

# Objective Functions for Tomographic Reconstruction from Randoms-Precorrected PET Scans

Mehmet Yavuz and Jeffrey A. Fessler  
Dept. of EECS, University of Michigan

## ABSTRACT

In PET, usually the data are precorrected for accidental coincidence (AC) events by *real-time* subtraction of the delayed window coincidences. Randoms subtraction compensates in mean for AC events but destroys the Poisson statistics. Furthermore, for transmission tomography the weighted least-squares (WLS) method leads to systematic biases, especially at low count rates. We propose a new “shifted” Poisson (SP) model for precorrected PET data, which properly matches the first and second order moments of the measurement statistics. Using simulations and analytic approximations, we show that estimators based on the “ordinary” Poisson (OP) model for the precorrected data lead to higher standard deviations than the proposed method. Moreover, if one zero-thresholds the data before applying the maximization algorithm, the OP model results in systematic bias. It is shown that the proposed SP model leads to penalized-likelihood estimates free of systematic bias, even for zero-thresholded data. The proposed SP model does not increase the computation requirements compared to OP model and it is robust to errors in the estimates of the AC event rates.

## I. INTRODUCTION

Accidental coincidence (AC) events are a primary source of background noise in PET measurements. AC events occur when photons that arise from separate annihilations are mistakenly registered as having arisen from the same annihilation. In transmission scans the photons that originate from different transmission sources (rod or sector sources rotating around the patient) cause AC events. Due to rod masking, ratio of these AC events to “true” events are usually small in transmission scans compared to emission scans. However the effect of AC events is significant for regions of high attenuation coefficients, because projections through the regions of high attenuation coefficients result in low true coincidence rates. These low count rates can become comparable to AC rates. In a conventional PET scan, the data are precorrected for AC events by real-time subtraction of the delayed-window coincidences. Real-time subtraction of delayed window coincidences [6] compensates in mean for AC events but destroys the Poisson statistics. To avoid this problem, one needs to maintain the transmission and randoms measurements as two separate sinograms [7]. However even if a PET system enables one to collect randoms (delayed coincidences) sino-

gram separately, this process doubles the storage space for the acquired data. So in practice most PET centers collect and archive only the randoms precorrected data.

Although our analysis and proposed model applies to both emission and transmission tomography, in this paper we focus on transmission tomography. We argue that for transmission scans, the WLS method and the ML method based on ordinary Poisson (OP) model lead to systematic bias and higher variance, respectively. Thus, we propose a “shifted” Poisson (SP) model which matches both the first and second-order moments of the model to the underlying statistics of the precorrected data. The corresponding log-likelihood function is shown to have better agreement with the exact log-likelihood function than the WLS and OP objective functions. We performed 2D simulations which showed that the proposed SP model yields lower variance in the reconstructed images than the OP model. Another observation was that the WLS method leads to unacceptably high systematic bias, especially for low count rates. Lastly, we investigated the effect of using estimates of AC rates. The SP estimator is found to be robust to the errors in the estimates of AC rates. The SP model has similar computation requirements for the maximization algorithm as that of the OP model.

## II. MEASUREMENT MODEL

In a conventional PET scan, the data are precorrected for AC events by *real-time* subtraction of the delayed-window coincidences [6]. The system detects coincidence events during two time windows: “prompt” window and “delayed” window. For each coincidence event in the prompt window, the corresponding sinogram bin is incremented. The statistics of these increments should be well approximated by a Poisson process. However, for coincidence events within the second delayed window, the corresponding sinogram bin is decremented, so the resultant “precorrected” events are *not* Poisson. Since prompt events and delayed events are independent Poisson processes, the precorrected measurements correspond to the difference of two independent Poisson random variables with variance equal to the sum of the means of the two random variables. In other words, randoms subtraction compensates in mean for AC events, but it also increases the variance of the measurement by an amount equal to the mean of AC events.

Let  $y = [y_1, \dots, y_N]'$  denote the vector of *precorrected* transmission scan measurements. The precorrected measurement for the  $n$ th coincidence detector pair is:

$$y_n = y_{n,p} - y_{n,d} \quad (1)$$

This work was supported in part by NIH grants CA-60711 and CA-54362.

where  $y_{n,p}$  and  $y_{n,d}$  are the number of coincidences within the prompt and delayed windows, respectively. Let  $\mu = [\mu_1, \dots, \mu_p]'$  denote the vector of unknown linear attenuation coefficients. We assume that  $y_{n,p}$  and  $y_{n,d}$  are statistically independent realizations of the random variables  $\{Y_{n,p}\}_{n=1}^N$  and  $\{Y_{n,d}\}_{n=1}^N$  having Poisson distributions with means  $\bar{y}_{n,p}$  and  $\bar{y}_{n,d}$  respectively as:

$$\bar{y}_{n,p}(\mu) = b_n e^{-l_n(\mu)} + r_n \quad (2)$$

$$\bar{y}_{n,d} = r_n \quad (3)$$

where  $l_n(\mu) = \sum_{j=1}^p a_{nj} \mu_j$  is the total attenuation between  $n$ th detector pair. The  $a_{nj} \geq 0$  factors have units of length and describe the tomographic system geometry. The  $b_n > 0$  factors denote the blank scan counts and the  $r_n$  factors denote the mean of AC events.

Since  $y_{n,p}$  and  $y_{n,d}$  are statistically independent:

$$E\{y_n\} = \bar{y}_n = \bar{y}_{n,p}(\mu) - \bar{y}_{n,d} = b_n e^{-l_n(\mu)},$$

$$\text{Var}\{y_n\} = \bar{y}_{n,p}(\mu) + \bar{y}_{n,d} = b_n e^{-l_n(\mu)} + 2r_n.$$

Thus the precorrected measurements ( $y_n$ 's) are clearly not Poisson distributed. One simple approach to image reconstruction would be to assume that the measurements have a Poisson distribution with means  $\bar{y}_n(\mu)$ , even though this model is incorrect we refer to this approach as the "ordinary Poisson" model. To illustrate the inaccuracy of the ordinary Poisson measurement model for  $y_n$ 's, we have performed a small Monte Carlo simulation similar to [2]. The circles in Fig. 1 show a simulated histogram for  $y_n$  generated by a pseudo-random number generator in accordance with the distribution described above (for 50,000 realizations) where  $\bar{y}_{n,p} = 8$  and  $\bar{y}_{n,d} = r_n = 1$  (corresponding to 12.5 percent randoms).

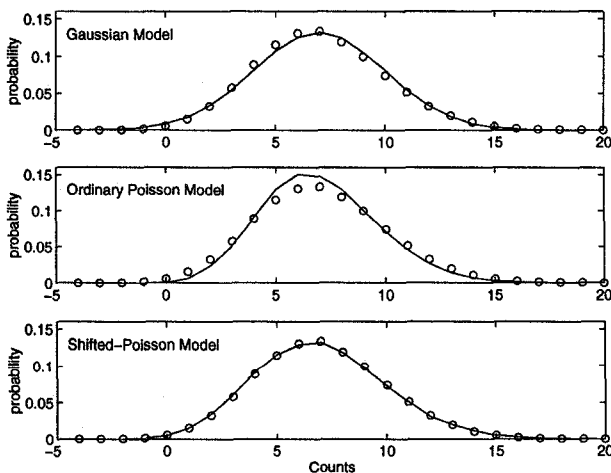


Fig. 1. Comparison of a) Gaussian, b) ordinary Poisson and c) shifted Poisson models (-) (with the moments matched to the moments of precorrected measurements), with the empirical distribution (o) of precorrected measurements.

Fig. 1a shows the approximation based on Gaussian distribution model with mean ( $\bar{y}_n$ ) and variance ( $\bar{y}_n + 2r_n$ ).

Fig. 1b shows the ordinary Poisson (OP) model where approximation is based on a Poisson model with mean ( $\bar{y}_n$ ), the ideal mean. Lastly, Fig. 1c shows the approximation based on a Poisson model with mean ( $\bar{y}_n + 2r_n$ ) and then shifted by  $-2r_n$ . The resultant approximation corresponds to a model with mean and variance that match both first and second order moments of  $y_n$ . The last approximation corresponds to our proposed "shifted" Poisson (SP) model and it has the best agreement with the precorrected measurement  $y_n$ . For large means, the shifted Poisson distribution is also approximately Gaussian by the Central Limit Theorem. However in transmission tomography the the projections through high attenuation regions (which is usually the region of interest) have lower count rates, and the above example illustrates that for low count rates the Gaussian approximation is less accurate than the SP model.

### III. OBJECTIVE FUNCTIONS

Let  $y_n$  given in (1) be a realization of statistically independent random variables  $\{Y_n\}_{n=1}^N$ . The exact log-likelihood for  $\mu$  can be formulated using total probability:

$$\begin{aligned} L(\mu) &= \log P(Y = y; \mu) \\ &= \sum_{n=1}^N \log \left( \sum_{m=[-y_n]_+}^{\infty} \frac{\bar{y}_{n,p}(\mu)^{(y_n+m)} \bar{y}_{n,d}^{(m)}}{(y_n+m)! m!} \right) \\ &\quad - (\bar{y}_{n,p}(\mu) + \bar{y}_{n,d}). \end{aligned} \quad (4)$$

This exact log-likelihood function contains infinite summations, so practically one cannot compute the exact value. In the light of the Monte Carlo simulation that we have performed in previous section, one can develop different approximations to the exact log-likelihood function. We describe three approximations below.

The quadratic approximation to the exact log-likelihood function results in the Weighted Least Squares objective function  $L_{WLS}(\mu)$  [8]:

$$L_{WLS}(\mu) = -\frac{1}{2} \sum_{n=1, y_n > 0}^N (l_n(\mu) - \hat{l}_n)^2 \frac{1}{\hat{\sigma}_n^2}, \quad (5)$$

where  $\hat{l}_n = \log(b_n/y_n)$  is the method-of-moments estimate of the *line integral of the attenuation*  $l_n(\mu)$  and the  $n$ th weighting factor  $\hat{\sigma}_n^2 = (y_n + 2r_n)/y_n^2$  is an estimate of the variance of  $\hat{l}_n(y_n)$  based on a second-order Taylor expansion around  $\hat{l}_n(\bar{y}_n)$ . This weighting is critical for the WLS method. The errors corresponding to projections with large values of  $y_n$  are weighted more heavily. These projections pass through less dense objects and consequently have higher SNR values.

The ordinary Poisson model for the precorrected data  $y_n$  with mean  $\bar{y}_n(\mu) = b_n e^{-l_n(\mu)}$  leads to the OP objective function:

$$L_{OP}(\mu) = \sum_{n=1}^N y_n \log(b_n e^{-l_n(\mu)}) - (b_n e^{-l_n(\mu)}), \quad (6)$$

disregarding constants independent of  $\mu$ .

In the light of Fig. 1c, a better approach, which matches both the first and the second order moments<sup>2</sup>, is to approximate the quantities  $(y_n + 2r_n)$  as realizations of independent Poisson random variables with means  $(\bar{y}_n(\mu) + 2r_n)$ . This model leads to our proposed SP objective function:

$$L_{SP}(\mu) = \sum_{n=1}^N h_n(l_n(\mu)), \quad (7)$$

where

$$h_n(l) = (y_n + 2r_n) \log(b_n e^{-l} + 2r_n) - (b_n e^{-l} + 2r_n).$$

(It can be shown that  $L_{WLS}(\mu)$  corresponds to the summations of second order Taylor series expansion of  $h_n(l_n(\mu))$  about  $h_n(\hat{l}_n)$  where  $\hat{l}_n = \log(b_n/y_n)$ .)

Fig. 2 compares the actual log-likelihood function and the approximations for noiseless data as a function of a single projection across the reconstructed image for 5% randoms rate.  $L_{SP}(\mu)$  agrees fairly well with the exact log-likelihood  $L(\mu)$ ; however  $L_{WLS}(\mu)$  and  $L_{OP}(\mu)$  depart significantly from the exact log-likelihood function. Note that all the curves have the maximum at the same point  $\hat{l}_n = \log(b_n/y_n)$ . This is due to the fact that all estimators works perfectly with the noiseless data (*i.e.*  $y_n = \bar{y}_n$ ). For noisy data the maximum of each curve will exhibit a variation around its mean value. We also observed that, for noisy data  $L_{SP}(\mu)$  agrees with the exact log-likelihood  $L(\mu)$  better than the other models.

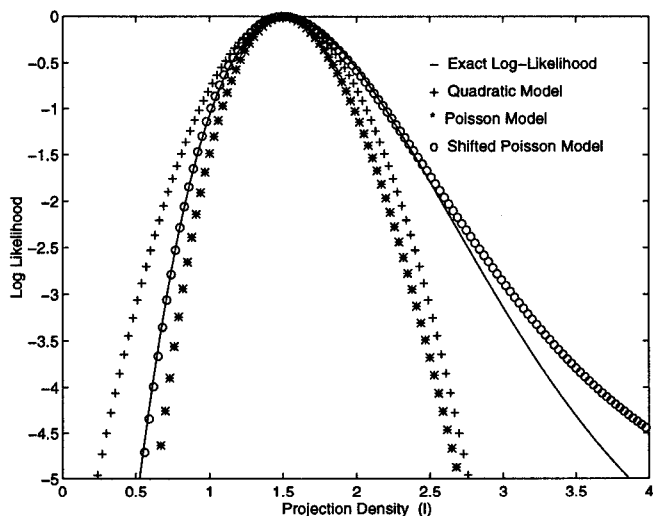


Fig. 2. Comparison of exact log-likelihood function with objective functions of different models for a single projection ray. The proposed shifted Poisson model agrees with exact log-likelihood better than the other models.

<sup>2</sup>Key Difference: Both  $L_{WLS}$  and  $L_{SP}$  match two moments, but in WLS the 2<sup>nd</sup> moment is "fixed" and it is the moments of  $\hat{l}_n$  rather than  $y_n$ .

## IV. BIAS-VARIANCE ANALYSIS

To analyze the bias and variance of each estimator analytically, we used the analytic approximations in [4]. For this purpose we considered a highly simplified version of transmission tomography where the unknown is a scalar parameter. This simplified problem provides insight into the estimator bias and variance without the undue notation of the multi-parameter case. Because of the space considerations we are not able to give detailed formulas for mean and variance approximations of different estimators. Using Cauchy-Schwarz inequality to the approximations of variances, we have shown analytically that SP estimator yields a lower variance than the OP estimator [9]. This result agrees with the 2-D simulations shown next.

### A. 2D Simulations

To study bias and variance properties of each estimator described above, we performed 2D simulations. For  $\mu$  we used the synthetic attenuation map shown in Fig. 3, which represents a human *abdomen* with linear attenuation coefficient 0.0096/mm. The image was a 128 by 64 array of 4.5 mm pixels. We simulated a PET transmission scan with 192 radial bins and 256 angles uniformly spaced over 180 degrees. The  $a_{nj}$  factors correspond to 6 mm wide strip integrals on 3 mm center-to-center spacing. The  $b_n$  factors were generated using pseudo-random log-normal variates with standard deviation of 0.3 to account for detector efficiency variations, and scaled so that  $\sum_n \bar{y}_n$  was one million counts. The  $r_n$  factors corresponded to a uniform field of 5% random coincidences. Pseudo-random transmission measurements were generated according to (2) and (3). For regularization, we used the modified quadratic penalty of [3], which improves the spatial resolution uniformity.



Fig. 3. Simulated abdomen attenuation map.

We generated 100 independent realizations of the transmission measurements. For each measurement realization, an estimate of the attenuation map was reconstructed, using 20 iterations of the grouped-coordinate ascent algorithm [5] applied to the objective functions (5), (6) and (7). We computed both the sample mean and sample standard deviation images for all three methods.

Fig. 4 shows horizontal profiles through the sample mean images. These profiles show that WLS is systematically negatively biased, whereas the OP and SP models are free of systematic bias. (One should note that the overshoot at the edges is due to the quadratic penalty used in the reconstruction. Even with noiseless data, this blurring effect will still be present.)

To study the standard deviation, we calculated the *ratio* of sample standard deviation images of different estima-

tors. Fig. 5 shows the histogram of the ratio of standard deviations, over all interior pixels. The OP model yields about 20% higher standard deviation than the SP model. In other words, to achieve the same noise level, the OP method would require about 40% greater scan time than our proposed SP method.

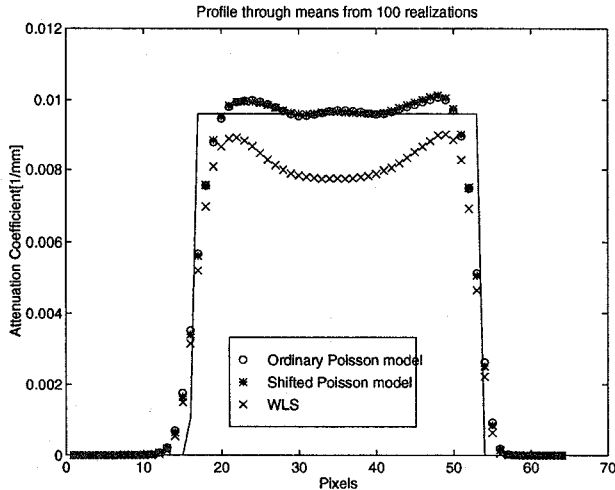


Fig. 4. The WLS method has a systematic negative bias. The ordinary Poisson and the shifted Poisson models yield negligible bias.

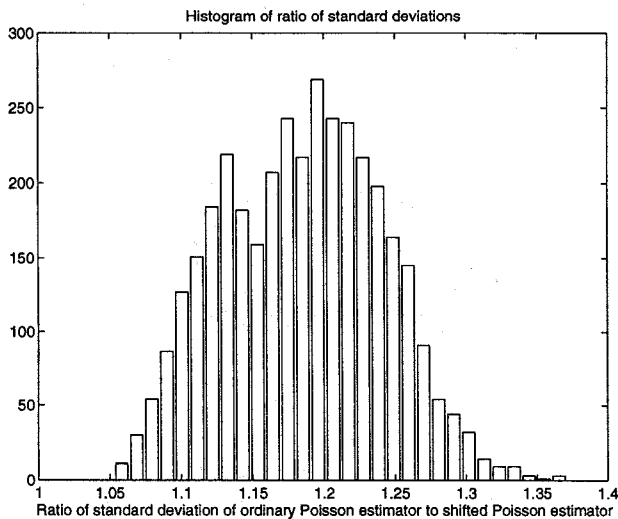


Fig. 5. Ordinary Poisson model yields, on the average 20%, higher standard deviation than proposed shifted Poisson model.

We performed additional simulations for the same PET system described above, but this time using the synthetic attenuation map shown in Fig. 6, which represents a human *thorax* with linear attenuation coefficients 0.0165/mm, 0.0096/mm, and 0.0025/mm, for bone, soft tissue, and lungs, respectively. Fig. 7 shows the profiles through the sample mean images. Similar to the observations from Fig. 4, the WLS estimator is negatively biased. Fig. 8 shows the histogram of the ratio of standard deviations.

Again the OP model yields, on the average 11%, higher standard deviation than SP model.

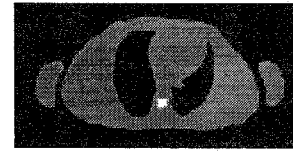


Fig. 6. Simulated thorax attenuation map.

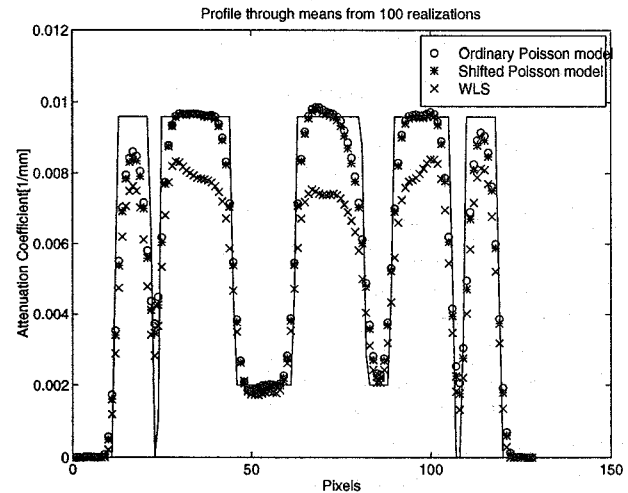


Fig. 7. The WLS method has a systematic negative bias. The ordinary Poisson and the shifted Poisson models yield negligible bias.

Note that the process of real-time subtraction of the delayed coincidence events from prompt events can lead to some negative values in the precorrected data. Since the mean of precorrected measurements is nonnegative, a natural choice might be to threshold the negative values in the precorrected data to zero before applying the maximization algorithm. (Moreover the likelihood objective function for emission tomography is not guaranteed to be concave if the measurements have negative values.) To study further the effects of zero-thresholding the data, we performed additional 2D simulations with zero-thresholded data and using the above phantoms [9]. The results have shown that the OP estimator was systematically negatively biased, especially for interior regions of the reconstructed image, and still had higher standard deviation than the SP estimator.

### B. Estimates of the AC rates ( $\hat{r}_n$ )

One needs to know the mean of the AC events ( $r_n$ ) in order to compute  $L_{SP}(\mu)$ . Since the  $r_n$  terms are not readily available from the real (precorrected) data, some estimates of the randoms must be used.

Fig. 9 displays the scatter plot of real delayed coincidence sinograms for blank scan and transmission scan data. Each point in the plot corresponds to a specific detector pair. The similarity of both delayed coincidence measurements suggests that one can simultaneously acquire the delayed

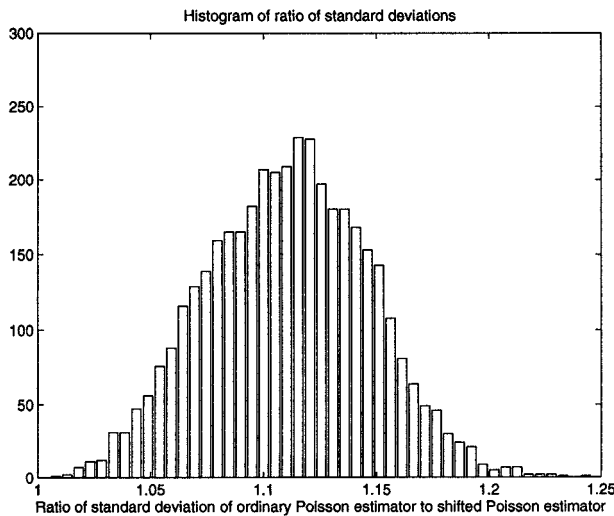


Fig. 8. Ordinary Poisson model yields, on the average 12%, higher standard deviation than shifted Poisson model.

coincidence events during the blank scan and use it (after properly normalizing for different scan durations) as an estimate of the AC rates for different transmission scans performed on the same PET system. For emission tomography the “singles” method suggested by Casey [1] can be used to obtain an estimate of the rate of AC events.

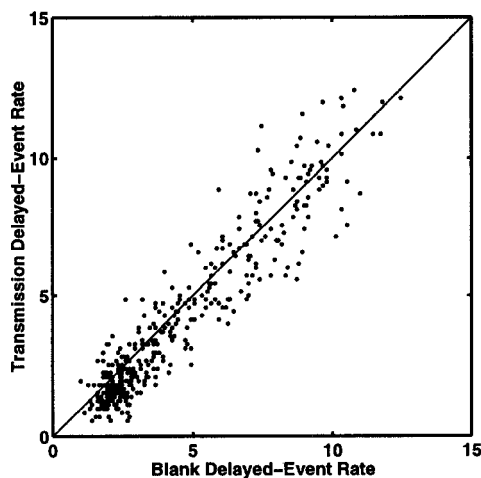


Fig. 9. Scatter plot of delayed coincidence event of blank and transmission scans.

To test the robustness of the SP estimator to the errors in estimates of AC rates, we performed additional simulations using the phantoms and the PET system described previously. We observed that, even using a constant value of  $\bar{r} = (1/N) \sum_n r_n$ , as an estimate of the AC event rates did not introduce any systematic bias for SP estimator and increased standard deviation only slightly (around 2%). Since the AC rates of transmission and blank scans are highly correlated, using AC rates obtained from blank scan measurements should yield better AC rates estimates, resulting in very similar bias/variance performance as the estimators obtained by using true AC rates.

## V. CONCLUSIONS

When the AC events are precorrected in PET, the measurement statistics are no longer Poisson. For transmission scans, WLS method and ML method based on ordinary Poisson (OP) model lead to systematic bias and higher variance, respectively. Thus, we proposed a shifted Poisson (SP) model for measurement statistics which matches both the first and second-order moments. The corresponding log-likelihood function was shown to have better agreement with the exact log-likelihood function than the other methods above. Using Taylor approximation, implicit function theorem and chain rule one can obtain analytic expressions for the mean and the variance of the different estimators and using Cauchy-Schwarz inequality, it can be shown analytically that OP model leads to higher variance than both the WLS and the SP methods [9]. We performed 2D simulations to support this observation. Another observation was that the WLS method leads to unacceptably high systematic bias, especially for low count rates. We have also shown that the effect of zero-thresholding the precorrected data leads to systematic negative bias for the OP model estimator. In addition, OP estimator still had a higher variance than the SP estimator [9].

Lastly, the effect of using estimates of AC rates, for the SP estimator was investigated. We observed that the SP model is robust to errors in the estimates of AC events. Namely, even using constant AC rates resulted in only a slight increase in the standard deviation without any systematic bias [9].

It should be noted that the SP model does not increase the computation requirements for the maximization algorithm over that of the OP model.

## REFERENCES

- [1] M E Casey and E J Hoffman. Quantitation in positron emission computed tomography: 7 A technique to reduce noise in accidental coincidence measurements and coincidence efficiency calibration. *J. Comp. Assisted Tomo.*, 10(5):845–850, 1986.
- [2] J A Fessler. Penalized weighted least-squares image reconstruction for positron emission tomography. *IEEE Tr. Med. Im.*, 13(2):290–300, June 1994.
- [3] J A Fessler. Resolution properties of regularized image reconstruction methods. Technical Report 297, Comm. and Sign. Proc. Lab., Dept. of EECS, Univ. of Michigan, Ann Arbor, MI, 48109-2122, August 1995.
- [4] J A Fessler. Mean and variance of implicitly defined biased estimators (such as penalized maximum likelihood): Applications to tomography. *IEEE Tr. Im. Proc.*, 5(3):493–506, March 1996.
- [5] J A Fessler, E P Ficaro, N H Clinthorne, and K Lange. Grouped-coordinate ascent algorithms for penalized-likelihood transmission image reconstruction. *IEEE Tr. Med. Im.*, 1996. To appear.
- [6] E J Hoffman, S C Huang, M E Phelps, and D E Kuhl. Quantitation in positron emission computed tomography: 4 Effect of accidental coincidences. *J. Comp. Assisted Tomo.*, 5(3):391–400, 1981.
- [7] E Ü Mumcuoğlu, R M Leahy, and S R Cherry. Bayesian reconstruction of PET images: methodology and performance analysis. *Phys. Med. Biol.*, 1996.
- [8] K Sauer and C Bouman. A local update strategy for iterative reconstruction from projections. *IEEE Tr. Sig. Proc.*, 41(2):534–548, February 1993.
- [9] M Yavuz and J A Fessler. Objective functions and reconstruction algorithms for randoms-precorrected PET scans, 1996. In preparation.