On Estimating the Variance of Smoothed MLEM Images.

Johan Nuyts

Abstract— The maximum-likelihood expectation- maximization (MLEM) algorithm is being used increasingly as a routine reconstruction algorithm in positron emission tomography (PET). After a finite number of iterations, the response of the MLEM-reconstruction algorithm is non-linear and not shift-invariant. As a result, two problems occur when MLEM is used for tracer kinetic modeling. First, resolution and recovery may be different for different positions and for different time frames, adversely affecting the modeling results. Second, it is not trivial to determine appropriate weights for weighted least squares fitting to image derived time-activity curves. The first problem can be remedied by applying a "relatively high" number of iterations, followed by smoothing with a shift invariant kernel. The smoothing kernel reduces the noise and imposes an (approximately) constant resolution. For the second problem, different approaches exist. Some authors have shown that the variance of pixel values in MLEM images is proportional to the mean of the pixel value. Others have suggested that the uncertainty about a reconstructed pixel value can be estimated from the Fisher information matrix. These two approaches produce different results. Our simulation study confirms that the pixel variances of unsmoothed MLEM images are approximately proportional to the pixel values. However, smoothing reduces the variance more in high count regions than in low count regions. As a result, the variances of smoothed MLEM pixel values correlate better with the reciprocal of the diagonal elements of the Fisher information matrix. Smoothing strongly affects the variance because neighboring pixel values are highly correlated. Because computing the mean of a region is a smoothing operation, the Fisher information can be used to compute appropriate weight values for model based analysis of time-activity curves.

Keywords— Maximum likelihood, Reconstruction, Variance.

I. Introduction

ALTHOUGH MLEM reconstruction is being used increasingly for many applications, filtered backprojection (FBP) is still the method of choice if the reconstructed images are used for tracer kinetic modeling. The main problem with MLEM is that it produces a position dependent resolution after a finite number of iterations. One solution is to iterate long while regularizing with a penalty designed to yield a uniform spatial resolution [1]. Another, simpler approach is to impose a constant point spread function by applying a high number of iterations, followed by smoothing. Since the maximum-likelihood solution is (virtually) unbiased [2], the resolution of the reconstruction should converge towards the sinogram resolution (or better if the finite resolution is taken into account) after a suffi-

cient amount of iterations. Consequently, subsequent convolution with a Gaussian kernel should impose a (nearly) position independent resolution, in particular if the width of the Gaussian is large compared to the sinogram resolution. In that case, the point spread function of the smoothed image is approximately equal to that kernel.

A second problem is the estimation of appropriate weights for weighted least squares fitting to time-activity curves (using individual pixel values or the mean value over a region-of-interest). This is more difficult for MLEM than for FBP, because MLEM is non-linear. Barrett and coauthors [3] have derived a recursive method to compute the variance after each MLEM iteration. Their simulations revealed that the variance image strongly resembles the mean image (or the noise-free MLEM reconstruction) [4]. A different approach is used by Fessler et al [5], [6] and Qi et al [7]. They study the characteristics of the ML (and penalized ML) solution, assuming that applying a sufficient amount of iterations yields an image with similar characteristics. They obtain good results assuming that the variance of a pixel value is approximately inversely proportional to the corresponding diagonal element of the Fisher information matrix (FIM). The results obtained with the two approaches are seemingly contradictory, because the image formed by the reciprocal of the diagonal of the FIM differs markedly from the mean image.

II. Theory

In emission tomography, it is reasonable to assume that the photon count y_i measured in detector i, is a Poisson realization of $\bar{y}_i$, with

$$\bar{y}_i = \sum_j c_{ij} \bar{\lambda}_j, \tag{1}$$

where $\bar{\lambda}_j$ is the (unknown) radioactivity in pixel j and c_{ij} is the probability that a photon emitted in j is detected in i. The MLEM algorithm produces estimates λ_j of $\bar{\lambda}_j$ by maximizing the logarithm of the likelihood [8]:

$$
\begin{aligned}
L_y(\lambda) &= \ln(p(y|\lambda)) \tag{2} \\
&= \sum_i \left(y_i \ln(\sum_j c_{ij}\lambda_j) - \sum_j c_{ij}\lambda_j - \ln(y_i!) \right)
\end{aligned}
$$

The elements f_{jk} of the associated FIM can be computed as [2]:

$$f_{jk} = E\left(-\frac{\partial^2 L}{\partial \lambda_j \partial \lambda_k}\right) = \sum_i \frac{c_{ij}c_{ik}}{\bar{y}_i}, \tag{3}$$

where E denotes expectation with respect to $p(y|\lambda)$. For an unbiased estimator, the following inequality holds [2]:

$$\text{var}(\lambda_j) \geq (FIM^{-1})_{jj} \geq \frac{1}{f_{jj}}, \qquad (4)$$

where λ_j is the estimated value for pixel j. If the estimator is efficient, the covariance matrix is the inverse of the FIM. For realistic applications, inversion of the FIM is prohibitive. In addition, MLEM is not truly efficient (the ML estimate is asymptotically efficient [2]), so the inverse of FIM may be a poor estimate of the true covariance matrix of the MLEM-reconstruction. Note that in practice, the FIM cannot be computed since $\bar{y}_i$ is unknown. Fortunately, f_{jk} is a smooth function (a modified backprojection) of $1/\bar{y}_i$, so a useful estimate of the FIM can be obtained by simply using the measured values y_i instead of their unknown means $\bar{y}_i$. Obviously, there is a problem when $\bar{y}_i$ equals zero. The FIM goes to infinity for vanishing $\bar{y}_i$ because the variance vanishes as well. To avoid zero devisions, we replace zeros with a small positive value (we used 0.5) in (3).

In [5], Fessler derived an approximate expression for the variance of penalized likelihood reconstructions. This expression involves inversion of the FIM, which is feasible only if the variance has to be computed in a few pixels. In [6], only the diagonal elements f_{jj} of the FIM are used to estimate the "certainty" about the pixel values λ_j provided by the likelihood. With this approach, Fessler et al are able to tune the local strength of a quadratic prior to that of the likelihood in order to obtain uniform resolution. The success of this approach suggests that $1/f_{jj}$ should be a reasonable measure of the uncertainty (or variance) about pixel value λ_j.

Wilson et al [4] have analyzed the noise properties of the MLEM algorithm with Monte Carlo simulations. They reported a linear relationship between the variance in a particular pixel and the total projection count. They also found a smooth monotonic relation between the pixel values and the pixel variances within the same image, although there was a deviation from linearity, which depended on the iteration number.

The results obtained by Fessler et al. and by Wilson et al. are seemingly contradictory, because there is no nearly linear relation between $1/f_{jj}$ and λ_j.

The k-th row or column of the FIM can be regarded as an image f_{jk} with fixed k, representing the amount of information about the value of pixel k carried by each of the pixel values in the reconstructed image (assuming we are using a (nearly) efficient estimator). As seen from (3) this image is obtained by backprojection of the very sparse sinogram c_{ik}/y_i. Consequently, the information is positive everywhere, tends to be highest in pixel k and decreases roughly proportional with the reciprocal of the distance between pixels j and k. The information is concentrated in a blob centered at k. Assuming that most rows of the FIM have such a shape, the columns of the corresponding covariance matrix would need to have some similarity to a deconvolution filter, with high variance v_{kk} and with

covariances v_{jk} which oscillate and decrease rapidly in amplitude with increasing distance between j and k. Wilson et al [4] show that profiles through the covariance matrices obtained at high iterations of the MLEM algorithm indeed show such behavior.

Smoothing of the reconstructed image has a strong effect on the variance:

$$\begin{aligned} \lambda_s' &= \sum_j w_{sj}\lambda_j \\ \text{var}(\lambda_s') &= \sum_j w_{sj}^2 v_{jj} + \sum_j \sum_{j \neq k} w_{sj} w_{sk} v_{jk}, \end{aligned} \qquad (5)$$

where λ_s' is the pixel value after smoothing, w_{sj} are the coefficients of the smoothing kernel, and v_{jj} and v_{jk} are the variances and covariances. Because the covariances $v_{jk,j \neq k}$ tend to be negative for neighboring pixels, the variance decreases rapidly with moderate smoothing. It follows that the variance of a pixel value overestimates the total uncertainty about that value. The findings by Fessler and Wilson can be reconciled by assuming that the covariances are stronger in regions with high activity, such that the variance resembles the mean. Smoothing the reconstruction should reduce the variance in a position dependent way, making the variance image more similar to the "uncertainty image" $1/f_{jj}$.

Consequently, we hypothesize that a useful measure for the variance of pixel value j after smoothing can be computed from the Fisher information matrix, as follows:

$$\text{var}(\sum_j w_j \lambda_j) \sim \sum_j w_j^2 \frac{1}{f_{jj}}, \qquad (6)$$

where w_j are the coefficients of a smoothing kernel. Computing the mean value of a region is a special case of smoothing, so (6) produces a measure of the variance on the region value, when $w_j = 1/N$ if j belongs to the region with N elements, and $w_j = 0$ otherwise.

III. EXPERIMENTS

Three Monte Carlo simulation experiments have been carried out:

1. Noisy sinograms were simulated and images were reconstructed by applying 1000 MLEM iterations, and smoothed with different smoothing masks. The relation between the mean, the variance and the Fisher information in every pixel is studied.
2. Typical radial profiles of the covariance images and Fisher information images are computed, to verify if the shape of the profiles depends on mean activity.
3. A cardiac PET ammonia study, consisting of 22 frames, has been simulated and 800 noisy sinograms were computed. All images are reconstructed by applying 44 MLEM iterations. For five regions, the time-activity curves and corresponding variances are computed and compared to the uncertainty predicted from the Fisher information.

2

A. Pixel variances

Several simulations using the Shepp-Logan phantom were carried out. A noise-free 2-D sinogram was computed, and 200 Poisson realizations were derived. From each noisy sinogram, an image was reconstructed by applying 1000 (unaccelerated) MLEM iterations. The resulting images were smoothed with a 2-D Gaussian convolution kernel. Then, for each pixel the mean and the variance was computed from the 200 smoothed reconstructions. The diagonal elements f_{jj} of the FIM were calculated, and a measure of the variance image was computed using (6). Different convolution kernels were evaluated, the full width at half maximum (FWHM) ranged from 0 to 8 pixels.

The simulations were repeated for attenuated sinograms, assuming attenuation as in positron emission tomography, and noise-free attenuation factors. The MLEM algorithm corrects for attenuation by including it in the detection probabilities c_{ij}. Similarly, computation of the values f_{ij} of the FIM is straightforward: attenuation is taken into account because it affects the detection probabilities c_{ij} and the mean $\bar{y}_i$ of the Poisson distribution. To estimate the variance from the mean MLEM pixel value, the mean value was divided by the mean attenuation for that pixel.

B. Covariance profiles

From the same 200 simulations, the covariance image v_{jk} for a fixed pixel k can be estimated. A radial profile is extracted by averaging the covariance values over circles centered at pixel k. This procedure has been applied to about 50 pixels from the ellipse with lowest activity (see figure 2), and the 50 profiles were averaged. The same was done for 30 pixels of the ellipse with highest activity, to obtain an average radial covariance profile for high count pixels.

Similarly, the values f_{jk} are computed with (3) for the same pixels, to produce two average radial profiles through the Fisher information images, one for pixels with high activity, and another one for pixels with low activity.

C. Dynamic PET study

A simplified thorax image was simulated. The thorax consists of five different compartments: the elliptic body, the two lungs, the left ventricular wall, the left ventricular cavity and the liver. Time-activity curves for each compartment have been derived from region-of-interest analysis of a clinical ammonia study. Figure 1 shows the 22 consecutive frames. The image has 150 by 150 pixels, there is only 1 slice. Sinograms were computed, taking into account the attenuation map also shown in figure 1. Eight hundred Poisson noise realizations were computed and reconstructed with MLEM using subset acceleration [9]. We used an iteration scheme with decreasing number of subsets, consisting of (iterations x subiterations): 1 x 15, 1 x 10, 1 x 6, 1 x 5, 1 x 3, 5 x 1, which is equivalent to 44 regular MLEM iterations. For each frame and each noise realization, the average activity in each of the 5 regions was computed. From the 800 noise realizations, the mean and
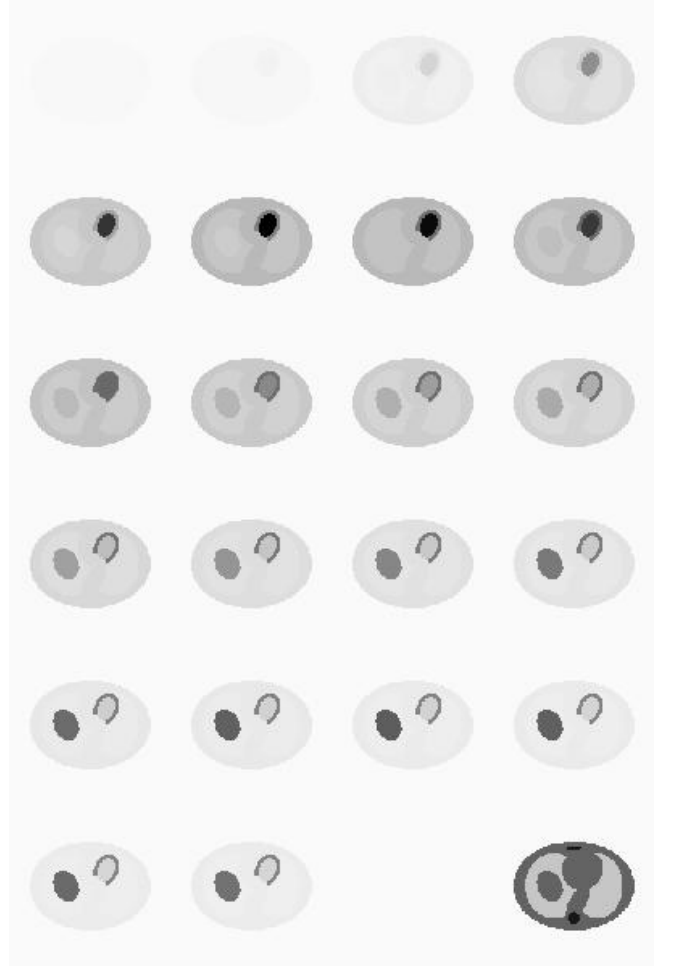

Fig. 1. The 22 frames of the simulated dynamic PET study. The last image is the attenuation map.

variance of each time point in the five time-activity curves were estimated. Note that the relative error on the variance computed from 800 realizations equals $\sqrt{2/800} = 0.05$.

For each time frame, the diagonal elements of the FIM were computed using (3), incorporating the attenuation in the detection probabilities c_{ij}. The measure for the variance on the region values is computed using (6).

The procedure was repeated, this time using a Gaussian smoothing kernel with full width at half maximum = 2.5 pixels, which was applied after the reconstruction.

IV. RESULTS

A. Pixel variances

We obtained similar results for simulations with and without attenuation. The figures are for the simulations without attenuation. Figure 2 shows the images of the mean pixel values obtained after 1000 MLEM iterations and smoothing with three different Gaussians (FWHM of 0, 3 and 8 pixels). The same figure displays the corresponding variance images computed from the 200 noise realizations, and the images of the variance measure obtained with (6). Figure 3 plots the mean pixel value and the FIM-based
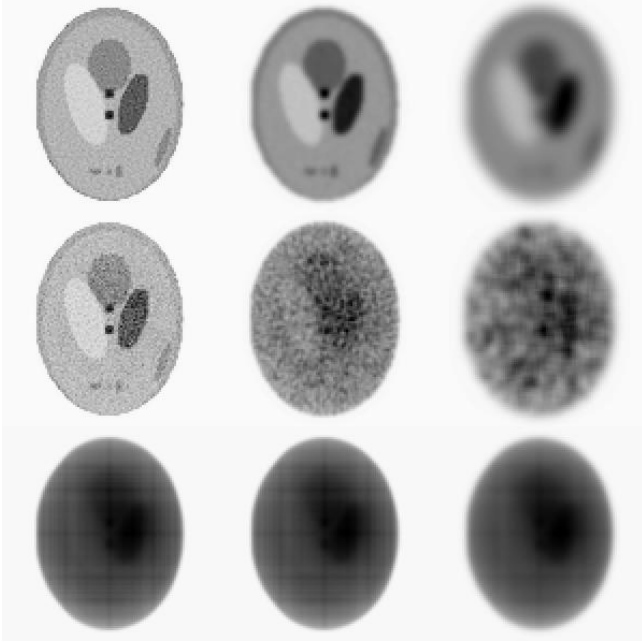
Fig. 2. Mean, variance and reciprocal of the diagonal of the FIM. Top row: the mean images for smoothing kernels of 0 (= no smoothing), 3 and 8 pixels. Middle row: the corresponding variance images. Bottom row: the variance predicted from the FIM. The images are scaled to their own maximum.
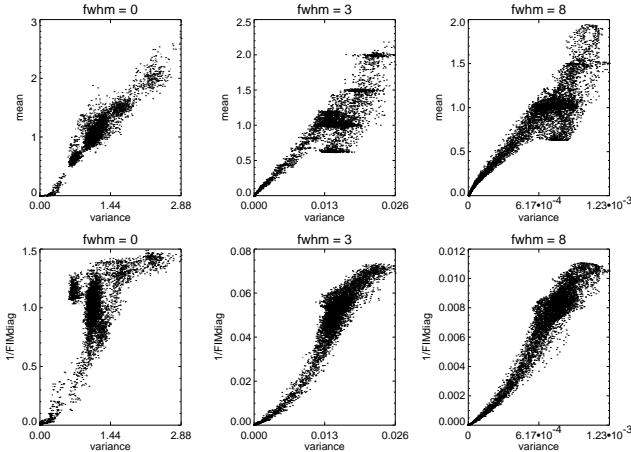

Fig. 3. The mean (top row) and the diagonal of the FIM (bottom row) as a function of the variances for the three filters (FWHM = 0, 3, 8 pixels).

variance measure as a function of the variance for the same smoothing kernels.

Both figures confirm that there is a (nearly) linear relationship between mean and variance if the MLEM images are not smoothed. Smoothing destroys this relationship, and the variance image becomes more similar to the reciprocal of the diagonal of the FIM.

B. Covariance profiles

Figure 4 shows the average radial profiles through the covariance and Fisher information images, for high and low count pixels. The low activity pixels have a higher Fisher
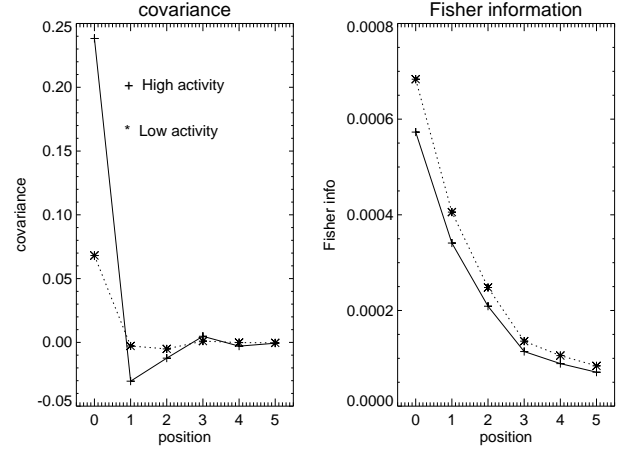

Fig. 4. Average radial profiles through the covariance (left) and Fisher information images (right). The position is expressed in pixels, position 0 is the central pixel.

information, because the variance on low count projection data is lower. These two profiles are similar in shape, indicating that the central value is a reasonable measure for the total amount of information about a pixel value. The covariance profiles are different: there are significant negative covariances for the high activity region, for the low activity region they are closer to zero.

C. Dynamic PET study

Figure 5 shows the variances (computed from the 800 realizations) as a function of the frame number. The mean region values and the uncertainty estimated with (6) are plotted in overlay, all curves are normalized to the same mean value. The FIM-based measure closely follows the variance, while for some regions (the liver in particular) the mean values show a different behavior. Figures 6 and 7 confirm that the variance correlates better with the FIM-based measure than with the mean.

The figures show the results for the non-smoothed MLEM reconstruction. The results obtained after smoothing are virtually identical, probably because the smoothing effect of averaging over regions is much stronger than that of a convolution with a Gaussian kernel of 2.5 pixels wide.

V. Discussion

Our simulations confirm that the variance in unsmoothed MLEM images is very similar to the image of the mean values, in agreement with the findings by Wilson et al [4]. Figure 4 confirms that the variance is higher in regions of high activity. The same figure illustrates that this higher variance is partly explained by that fact that the data provide less information about high activity regions (the Fisher information profile has a lower amplitude for more active regions). However, it also shows that the higher deviations from the mean are partly compensated by deviations in the other direction in neighboring pixels. In other words: the neighboring pixels seem to carry a larger fraction of the information in high activity regions, which causes an
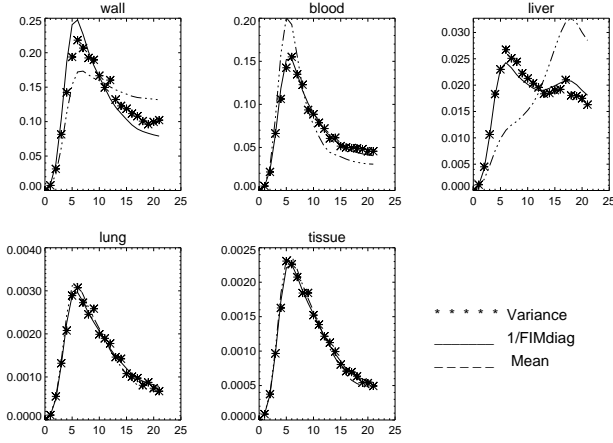
Fig. 5. The estimated variances (*) plotted as function of the frame number, for each of the five regions. Also the variance estimated from the FIM (solid line) and the mean region value (dashed line) are plotted. All curves are normalized to the same mean.
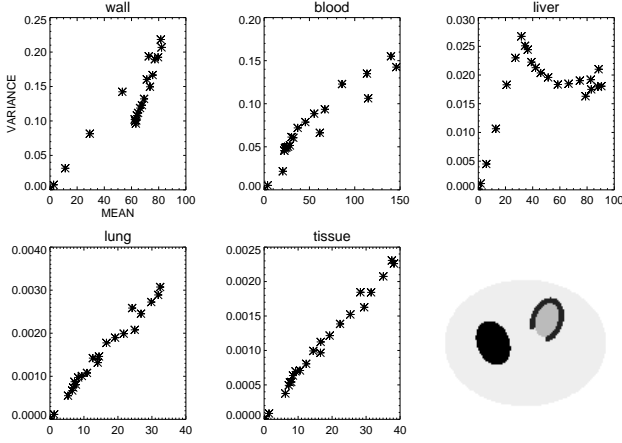

Fig. 6. The variance plotted as a function of the mean for each of the five regions. Each point corresponds to a single time frame. The image shows the activity distribution in the last frame.
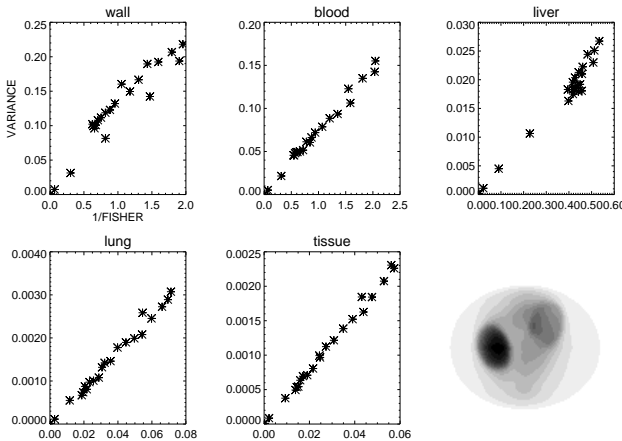

Fig. 7. The variance plotted as a function of the FIM-based value (equation (6)) for each of the five regions. Each point corresponds to a single time frame. The image shows the reciprocal of FIM diagonal $1/f_{jj}$ for the last frame.

additional increase of variance. In every MLEM-iteration, the deviations between predicted and measured projection values are backprojected and then multiplied with the current reconstruction value. Therefore, it is not very surprising that pixels with high reconstruction values tend to accumulate more noise in the first iterations. As shown in figure 2, the effect is still prominent after 1000 iterations. It is unclear if this behavior is a true feature of the ML-solution, or if even more than 1000 iterations are needed to erase the transient features of the maximization algorithm we are using.

Smoothing combines the information from neighboring voxels, causing a more rapid decrease of variance in regions where the covariances are significant. As a result, the variance on the smoothed reconstruction values is no longer proportional to the mean and is better predicted by the reciprocal of the FIM diagonal. Computing the average of a region is a smoothing operator, and the simulations confirm that the variance on the mean region value is proportional to the sum of the uncertainties about the contributing pixel values.

In tracer kinetic modeling analysis, the parameters of a compartmental model are often estimated by weighted least squares fitting to the measured time-activity curves. The absolute values of the weights are unimportant, the weights must be determined only except for a constant factor. Our findings suggest a simple way to compute the weights for the frame points. Computation of the FIM only involves a single modified backprojection per frame point, so the computational burden is very low compared to that of the MLEM reconstruction.

REFERENCES

[1] JW Stayman, FA Fessler. "Regularization for uniform spatial resolution properties in penalized-likelihood image reconstruction." *IEEE Trans Med Imaging* vol 19, pp 601-615, 2000.
[2] HH Barrett, JL Denny, RF Wagner, KJ Myers. "Objective assessment of image quality. II. Fisher information, Fourier crosstalk, and figures of merit for task performance." *J Opt Soc Am A* vol 12, pp 834-852, 1995.
[3] HH Barrett, DW Wilson, BMW Tsui. "Noise properties of the EM algorithm: I. Theory." *Phys Med Biol* vol 39, pp 833-846, 1994.
[4] DW Wilson, BMW Tsui, HH Barrett. "Noise properties of the EM algorithm: II. Monte Carlo simulations." *Phys Med Biol* vol 39, pp 847-871, 1994.
[5] JA Fessler. "Mean and variance of implicitly defined biased estimators (such as penalized maximum likelihood): applications to tomography." *IEEE Trans Image Proc* vol 5, pp 493-506, 1996.
[6] JA Fessler, WL Rogers. "Spatial resolution properties of penalized-likelihood image reconstruction: space-invariant tomographs." *IEEE Trans Image Proc.* vol 5, pp 1346-1358, 1996.
[7] J Qi, RM Leahy. "A theoretical study of the contrast recovery and variance of MAP reconstructions from PET data." *IEEE Trans Med Imaging* vol 18, 293-305, 1999.
[8] LS Shepp, Y Vardi. "Maximum likelihood reconstruction for emission tomography," *IEEE Trans Med Imaging*, vol MI-1, pp. 113-122, 1982.
[9] MH Hudson, RS Larkin. "Accelerated image reconstruction using ordered subsets of projection data," *IEEE Trans Med Imaging* 1994; 13: 601-609.