Maximum Likelihood Emission Image Reconstruction for Randoms-precorrected PET Scans

Mehmet Yavuz and Jeffrey A. Fessler

Dept. of Electrical Engineering and Computer Science , University of Michigan, Ann Arbor, MI 48109-2122

Abstract

Most PET scans are compensated for accidental coincidence (AC) events by real-time subtraction of delayed-window coincidences. Real time subtraction of delayed coincidences compensates for the average of AC events, but also destroys the Poisson statistics. Moreover, negative values result during the real-time subtraction which would cause conventional penalized maximum likelihood algorithms to diverge, and setting these negative values to zero introduces a systematic positive bias. We have previously developed and compared two new methods for reconstructing transmission scans from randoms precorrected measurements: one based on a "shifted Poisson" (SP) model, and the other based on saddle-point (SD) approximations. Simulations and experimental phantom studies of transmission scans showed that both SP and SD methods lead to significantly lower variance than the conventional maximum likelihood methods (based on the ordinary Poisson (OP) model). We have now extended these methods to emission scans. In situations like 3D PET emission scans (with low counts per ray but many total counts and high randoms rates), we show that the proposed methods not only avoid the systematic positive bias of OP method but also lead to significantly lower variance. The new methods offer improved image reconstruction in PET through more realistic statistical modeling, yet with negligible increase in computation over the conventional OP method.

I. INTRODUCTION

In PET emission scans, a significant portion of the collected data generally is accidental coincidence (AC) events, and these are a primary source of background noise [1–3]. Moreover, AC rates increase as the *square* of the amount of radio-isotope injected to the patient, while true coincidences increase only *linearly* with the radio-isotope concentration. This count rate limitation, along with detector deadtime, determines the upper limit on the injected radio-isotope dose for many PET studies.

In conventional PET scans the system detects coincidence events during two time windows: the "prompt" window and the "delayed" window, and the data are pre-corrected for AC events by *real-time* subtraction of delayed window coincidences [2]. Each such pre-corrected measurement is the difference of two independent Poisson random variables, which compensates in mean for AC events, but which also increases the measurement variance. Moreover, negative values

result during the real-time subtraction of delayed coincidences (especially in high-resolution 3-D PET systems where counts per ray can be very low). Such negative values would cause conventional penalized maximum likelihood algorithms to diverge. Setting the negative values to zero alleviates this problem but introduces a systematic positive bias in the resulting images [1,4].

We have previously developed and compared two methods for reconstructing *transmission* scans from randoms precorrected measurements: one based on a "shifted Poisson" (SP) model [5–7], and the other based on saddle-point (SD) approximations [6,7]. In this paper we examine these methods in *emission* scans. In situations like 3D PET emission scans (with low counts per ray but many total counts), we show that the proposed methods not only avoid the systematic positive bias of the conventional approach, but also lead to significantly lower variance. Emission scan results also show an unexpected benefit of the SD method that had been unrealized in our previous investigation of the transmission reconstruction problem.

II. MEASUREMENT MODEL AND EXACT LOG-LIKELIHOOD

Let $Y = [Y_1, \ldots, Y_N]'$ denote the vector of *precorrected* measurements, where " ′ " denotes vector and matrix transpose. The precorrected measurement for the nth coincidence detector pair is $Y_n = Y_n^{\text{prompt}} - Y_n^{\text{delay}}$, where Y_n^{prompt} and Y_n^{delay} are the number of coincidences within the prompt and delayed windows, respectively.

Let $\lambda = [\lambda_1, \ldots, \lambda_P]'$ denote the vector of unknown radio-isotope concentration values. For emission scans, we assume that Y_n^{prompt} and Y_n^{delay} are statistically independent Poisson random variables with means $\bar{y}_n^{\text{p}}$ and $\bar{y}_n^{\text{d}}$ given respectively as:

$$\text{E}\left\{Y_n^{\text{prompt}}\right\} = \bar{y}_n^{\text{p}}(\lambda) = \sum_{j=1}^{P} g_{nj}\lambda_j + r_n \qquad (1)$$

$$\text{E}\left\{Y_n^{\text{delay}}\right\} = \bar{y}_n^{\text{d}} = r_n, \qquad (2)$$

where $r_n > 0$ denote the mean of the AC events for the nth ray and $G = \{g_{nj}\}$ represents the system matrix including ray-dependent factors such as attenuation and detector efficiency.

Let $y = [y_1, \ldots, y_N]'$ be an observed realization of Y. Since the measurements are independent, one can express the exact log-likelihood as follows [6]:

$$L(\lambda) = \sum_{n=1}^{N} h_n(l_n(\lambda)), \qquad (3)$$

with $l_n(\lambda) = \sum_{j=1}^{P} g_{nj}\lambda_j$ and ignoring constants independent of λ throughout:

$$h_n(l) \triangleq \log\left(\sum_{m=[-y_n]_+}^{\infty} \frac{(l+r_n)^{y_n+m}}{(y_n+m)!} \frac{r_n^m}{m!}\right) - (l+2r_n), \quad (4)$$

where $[x]_+ = x$ if $x > 0$ and is 0 otherwise.

Since image reconstruction is ill conditioned, we combine a roughness penalty $R(\lambda)$ with the log-likelihood to form a penalized-likelihood objective function: $\Phi(\lambda) = L(\lambda) - R(\lambda)$. The goal is to estimate λ by maximizing $\Phi(\lambda)$ over the nonnegative cone. The exact log-likelihood function (4) has a complicated form because of the lower and upper summation limits. Although one can express the exact log-likelihood function using modified complex Bessel functions [8, 9], it does not yield in practical maximization methods [10]. Next we describe practical and yet accurate approximations to the exact log-likelihood.

III. APPROXIMATIONS TO THE EXACT LOG-LIKELIHOOD

In this section, we briefly review the three practical approximations to $L(\lambda)$. All log-likelihood approximations have the form (3) for different choices for $h_n(l, y_n)$.

A. Ordinary Poisson (OP) Approximation

The conventional approach is to ignore the random coincidences by assuming that $\{[Y_n]_+\}_{n=1}^{N}$ are distributed as independent Poisson random variables with means $\bar{y}_n(\lambda)$ given by (1). The log-likelihood $L_{\text{OP}}(\lambda)$ corresponding to this OP approximation is of the form (3) with

$$h_n^{\text{OP}}(l) = [y_n]_+ \log(l) - l. \quad (5)$$

The thresholding $[y_n]_+$ ensures concavity of the OP objective function.

B. Shifted Poisson (SP) Approximation

A better approach is to match both the first and second moments by approximating the random variables $\{Y_n + 2r_n\}_{n=1}^{N}$ as having Poisson distributions with means $\{\bar{y}_n(\lambda) + 2r_n\}$. This idea leads to the SP approximation $L_{\text{SP}}(\lambda)$ [11] of the form (3) with

$$h_n^{\text{SP}}(l) = [y_n + 2r_n]_+ \log(l + 2r_n) - (l + 2r_n), \quad (6)$$

where again the zero thresholding of $(y_n + 2r_n)$ ensures that the objective function is concave.

C. Saddle-point (SD) Approximation

An better alternative to the previous approximations for the exact pmf of precorrected measurements is to make second order Taylor series approximations in the z-transform domain (*i.e.*, on the probability generating function) and then to perform the inverse transform. For emission tomography this

saddle point (SD) approximation [6,7] is of the form (3) with:

$$h_n^{\text{SD}}(l) = y_n \log\left(\frac{l+r_n}{z_n+u_n(l)}\right) - l + u_n(l) - \frac{1}{2}\log(u_n(l)), \quad (7)$$

where

$$z_n = \begin{cases} y_n + 1, & y_n \geq 0 \\ y_n - 1, & y_n < 0, \end{cases} \quad (8)$$

and

$$u_n(l) = \sqrt{z_n^2 + 4(l+r_n)r_n}. \quad (9)$$

D. Exact Log-likelihood for Prompt Data

If one has access to the prompt data y_n^{p} separately, then the exact log-likelihood $L_{\text{PR}}(\mu)$ can be written in the form (3) with

$$h_n^{\text{PR}}(l) = y_n^{\text{p}} \log(l + r_n) - (l + r_n). \quad (10)$$

We include the exact log-likelihood model for prompt data for comparing the bias and variance results with the methods for randoms-precorrected data.

IV. CONCAVITY AND LOG-LIKELIHOOD MAXIMIZATION

The second partial derivatives of the OP (5) and the SP (6) objective functions and the PR log-likelihood (10) are:

$$-\frac{\partial^2}{\partial \lambda_j \lambda_k} L(\lambda) = \sum_{n=1}^{N} g_{nj} g_{nk} \frac{x_n}{(\bar{y}_n(\lambda) + d_n)^2},$$

with

$$d_n \triangleq \begin{cases} 0, & \text{OP} \\ 2r_n, & \text{SP} \\ r_n, & \text{PR}, \end{cases} \quad \text{and} \quad x_n \triangleq \begin{cases} [y_n]_+, & \text{OP} \\ [y_n + 2r_n]_+, & \text{SP} \\ y_n^{\text{p}}, & \text{PR}. \end{cases}$$

Thus, $L(\lambda)$ is globally concave when $x_n > 0$, hence the zero thresholds in (5, 6). Since the "thresholding function" $[y_n]_+$ is not differentiable at $y_n = 0$, it is difficult to derive accurate analytic approximations for the mean and variance of these estimators. However, one can explain the overall effect of zero-thresholding as follows: setting negative precorrected data values to zero increases the mean of the precorrected data. For the emission problem the data is linearly related to emission rates, thus the increase in the mean value of the precorrected data causes the estimator to introduce a systematic *positive bias* for the estimated emission rates [1, 4]. A detailed concavity analysis [10] of the SD method shows that the $h_n^{\text{SD}}(l)$'s are concave for $l \in [0, \infty)$ without any zero thresholding, which makes it free of any systematic positive bias.

In this study we used the paraboloid surrogates maximization algorithm [12] which requires certain convexity conditions of the derivatives of the $h_n^{\text{SD}}(l)$'s [13]. In the paraboloid surrogates coordinate ascent (PSCA) maximization method [12] one maximizes a surrogate function (which is parabolic). The surrogate function is composed such that the log-likelihood function converges to the true maximizer. This is

achieved by forming a summation of 1-D surrogate functions. Then we use the fast coordinate ascent method for maximizing the parabolic function. The optimum curvature for the PSCA method requires certain conditions from the second and third derivatives of $h_n(l)$ [12], which are satisfied for the OP, SP and PR methods [10]. For the SD method we showed that these conditions are satisfied almost for all values of l and y_n [10]. For the remaining cases we use the maximum curvature [13], which ensures monotonicity based on the generalized mean value theorem for twice differentiable functions [10].

V. RESULTS

To study the bias and variance properties of the estimators that are based on the above likelihood approximations, we performed 2D simulations. For λ we used the synthetic emission phantom shown in Fig. 1. The spine, lungs, soft tissue, and heart had relative radioactivity concentrations of 0, 1, 2 and 4 respectively. The sinograms had 200 radial bins and 300 angles uniformly sampled over 180 degrees. The system geometry was 2.8 mm wide strip integrals and 2.8 mm ray spacing. The reconstructed images were 64 by 64 with 9 mm pixels. The r_n factors corresponded to a uniform field of 50% random coincidences.

We generated 300 pseudo-random emission measurements according to (1) and (2). For each realization, we reconstructed an estimate of the emission phantom using 30 iterations of the paraboloid surrogates algorithm [12] applied to objective functions (5), (6) and (7).

For regularization, we used the modified quadratic penalty [14, 15], which improves resolution uniformity and enables matching of the spatial resolutions of different methods. We matched the resolution of the reconstructed images for all methods to 1.9 pixels FWHM.

Since in these simulations we had access to Y_n^{prompt} and Y_n^{delay} values separately, we also performed conventional penalized maximum likelihood reconstruction with prompt counts (PR) for comparison purposes. In the PR case the data is not precorrected for randoms and we have access to r_n values separately. Thus, this method is expected to perform better than the randoms-precorrected methods. We include this method in our simulations for comparison purposes only.

Figs. 1 and 2 show the sample mean and standard deviation images of different methods (including filtered back projection (FBP) method for comparison) for a total of 50K counts[2]. Fig. 3 shows the profiles through the sample mean images and Fig. 4 shows the histogram of the ratio of the standard deviation of different methods with respect to the PR method. The OP method results in severe bias and the SP results in some bias in the reconstructed images. However, the SP and SD methods yield similar standard deviations. Fig. 5 shows the profiles through the sample mean images for a total of 500K counts. Again, the OP method results in systematic positive bias while SP and SD methods are free of such bias. Also, Fig. 6 shows the histogram of the ratio of the standard deviation of different

---

[2]To emulate 3D PET emission scans with very low counts per ray.

methods with respect to the PR method. Both, the SP and SD methods yield lower standard deviation than OP method.

VI. CONCLUSIONS

For randoms pre-corrected PET emission measurements, we have described practical approximations for the complicated exact log-likelihood. The results at different count levels show that the proposed models not only avoid the systematic positive bias of OP method but also lead to lower variance. The SP model is shown to be slightly biased for emission scans with very low count rates, whereas the SD model is free of any systematic bias and performs almost identically as the exact log-likelihood. The new methods offer improved image reconstruction in PET through more realistic statistical modeling, yet with negligible increase in computation over the conventional OP method.
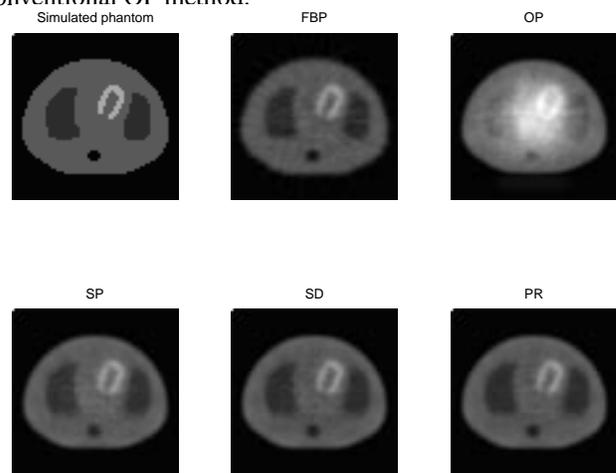

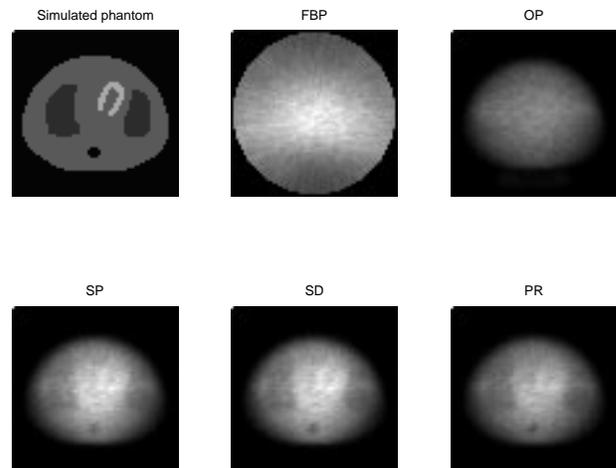Figure 1: Sample mean images of different methods from 300 realization with 50K counts per scan.

Figure 2: Sample standard deviation images of different methods from 300 realization with 50K counts per scan.
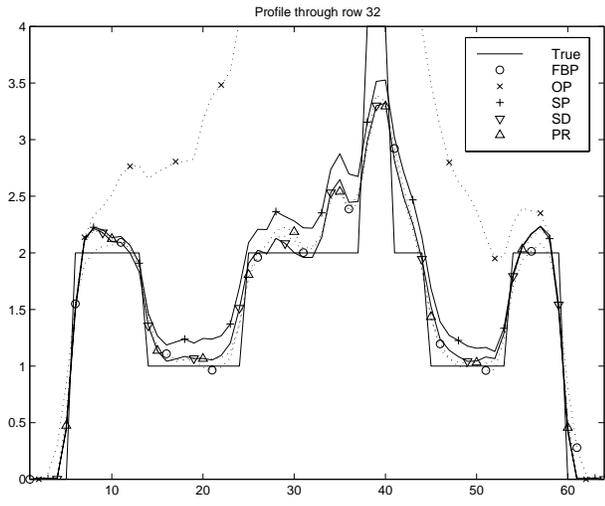
Figure 3: Profile through the sample mean images of different methods from 300 realization with 50K counts per scan.
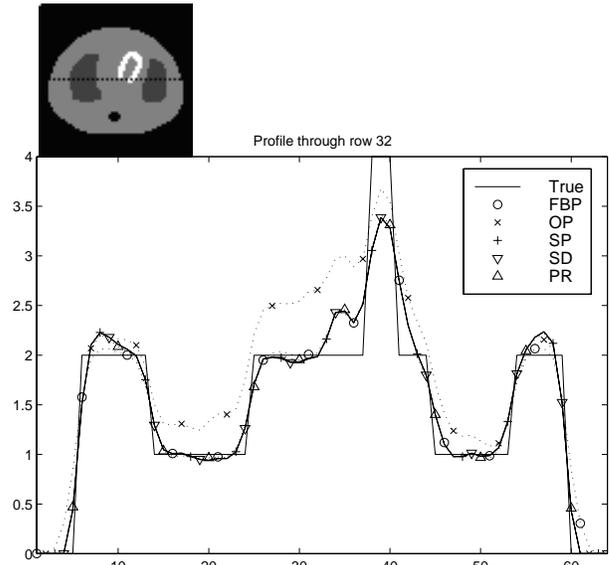

Figure 5: Profile through the sample mean images of different methods from 300 realization with 500K counts per scan.
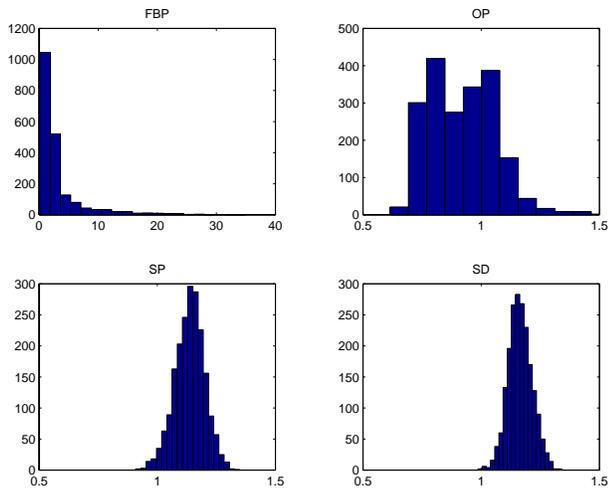

Figure 4: Histogram of the ratio of standard deviation of different methods to the standard deviation of PR method with 50K counts per scan.
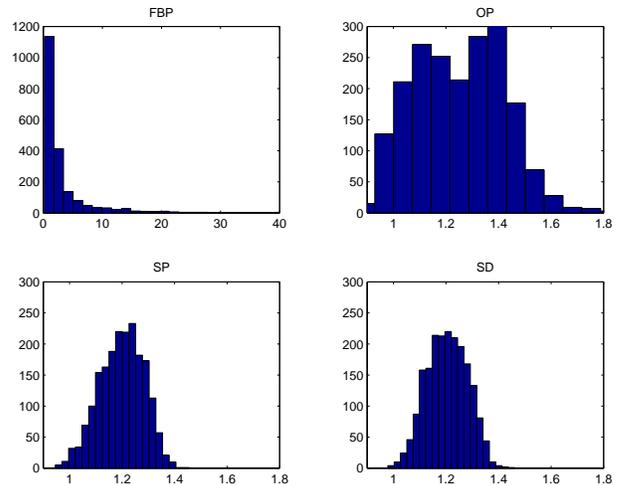

Figure 6: Histogram of the ratio of standard deviation of different methods to the standard deviation of PR method with 500K counts per scan.

VII. REFERENCES

[1] D. G. Politte and D. L. Snyder, "Corrections for accidental coincidences and attenuation in maximum-likelihood image reconstruction for positron-emission tomography," *IEEE Tr. Med. Im.*, vol. 10, no. 1, pp. 82–89, March 1991.

[2] 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.*, vol. 5, no. 3, pp. 391–400, 1981.

[3] T. J. Spinks, T. Jones, M. C. Gilardi, and J. D. Heather, "Physical performance of the latest generation of commercial positron scanner," *IEEE Tr. Nuc. Sci.*, vol. 35, no. 1, pp. 721–725, February 1988.

[4] L. Xuan, C. Comtat, C. Michel, P. Kinahan, M. Defrise, and D. Townsend, "Comparison of 3D reconstruction with OSEM and FORE+OSEM for PET," in *Proc. of the 1999 Intl. Mtg. on Fully 3D Im. Recon. in Rad. Nuc. Med.*, pp. 39–42, 1999.

[5] M. Yavuz and J. A. Fessler, "Objective functions for tomographic reconstruction from randoms-precorrected PET scans," in *Proc. IEEE Nuc. Sci. Symp. Med. Im. Conf.*, volume 2, pp. 1067–71, 1996.

[6] M. Yavuz and J. A. Fessler, "New statistical models for randoms-precorrected PET scans," in *Information Processing in Medical Im.*, J. Duncan and G. Gindi, editors, volume 1230 of *Lecture Notes in Computer Science*, pp. 190–203, Springer Verlag, Berlin, 1997.

[7] M. Yavuz and J. A. Fessler, "Statistical image reconstruction methods for randoms-precorrected PET scans," *Med. Im. Anal.*, vol. 2, no. 4, pp. 369–378, 1998.

[8] J. G. Skellam, "The frequency distribution of the difference between two Poisson variates belonging to different populations," *J. Royal Stat. Soc.*, vol. 109, no. 3, pp. 296, 1946.

[9] G. de Castro, "Note on differences of Bernoulli and Poisson variables," *Portugaliae Math.*, vol. 11, pp. 173–5, 1952.

[10] M. Yavuz, *Objective functions for tomographic reconstruction from randoms-precorrected PET scans*, PhD thesis, Univ. of Michigan, Ann Arbor, MI, 48109-2122, Ann Arbor, MI., 1999.

[11] M. Yavuz and J. A. Fessler, "Penalized-likelihood estimators and noise analysis for randoms-precorrected PET transmission scans," *IEEE Tr. Med. Im.*, vol. 18, no. 8, pp. 665–74, August 1999.

[12] J. A. Fessler and H. Erdoğan, "A paraboloidal surrogates algorithm for convergent penalized-likelihood emission image reconstruction," in *Proc. IEEE Nuc. Sci. Symp. Med. Im. Conf.*, volume 2, pp. 1132–5, 1998.

[13] H. Erdoğan and J. A. Fessler, "Monotonic algorithms for transmission tomography," *IEEE Tr. Med. Im.*, vol. 18, no. 9, pp. 801–14, September 1999.

[14] J. A. Fessler and W. L. Rogers, "Spatial resolution properties of penalized-likelihood image reconstruction methods: Space-invariant tomographs," *IEEE Tr. Im. Proc.*, vol. 5, no. 9, pp. 1346–58, September 1996.

[15] J. W. Stayman and J. A. Fessler, "Regularization for uniform spatial resolution properties in penalized-likelihood image reconstruction," *IEEE Tr. Med. Im.*, vol. 19, no. 6, pp. 601–15, June 2000.