha^{(k+1)} &= \frac{\alpha^{(k)}}{\sum_{l,t'} n_l s_{l,t'}} \sum_{i,t} \frac{n_i s_{i,t} y_{i,t}}{\hat{y}_{i,t}^{(k)}} \end{aligned} \tag{5}$$

and P updates of the attenuation factors, with $a_i^{(k,0)} = a_i^{(k)}/C^{(k)}$, and for $p = 0, \ldots, P-1$,

$$\begin{aligned} a_i^{(k,p+1)} &= \min\Big\{1, \frac{a_i^{(k,p)}}{\sum_{t'} p_{i,t'}^{(k+1)}} \\ &\times \sum_t \frac{p_{i,t}^{(k+1)} y_{i,t}}{n_i a_i^{(k,p)} p_{i,t}^{(k+1)} + n_i \alpha^{(k+1)} s_{i,t} + r_i}\Big\} \\ a_i^{(k+1)} &= a^{(k,P)} \end{aligned} \tag{6}$$

The estimated prompts in (5) are given by equation (4), with a_i replaced by its current estimate $a_i^{(k)}$. The global scaling factor $C^{(k)}$ restores at each iteration the correct scaling of the activity and attenuation factors, which is undetermined since only products $a_i \lambda_j$ appear in the data model (1). We calculate $C^{(k)}$ to match the mean attenuation factor in a subset of LORs to the exact value, which is known for this simulation study. An alternative is to match the mean activity in some region of interest. Note that this rescaling must be done at each iteration lest the physical constraint $a_i \leq 1$ in (6) be invalid.

III. Consistency

Do the TOF PET data uniquely determine the scatter scale factor? When the attenuation is known, the problem is linear and the answer is intuitive: the solution is unique if the scatter is inconsistent, namely if it is impossible to find a "scatter image" $\rho \in \mathbb{R}^M$ such that

$$s_{i,t} = a_i \sum_{j=1}^{M} c_{i,t,j} \rho_j \quad i = 1, \ldots, N; t = 1, \ldots, T \tag{7}$$

Note that the ρ image need not be non-negative. The increase of the image covariance due to estimating α can be calculated for a weighted least-squares (WLS) reconstruction. If V is the $NT \times NT$ data covariance matrix, and A the $NT \times M$ system matrix with elements $A_{(i,t),j} = a_i c_{i,j,t}$, one can show that the covariance matrix of the reconstructed activity is

$$\text{cov}(\lambda_{WLS}) = F^{-1} + \frac{1}{||s_{inc}||_V^2} \rho_{WLS} \cdot \rho_{WLS}^t \tag{8}$$

The first term is the covariance when the scatter scale factor is known, it is equal to the inverse of the $M \times M$ Fisher matrix $F = A^t V^{-1} A$, which we assume non-singular. The second term is the additional covariance due to the estimation of α, where $\rho_{WLS} = F^{-1} A^t V^{-1} s$ is the WLS reconstruction of the scatter, $s_{inc} = s - A\rho_{WLS}$ is the inconsistent part of the scatter background, and $||g||_V^2 = g^t \cdot V^{-1} \cdot g$. Equation (8) confirms the intuition that the scatter background should be as inconsistent as possible to allow reliable scatter correction.

Unfortunately this simple analysis cannot be extended to the non-linear CT-less problem. In particular, if the scatter is "CT-less consistent" in the sense that

$$s_{i,t} = \tilde{a}_i \sum_{j=1}^{M} c_{i,t,j} \rho_j \quad i = 1, \ldots, N; t = 1, \ldots, T \tag{9}$$

for some attenuation factors $0 < \tilde{a}_i \leq 1$ and some scatter image ρ, it does not necessarily mean that uniqueness does not hold (unless $a = \tilde{a}$). Uniqueness of the scatter scale factor remains an open question when the attenuation is unknown.

IV. Simulation

We have simulated 2D TOF data for one slice of the phantom shown in figure 1. Noise-free 2D true data and attenuation factors were simulated analytically for a mathematical model of this phantom, with 13 TOF bins (FWHM$_{TOF}$ = 87 mm, TOF bin size 46.9 mm), 256 uniform radial samples spaced 2.07 mm, and 168 angular samples. Activity images were reconstructed on a 256×256 grid with pixel side 2.07 mm. The phantom is almost identical to the NEMA IQ phantom in shape, contrast and locations of the cold and rod structures (Fig. 1), except for the external shape which on one side is ellipsoidal rather than flat. We assume uniform sensitivity $n_i = 1$ and no randoms $r_i = 0$. For the scatter background $s_{i,t}$ we used the single scatter ("SSS") estimate calculated from a

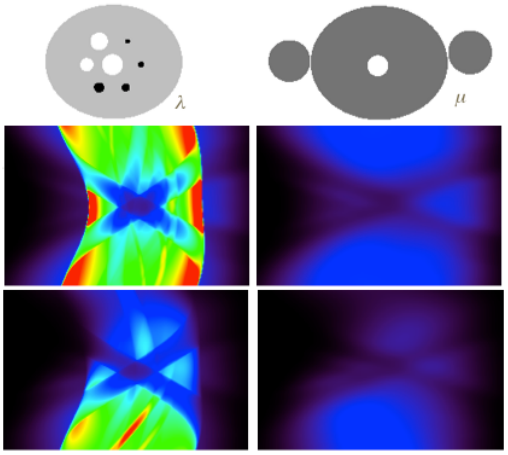

Figure 1: Top row: activity and attenuation of the studied slice of the mathematical model of the IQ phantom. The axes of the external ellipse are 288 mm and 238 mm. The ratio between the activity in the hot rods and the background is 4.0. Middle and lower rows: sinograms of time bins 0 and 2. Left: prompts, right: scatter background.

7200 s scan of the physical phantom on the mCT scanner. This scatter is added to the simulated trues to obtain a scatter fraction of 0.25, with a scatter scaling factor $\alpha_{exact} = 1.5$. MLEM and MLACF iterations are initialized with a uniform activity $\lambda_j = 1$ and attenuation $a_i = 1$, and with $\alpha = 1$.

To test the algorithms with decreasing amounts of "out of patient" LORs, we performed reconstructions with three different field-of-view (FOV): the *full FOV* (256×168 sinograms), the *attenuated FOV* (using only LORs that intersect the attenuation map, such that $a_i < 1$), and the *active FOV* (using only LORs with $\sum_t p_{i,t} > 0$). In all cases, no information on the support of the activity is assumed during MLEM or MLACF reconstruction, so that, with these hypotheses, tail fitting could not be applied. When the attenuation is known, this assumption is artificial because one can safely assume that $\lambda = 0$ where $\mu = 0$. For CT-less PET in contrast, there is normally no prior information on the support of the activity. Note that with the *active FOV*, there are no out of patient LORs at all, this example is also artificial and only meant to illustrate that the active LORs do contain information on the scatter scale.

A. Consistency

A scatter image ρ was reconstructed by applying TOF filtered-backprojection (rectangular window at Nyquist frequency) to the scatter background corrected with the known attenuation factors. That image was then reprojected and attenuated, and Figure 2 shows that the scatter background has a significant inconsistent component. The inconsistent fraction was $||s_{inc}||/||s|| = 0.47$, and $||s_{inc}||_1/||s||_1 = 0.40$. Note that the TOF summed scatter $\bar{s}_i = \sum_t s_{i,t}$ is also inconsistent, we found in that case $||\bar{s}_{inc}||/||\bar{s}|| = 0.56$ and $||\bar{s}_{inc}||_1/||\bar{s}||_1 = 0.46$ (these two last figures are not reliable however because the non TOF scatter image had a larger support than the image matrix used and the reprojection was therefore not complete).

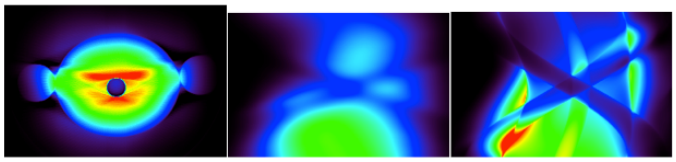

Figure 2: Known attenuation. Left: TOF FBP reconstruction ρ of the scatter background $s_{i,t}$. TOF sinogram for time bin 2: $s_{i,t}$ (middle) and reprojected and attenuated scatter image (right).

B. Noise-free data

We use ordered subset versions of the algorithms, with 14 subsets and up to 50 iterations. Figure 3 shows the good convergence to the exact value 1.5 of the estimated scatter scale $\alpha^{(k)}$ for the three different FOVs when the attenuation factors a_i are known (MLEM). The slowest convergence is observed with the *active FOV*. Figure 4 shows the corresponding convergence for CT-less PET (MLACF), here again, no tail information is available and good, though slower convergence, is observed in the three cases. Profiles through the activity and attenuation sinogram are shown in figures 5 and 6.

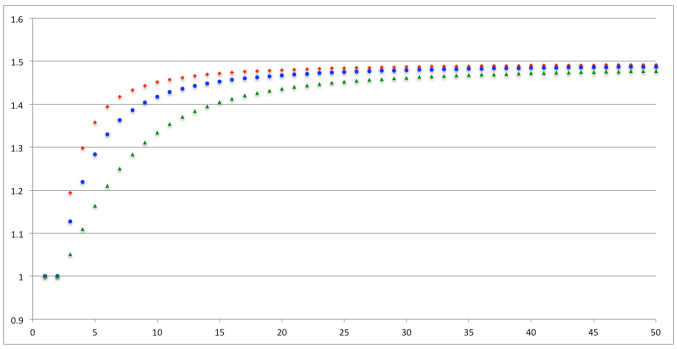

Figure 3: Known attenuation and noise-free data. Evolution of the estimated scatter scale $\alpha^{(k)}$ versus number of iterations (MLEM, 14 subsets). Red symbols: full FOV. Blue symbols: attenuated FOV. Green symbol: active FOV.

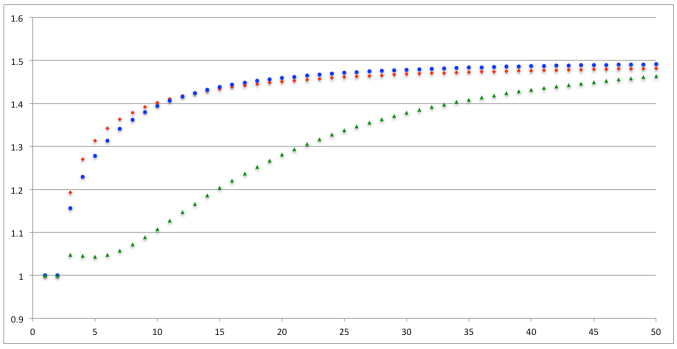

Figure 4: Unknown attenuation and noise-free data. Evolution of the estimated scatter scale $\alpha^{(k)}$ versus number of iterations (MLACF, 14 subsets). Red symbols: full FOV. Blue symbols: attenuated FOV. Green symbol: active FOV.

C. Noisy data

Noisy data were generated for three count levels of 10^7, 10^6 and 10^5 prompt events for the slice under study, this corre-

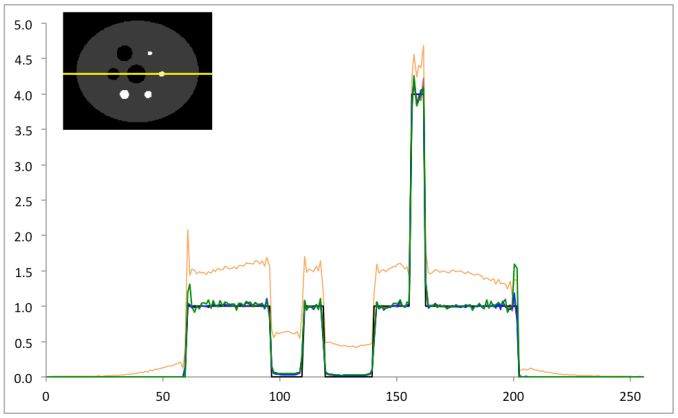

Figure 5: Unknown attenuation and noise-free data. Profile of the reconstructed activity (MLACF, 14 subsets, 50 iterations). Red: full FOV. Blue: attenuated FOV. Green: active FOV. Orange: no scatter correction. Black: reference. The profiles with the three FOVs can hardly be distinguished on this plot, except for the overshoot at the edges of the phantom with the active FOV.

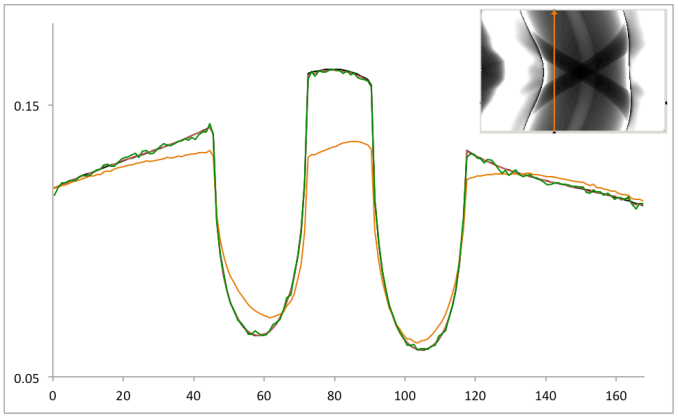

Figure 6: Unknown attenuation and noise-free data. Profile of the reconstructed attenuation sinogram (MLACF, 14 subsets, 50 iterations) at fixed radial variable, the horizontal scale in the plot is the azimuthal angular index. Red: full FOV. Blue: attenuated FOV. Green: active FOV. Orange: no scatter correction. The profile shown is within the active FOV; as shown by the insert MLACF is unable to reconstruct the attenuation outside the active FOV.

sponds to maximum mean counts $\max_{i,t}(<y_{i,t}>) = 211, 21$ and 2.1. For each noise level, 100 noise realizations were generated and reconstructed.

The most important observation is an increasing negative bias of the estimated scatter scaling factor with increasing noise levels, as shown in Table I for the case with known attenuation, whereas the standard deviation remains much smaller than the bias. Table I also shows, as expected, that the bias disappears if a linear estimator is used instead of the Poisson ML estimation (for both λ and α). A numerically efficient alternative to decrease the bias on α is to perform a first pass reconstruction using MLEM with a coarser image sampling (Table I). This was done keeping the original data sampling, without rebinning the sinograms on a coarser grid.

The bias is even more important with MLACF (Table II), resulting in an increasing under-correction of the scatter background as the noise level decreases (Figure 7,8). The CT-less problem is not linear, and least-squares estimation does

MLEM, 2.07 mm pixels

Counts	Full FOV	Attenuated FOV	Active FOV
10^7	1.44 (0.003)	1.42 (0.003)	1.40 (0.004)
10^6	1.31 (0.008)	1.26 (0.011)	1.19 (0.016)
10^5	1.14 (0.028)	1.00 (0.029)	0.69 (0.041)

Unweighted least-squares

Counts	Full FOV	Attenuated FOV	Active FOV
10^7	1.50 (0.005)	1.52 (0.006)	1.52 (0.007)
10^6	1.49 (0.017)	1.53 (0.019)	1.52 (0.022)
10^5	1.48 (0.050)	1.52 (0.053)	1.54 (0.066)

MLEM, 8.28 mm pixels

Counts	Full FOV	Attenuated FOV
10^7	1.41 (0.003)	1.40 (0.004)
10^6	1.36 (0.010)	1.35 (0.013)
10^5	1.40 (0.032)	1.39 (0.037)

Table I: Known attenuation and noisy data. Mean estimated α and (standard deviation), for ML estimation, unweighted least-squares estimation (preconditioned conjugate gradient, 50 iterations), and ML estimation on a coarse 64×64 image matrix.

MLACF, 2.07 mm pixels

Counts	Full FOV	Attenuated FOV	Active FOV
10^7	1.28 (0.004)	1.40 (0.004)	1.43 (0.011)
10^6	0.69 (0.024)	1.13 (0.011)	1.22 (0.031)
10^5	0.01 (0.01)	0.25 (0.016)	0.17 (0.029)

MLACF, 2.07 mm pixels, pre-smoothed data

Counts	Full FOV
10^7	1.47 (0.003)
10^6	1.44 (0.008)
10^5	1.33 (0.030)

Table II: Unknown attenuation and noisy data. Mean estimated α and (standard deviation), for MLACF estimation, and MLACF estimation applied to pre-smoothed data.

not guarantee unbiased estimation. Attempts to reconstruct on a coarse image grid, as in Table I, did not decrease the bias significantly. Table II shows that the bias can be reduced by strongly smoothing the data $y_{i,t}$ and the scatter $s_{i,t}$ (here with a 2D gaussian filter in the radial and azimuthal sinogram variables, with FWHM equal to 8 samples).

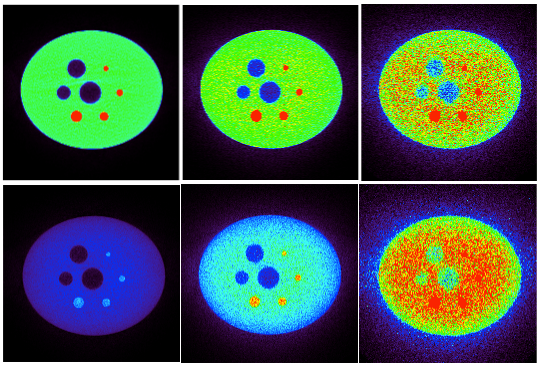

Figure 7: Unknown attenuation, full FOV. Mean (top row, color scale (0, 2)) and standard deviation (bottom row, color scale (0,4)) of the activity reconstructed from 100 noise realizations. MLACF (14 subsets, 50 iterations). From left to right: 10^7, 10^6, 10^5 prompts.

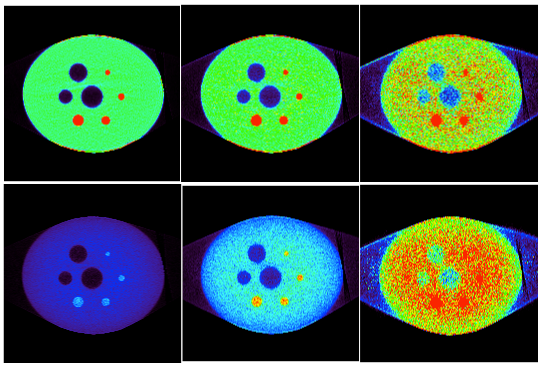

Figure 8: Unknown attenuation, attenuated FOV. Mean (top row, color scale (0, 2)) and standard deviation (bottom row, color scale (0,4)) of the activity reconstructed from 100 noise realizations. MLACF (14 subsets, 50 iterations). From left to right: 10^7, 10^6, 10^5 prompts.

V. Conclusions

The goal of this work is to investigate the simultaneous maximum likelihood estimation of the activity and scatter scaling factor from TOF PET data, for configurations where the FOV contains few ("attenuated FOV") or none ("Active FOV") lines of response without activity. We assumed in this work the absence of any information on the support of the activity image, and therefore the tail fitting method does not apply. When such information is available, tail fitting is applicable and yields accurate estimates of α. In that case, even better results can be obtained by simultaneous estimation of λ and α using all LORs, as noted in [16] in the case of known attenuation. This only requires enforcing the activity to be zero in some sufficiently large region of the FOV, as we also verified with our simulation (results not shown).

The case with known attenuation was previously investigated in [17], that work is here extended to CT-less PET. This preliminary study on simulated 2D TOF data shows that TOF data contain enough information to determine the scatter scale factor, even when the data only contain LORs that intersect the activity image. A second observation is that the ML estimator of the scatter scale is strongly biased for low count data. The bias is negative in our study, suggesting that the scatter scale can be considered as a "hot" pixel. A linear estimator such as unweighted least-squares can be used to eliminate the bias when the attenuation is known, but this approach does not apply to the bi-linear problem with unknown attenuation. Alternatively the bias can be reduced by means of a first pass reconstruction on a coarse image matrix or by using strongly smoothed data.

We used a scatter estimate obtained from a model-based single scatter simulation. This leaves open the issue of how to obtain an initial estimate of the activity, and (for CT-less PET) of the attenuation, as these quantities are needed to calculate the scatter estimate. This issue is common to other methods and it is typically hoped that an iterative application may provide a solution. The method proposed here however could also be applied if the scatter estimate is obtained from one or several secondary energy windows.

In this study the trues and the attenuation have been simulated analytically. Instead of generating a scatter model by convolving the trues with some low pass filter, we have used the model-based single scatter estimate calculated from a physical measurement of the NEMA IQ phantom. Since the mathematical and physical phantoms do not perfectly coincide, the results will need to be confirmed by reconstructing Monte-Carlo and real data. Another limitation of this study is the assumption that the scatter estimate $s_{i,t}$ is exact, up to the scale factor, the influence of multiple scatter has been neglected.

References

[1] C Tsoumpas, 2004, Implementation and evaluation of scatter estimation algorithms in PET, Master thesis, U. of Patras.

[2] H Zaidi, K F Koral, 2004, Scatter modeling and compensation in emission tomography, *Eur J Nucl Med Mol Imaging*, **31**, 761-782.

[3] S Grootoonk, T J Spinks, D Sashin, N M Spyrou, T Jones, 1996, Correction for scatter in 3D brain PET using a dual energy window method, *Phys Med Biol*, **41**, 2757-2774.

[4] L E Adam, J S Karp, R Freifelder, 2000, Energy-Based Scatter Correction for 3-D PET Scanners Using NaI(Tl) Detectors, *IEEE Trans Med Imag*, **19 (5)**, 513-521.

[5] B Bendriem, R Trebossen, V Frouin, A Syrota, 1993, A PET scatter correction using simultaneous acquisitions with low and high lower energy thresholds, *Proc. IEEE Med Imag Conf (Seattle)*, 1779-83.

[6] J M Ollinger, 1996, Model-based scatter correction for fully 3D PET., *Phys Med Biol* **41** 153-176.

[7] C C Watson, 2000, New, Faster, Image-Based Scatter Correction for 30 PET, *IEEE Trans Nucl Sc* **47**(4) 1587-1594.

[8] C C Watson, 2007, Extension of Single Scatter Simulation to Scatter Correction of Time-of-Flight PET, *IEEE Trans Nucl Sc* **54**(5) 1679-1686.

[9] A Werling, O Bublitz, J Doll, L E Adam, G Brix, 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.

[10] S D Wollenweber, 2002, Parameterization of a model-based 3-D PET scatter correction, *IEEE Trans Nucl Sc* **49**(3) 722-727.

[11] N C Ferreira, R Trebossen, C Lartizien, V Brulon, P Merceron, B Bendriem, 2002, A hybrid scatter correction for 3D PET based on an estimation of the distribution of unscattered coincidences: implementation on the ECAT EXACT HR+, *Phys Med Biol* **47** 1555-1573.

[12] J-C Cheng, S Blinder, A Rahmim, V Sossi, 2010, A Scatter Calibration Technique for Dynamic Brain Imaging in High Resolution PET, *IEEE Trans Nucl Sc* **57**(1) 225-233.

[13] B Guerin, 2010, PhD thesis, Université Pierre et Marie Curie.

[14] B Guerin, G El Fakhri, 2011, Novel scatter compensation of list-mode PET data using spatial and energy dependent corrections, *IEEE Trans Med Imag*, **30**, 759-773.

[15] Y Berker, F Kiessling, V Schulz, 2014, Scattered PET data for attenuation-map reconstruction in PET/MRI, *Med Phys*, **41**, 102502, 13p.

[16] K Thielemans, R M Manjeshwar, C Tsoumpas, F P Jansen, 2007, A New Algorithm for Scaling of PET Scatter Estimates Using all Coincidence Events, *Proc. IEEE Med Imag Conf (Dresden)*, 3586-90.

[17] V Panin, 2012, Scatter Estimation Scaling with All Count. Use by Employing Discrete Data Consistency Conditions, *Proc. IEEE Med Imag Conf (Anaheim)*, 2998-3004.

[18] A Rezaei, M Defrise, G Bal, C Michel, M Conti, C Watson, J Nuyts, 2012, Simultaneous reconstruction of activity and attenuation in time-of-flight PET, *IEEE Trans Med Imag*, **31 (12)**, 2224-2233.

[19] J Nuyts, P Dupont, S Stroobants, R Benninck, L Mortelmans, P Suetens, 1999, Simultaneous maximum a-posteriori reconstruction of attenuation and activity distributions from emission sinograms, *IEEE Trans Med Imag*, **18 (5)**, 393-403.

[20] A Rezaei, M Defrise, J Nuyts, 2014, ML-reconstruction for TOF-PET with simultaneous estimation of the attenuation factors, *IEEE Trans Med Imag*, **33 (7)**, 1563-1567.