Medical Image Analysis (1998) volume 2, number 4, pp 369–378

Statistical image reconstruction methods for randoms-precorrected PET scans

Mehmet Yavuz and Jeffrey A. Fessler*

Department of EECS, University of Michigan, Ann Arbor, MI 48109-2122, USA

Abstract

Positron emission tomography (PET) measurements are usually precorrected for accidental coincidence events by real-time subtraction of the delayed-window coincidences. Randoms subtraction compensates on average for accidental coincidences but destroys the Poisson statistics. We propose and analyze two new approximations to the exact log-likelihood of the precorrected measurements, one based on a 'shifted Poisson' model, the other based on saddle-point approximations to the measurement of probability mass function (PMF). The methods apply to both emission and transmission tomography; however, in this paper we focus on transmission tomography. We compare the new models to conventional data-weighted least-squares (WLS) and conventional maximum-likelihood methods [based on the ordinary Poisson (OP) model] using simulations and analytic approximations. The results demonstrate that the proposed methods avoid the systematic bias of the WLS method, and lead to significantly lower variance than the conventional OP method. The saddle-point method provides a more accurate approximation to the exact log-likelihood than the WLS, OP and shifted Poisson alternatives. However, the simpler shifted Poisson method yielded comparable bias-variance performance to the saddle-point method in the simulations. The new methods offer improved image reconstruction in PET through more realistic statistical modeling, yet with negligible increase in computation time over the conventional OP method.

Keywords: maximum-likelihood image reconstruction, randoms-precorrected PET, statistical approximations, statistical image reconstruction

Received October 28, 1997; accepted April 1, 1998

1. INTRODUCTION

In PET measurements, accidental coincidence (AC) events are a primary source of background noise. 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. The ratio of total AC events to 'true' events is usually small in transmission scans compared to emission scans. Nevertheless, the effect of AC events becomes severe for regions of high attenuation coefficients, because projections through such regions result in low true coincidence rates. These low count rates can become comparable to AC rates. Thus estimates of the AC events are needed. One can use the 'singles' method (Casey and Hoffman, 1986) for this purpose, however this approach is not used widely because of the necessity for additional hardware, and moreover usually singles rates vary during data acquisition (Ollinger and Fessler, 1997). Therefore, in most PET scans, the AC rates are estimated using delayed-window coincidences and the data are precorrected for AC events by real-time subtraction. Real-time subtraction of delayed-window coincidences compensates on average for AC events, but destroys the Poisson statistics (Hoffman *et al.*, 1981). To avoid this problem, one needs to maintain the transmission and randoms measurements as two separate sinograms (Politte and Snyder, 1991; Mumcuoğlu *et al.*, 1996). However even if a PET system allows one to collect a randoms (delayed coincidences) sinogram separately, this

*Corresponding author
(e-mail: fessler@umich.edu)

process would double the storage space for the acquired data. So in practice most PET centers collect and archive only the randoms-precorrected data. We recommend separate acquisition and storage of delayed coincidences wherever feasible. The purpose of this paper is to provide accurate statistical methods for PET measurements with presubtracted delayed coincidences. Although our analysis and proposed models apply to both emission and transmission tomography, in this paper we focus on transmission tomography.

The exact log-likelihood for randoms-precorrected data is intractable, so we describe and compare several approximations. For completeness, we first review the data-weighted least-squares (WLS) method and the log-likelihood for the ordinary Poisson (OP) model for PET measurements. Then, we introduce a new 'shifted' Poisson (SP) model (Yavuz and Fessler, 1996) which matches both the first- and second-order moments of the model to the underlying statistics of the precorrected data. We derive approximate analytic expressions for the variance of the different estimators and use the Cauchy–Schwarz inequality to show analytically that the proposed SP method yields lower variance than the OP method.

Secondly, we introduce a new saddle-point (SD) approximation for the probability mass function (PMF) of precorrected measurements. The corresponding log-likelihood function is shown to have better agreement with the exact log-likelihood than the previous approximations. We apply the fast grouped-coordinate ascent algorithm (Fessler *et al.*, 1997) (with a few simple modifications) to maximize the proposed saddle-point objective function.

We also show results of two-dimensional (2-D) simulations showing that the WLS method leads to a systematic bias and that the OP method leads to a higher variance than SP and SD methods. We also observe that the SP and SD methods yield equivalent bias/variance performance, whereas SP requires less computation time. The contribution of this work lies in the fact that the proposed methods offer significant improvements in accuracy with minor increases in computation time.

2. MEASUREMENT MODEL

In conventional PET scans, the data are precorrected for AC events by real-time subtraction of the delayed-window coincidences (Hoffman *et al.*, 1981). The system detects coincidence events during two time windows: the '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 delayed window, the corresponding sinogram bin is decremented, so the resultant 'precorrected' measurements 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 on average for AC events, but it also increases the variance of the measurement by an amount equal to the mean of AC events.

Let $\underline{Y} = [Y_1, \ldots, Y_N]'$ denote the vector of precorrected measurements. The precorrected measurement for the nth coincidence detector pair is

$$Y_n = Y_n^{\text{prompt}} - Y_n^{\text{delay}}, \tag{1}$$

where Y_n^{prompt} and Y_n^{delay} are the number of coincidences within the prompt and delayed windows respectively. Let $\mu = [\mu_1, \ldots, \mu_M]'$ denote the vector of unknown linear attenuation coefficients. For transmission 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 $\bar{y}_n^{\text{d}}$ respectively as

$$\mathrm{E}\{Y_n^{\text{prompt}}\} = \bar{y}_n^{\text{p}}(\mu) = b_n \mathrm{e}^{-l_n(\mu)} + r_n \tag{2}$$

$$\mathrm{E}\{Y_n^{\text{delay}}\} = \bar{y}_n^{\text{d}} = r_n, \tag{3}$$

where $l_n(\mu) = \sum_{j=1}^{M} a_{nj}\mu_j$ is the total attenuation between the nth 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 $r_n \geq 0$ factors denote the mean of AC events.

Since Y_n^{prompt} and Y_n^{delay} are statistically independent and Poisson:

$$\mathrm{E}\{Y_n\} = \bar{y}_n^{\text{p}}(\mu) - \bar{y}_n^{\text{d}} = b_n \mathrm{e}^{-l_n(\mu)}, \tag{4}$$

$$\mathrm{Var}\{Y_n\} = \bar{y}_n^{\text{p}}(\mu) + \bar{y}_n^{\text{d}} = b_n \mathrm{e}^{-l_n(\mu)} + 2r_n. \tag{5}$$

3. EXACT LOG-LIKELIHOOD

Let $\underline{y} = [y_1, \ldots, y_N]'$ be a realization of statistically independent random variables $\underline{Y}$ given in (1). Under the usual assumption of independence between different rays, one can express the exact distribution of $\underline{Y}$ using total probability:

$$\begin{aligned} P(\underline{Y} = \underline{y}; \mu) &= \prod_{n=1}^{N} \sum_{m=0}^{\infty} P(Y_n^{\text{prompt}} = y_n + m; \mu) \\ &\quad \times P(Y_n^{\text{delay}} = m) \\ &= \prod_{n=1}^{N} \sum_{m=\lfloor -y_n \rfloor_+}^{\infty} \frac{\left[\bar{y}_n^{\text{p}}(\mu)\right]^{y_n+m} \mathrm{e}^{-\bar{y}_n^{\text{p}}(\mu)}}{(y_n+m)!} \frac{r_n^m \mathrm{e}^{-r_n}}{m!}, \end{aligned} \tag{6}$$

where $\lfloor x \rfloor_+ = x$ if $x > 0$ and is 0 otherwise. The exact log-likelihood for μ becomes

$$\begin{aligned} L(\mu) &= \log P(\underline{Y} = \underline{y}; \mu) \\ &= \sum_{n=1}^{N} \log \left(\sum_{m=\lfloor -y_n \rfloor_+}^{\infty} \frac{\left[\bar{y}_n^{\mathrm{p}}(\mu)\right]^{y_n+m}}{(y_n+m)!} \frac{r_n^m}{m!} \right) \\ &\quad - (\bar{y}_n^{\mathrm{p}}(\mu) + r_n). \end{aligned} \tag{7}$$

Since image reconstruction is ill conditioned, usually one includes a roughness penalty $R(\mu)$ in the objective function. From the Bayesian point of view, this roughness penalty can be thought of as a log-prior for μ. Combining this penalty with the log-likelihood yields a penalized-likelihood objective function:

$$\Phi(\mu) = L(\mu) - R(\mu). \tag{8}$$

The goal is to estimate μ by maximizing $\Phi(\mu)$ over the non-negative cone:

$$\hat{\mu} = \arg\max_{\mu \geq 0} \Phi(\mu). \tag{9}$$

Since the exact log-likelihood function in Equation (7) contains infinite summations, the above maximization is intractable. The following two sections develop tractable yet accurate approximations to $L(\mu)$.

4. SIMPLE APPROXIMATIONS TO THE EXACT LOG-LIKELIHOOD

In this section, we first review the conventional approximations to $L(\mu)$: the WLS model and the conventional OP model. Then we introduce the SP model (Yavuz and Fessler, 1996).

4.1. Quadratic approximations

The quadratic approximation to the exact log-likelihood function results in the data-weighted least-squares objective function $L_{\mathrm{WLS}}(\mu)$ (Sauer and Bouman, 1993):

$$L_{\mathrm{WLS}}(\mu) = -\frac{1}{2} \sum_{n=1,\ y_n>0}^{N} (l_n(\mu) - \hat{l}_n)^2 \frac{1}{\hat{\sigma}_n^2}, \tag{10}$$

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 $\hat{\sigma}_n^2 = (y_n + 2r_n)/y_n^2$. The nth weighting factor $\hat{\sigma}_n^2$ 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 signal to noise ratio (SNR) values.

Alternatively, the choice of $\hat{\sigma}_n^2 = 1$ results in the unweighted least-squares (ULS) approach, which leads to estimates with much higher variance.

4.2. Ordinary Poisson (OP) approximation

The conventional approach is to assume (approximate) that $\{Y_n\}_{n=1}^N$ are distributed as independent Poisson random variables with mean $\bar{y}_n(\mu) = b_n \mathrm{e}^{-l_n(\mu)}$ as in (4), i.e.

$$P(\underline{Y} = \underline{y}; \mu) \approx \prod_{n=1}^{N} \frac{[\bar{y}_n(\mu)]^{y_n}\ \mathrm{e}^{-\bar{y}_n(\mu)}}{y_n!}. \tag{11}$$

The log-likelihood corresponding to this OP approximation is

$$\begin{aligned} L_{\mathrm{OP}}(\mu) &= \sum_{n=1}^{N} y_n \log \bar{y}_n(\mu) - \bar{y}_n(\mu) \\ &= \sum_{n=1}^{N} y_n \log(b_n \mathrm{e}^{-l_n(\mu)}) - b_n \mathrm{e}^{-l_n(\mu)}, \end{aligned} \tag{12}$$

disregarding the constants independent of μ.

4.3. Shifted Poisson (SP) approximation

A better approach is to match both the first- and second-order moments by approximating the random variables $\{Y_n + 2r_n\}_{n=1}^N$ as having Poisson distributions with means $\{\bar{y}_n(\mu) + 2r_n\}$. This model leads to our proposed SP objective function:

$$L_{\mathrm{SP}}(\mu) = \sum_{n=1}^{N} (y_n + 2r_n) \log(\bar{y}_n(\mu) + 2r_n) - (\bar{y}_n(\mu) + 2r_n),$$

where $\bar{y}_n(\mu)$ is defined by Equation (4). Note that although both L_{WLS} and L_{SP} match two moments of the measurement distribution, in WLS the second moment of $\hat{l}_n(y_n)$ is 'fixed' to $\hat{\sigma}_n^2$ independently of μ, whereas in the SP model the moments vary with $\bar{y}_n(\mu)$ appropriately.

We have previously shown empirically that the SP model better agrees with the exact log-likelihood than either the WLS or OP model (Yavuz and Fessler, 1996). Next we provide an analytical result that corroborates those results.

4.4. Variance analysis

To analyze the variance of each estimator, we apply the analytic approximations suggested by Fessler (1996). If $\bar{Y} = \mathrm{E}\{\underline{Y}\}$, then using a first-order Taylor expansion of $\hat{\mu}(\underline{Y})$ leads to the following approximation for the covariance of $\hat{\mu}$ (Fessler, 1996):

$$\mathrm{Cov}\{\hat{\mu}\} \approx P\ \mathrm{Cov}\{\underline{Y}\}\ P^{\mathrm{T}} \tag{13}$$

where $P = [-\nabla^{20}\Phi(\bar{\mu}, \bar{Y})]^{-1}\nabla^{11}\Phi(\bar{\mu}, \bar{Y})$ and $\bar{\mu} = \arg\max_\mu \Phi(\mu, \bar{Y})$. The (j,k)th element of operator ∇^{20} is $\partial^2/\partial\mu_j\partial\mu_k$ and the (j,n)th element of operator ∇^{11} is $\partial^2/\partial\mu_j\partial Y_n$.

We apply Equation (13) to find approximate expressions for the variance of the maximum-likelihood estimators: $\hat{\mu}_{\mathrm{OP}} = \arg\max_\mu L_{\mathrm{OP}}(\mu)$ and $\hat{\mu}_{\mathrm{SP}} = \arg\max_\mu L_{\mathrm{SP}}(\mu)$. For this purpose we considered a highly simplified version of transmission tomography where the unknown is a scalar parameter, i.e. $M = 1$. This highly simplified problem provides insight into the estimator bias and variance without the undue notation of the multi-parameter case. The objective functions described above can be expressed in the form:

$$\Phi(\mu, Y) = \sum_{n=1}^{N} h_n(\mu, Y).$$

Since the measurements are statistically independent, for the scalar problem the above approximation Equation (13) reduces to

$$\mathrm{Var}\{\hat{\mu}\} \approx \left(\sum_{n=1}^{N} \frac{\partial^2 h_n(\bar{\mu}, \bar{Y})}{\partial\mu^2}\right)^{-2} \times \sum_{n=1}^{N}\left[\frac{\partial^2 h_n(\bar{\mu}, \bar{Y})}{\partial\mu\ \partial Y_n}\right]^2 \mathrm{Var}\{Y_n\}. \quad (14)$$

With some tedious algebra, one can derive the following approximate expressions for the variance of $\hat{\mu}_{\mathrm{OP}}$ and $\hat{\mu}_{\mathrm{SP}}$:

$$\mathrm{Var}\{\hat{\mu}_{\mathrm{OP}}\} \approx \frac{\sum_{n=1}^{N} a_n^2(\bar{y}_n(\mu_t) + 2r_n)}{\left(\sum_{n=1}^{N} a_n^2\bar{y}_n(\mu_t)\right)^2} \quad (15)$$

$$\mathrm{Var}\{\hat{\mu}_{\mathrm{SP}}\} \approx \left[\sum_{n=1}^{N} \frac{a_n^2\bar{y}_n(\mu_t)^2}{\bar{y}_n(\mu_t) + 2r_n}\right]^{-1}, \quad (16)$$

where μ_t denotes the true attenuation coefficient value and $\bar{y}_n(\mu) = b_n \mathrm{e}^{-a_n\mu}$.

Letting $s_n = a_n^2\bar{y}_n(\mu_t)$ and $t_n = a_n^2(\bar{y}_n(\mu_t) + 2r_n)$, one can rewrite Equations (15) and (16) as

$$\frac{1}{\mathrm{Var}\{\hat{\mu}_{\mathrm{OP}}\}} \approx \frac{\left(\sum_n s_n\right)^2}{\sum_n t_n}, \qquad \frac{1}{\mathrm{Var}\{\hat{\mu}_{\mathrm{SP}}\}} \approx \sum_n \frac{s_n^2}{t_n}.$$

Let $\underline{a},\ \underline{b} \in \mathbb{R}^n$ such that $a_n = s_n/\sqrt{t_n},\ b_n = \sqrt{t_n}$. Using the Cauchy–Schwarz inequality: $|\underline{a}^{\mathrm{T}}\underline{b}| \le \|\underline{a}\|_2\,\|\underline{b}\|_2$,

$$\sum_n s_n \le \left(\sum_n \frac{s_n^2}{t_n}\right)^{1/2}\left(\sum_n t_n\right)^{1/2}$$

$$\left(\sum_n \frac{s_n^2}{t_n}\right)^{-1} \le \frac{\sum_n t_n}{\left(\sum_n s_n\right)^2},$$

so that within the accuracy of Equation (13)

$$\mathrm{Var}\{\hat{\mu}_{\mathrm{SP}}\} \le \mathrm{Var}\{\hat{\mu}_{\mathrm{OP}}\}, \quad (17)$$

with equality if, and only if, the $r_n/\bar{y}_n$ ratios are equal. For PET systems, these ratio terms are never constant, and in fact can be quite disparate. Thus we have shown the following result for the above scalar problem: the variance of the SP estimator is lower than the variance of the OP estimator. This analytic result is corroborated by multi-parameter empirical results in Section 6.

5. SADDLE-POINT (SD) APPROXIMATION

An alternative to the previous approximations for the exact PMF (6) of precorrected measurements is to make second-order Taylor series approximations in the z-transform domain. Snyder *et al.* (1995) (see also Helstrom, 1978) have applied the saddle-point approximation to the distribution of the sum of independent Gaussian and Poisson random variables. Here we apply the saddle-point method to the distribution of the difference of two independent Poisson random variables. We performed a quadratic approximation to the probability generating function and then carried out the inverse transform to find the PMF.

Let $U \sim \mathrm{Poisson}(\alpha)$, $V \sim \mathrm{Poisson}(\beta)$ and $Y = U - V$ with PMFs $P_U(k)$, $P_V(k)$ and $P_Y(k)$ respectively. When U and V are independent the generating function of Y is

$$G_Y(z) = \sum_k z^k P_Y(k) = G_U(z)\,G_V(z^{-1})$$

where $G_U(z) = \exp(\alpha(z-1))$ and $G_V(z) = \exp(\beta(z-1))$. In terms of the generating function, $P_Y(k)$ is given by the contour integral

$$P_Y(k) = \frac{1}{2\pi\mathrm{j}}\oint_{C^+} z^{-k-1}G_Y(z)\,\mathrm{d}z = \frac{1}{2\pi\mathrm{j}}\oint_{C^+} \mathrm{e}^{\Phi_k(z)}\,\mathrm{d}z, \quad (18)$$

where $\mathrm{j} = \sqrt{-1}$ and the contour C^+ must lie in the region of convergence of $G_Y(z)$ and enclose the origin, and

$$\begin{aligned}
\Phi_k(z) &= -(k+1)\log(z) + \alpha(z-1) + \beta(z^{-1}-1)\\
\frac{\mathrm{d}\Phi_k(z)}{\mathrm{d}z} &= \Phi_k^{(1)}(z) = -\frac{(k+1)}{z} + \alpha - \frac{\beta}{z^2}\\
\frac{\mathrm{d}^2\Phi_k(z)}{\mathrm{d}z^2} &= \Phi_k^{(2)}(z) = \frac{(k+1)}{z^2} + \frac{2\beta}{z^3}.
\end{aligned}$$

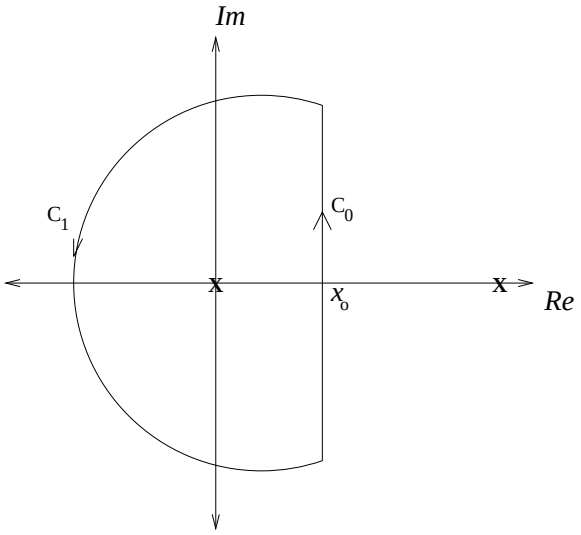

Figure 1. Deformation of the contour C^+ in the complex plane into a vertical line C_0 through the saddle point x_0 and a semicircle C_1 around the left half-plane at infinity. The singularities of the integrand are at $z = 0$ and $z = \infty + \mathrm{j}0$ for $k \geq 0$.

We observe that $\Phi_k(z)$ (and hence the integrand $\mathrm{e}^{\Phi_k(z)}$) is convex for $z \in \mathbb{R}$, $z > 0$ and $k \geq 0$. The integrand has a minimum at $x_0 \in \mathbb{R}$, $x_0 > 0$ which is called the saddle point, i.e.

$$\Phi_k^{(1)}(x_0) = -\frac{(k+1)}{x_0} + \alpha - \frac{\beta}{x_0^2} = 0 \qquad \text{and} \qquad x_0 > 0$$

which yields

$$x_0 = \frac{(k+1) + v_k}{2\alpha} = \frac{2\beta}{-(k+1) + v_k}, \tag{19}$$

where $v_k = x_0^2 \Phi_k^{(2)}(x_0) = \sqrt{(|k|+1)^2 + 4\alpha\beta}$.

Following Helstrom (1978), we deform the contour C^+ in Equation (18) into a vertical line C_0 through the saddle point x_0, as $z = x_0 + \mathrm{j}y$, $-\infty < y < \infty$ and a semicircle C_1 around the left half-plane at infinity (Figure 1). This contour is permissible for $k \geq 0$, since the only singularities of the integrand are at $z = 0$ and $z = \infty + \mathrm{j}0$. If $|z| \to \infty$ for $\mathrm{Re}[z] < x_0$ then $\mathrm{e}^{\Phi_k(z)} \to 0$. Hence the contribution of the semicircle around the left half-plane at infinity vanishes and Equation (18) reduces to

$$P_Y(k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{\Phi_k(x_0 + \mathrm{j}y)}\, \mathrm{d}y. \tag{20}$$

Expanding $\Phi_k(z)$ in Taylor's series around $z = x_0$, one obtains:

$$\begin{aligned} \exp[\Phi_k(z)] &= \exp\Big[\Phi_k(x_0) + \tfrac{1}{2}\Phi_k^{(2)}(x_0)(z-x_0)^2 \\ &\qquad + \sum_{l=3}^{\infty} \frac{1}{l!}\Phi_k^{(l)}(x_0)(z-x_0)^l\Big] \\ &= \exp\big[\Phi_k(x_0) + \tfrac{1}{2}\Phi_k^{(2)}(x_0)(z-x_0)^2\big] \\ &\qquad \times\big[1 + \tfrac{1}{6}\Phi_k^{(3)}(x_0)(z-x_0)^3 + \cdots\big], \end{aligned}$$

since $\Phi_k^{(1)}(x_0) = 0$. The integral (20) becomes

$$\begin{aligned} P_Y(k) &= \frac{\mathrm{e}^{\Phi_k(x_0)}}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{\frac{1}{2}\Phi_k^{(2)}(x_0)(\mathrm{j}y)^2} \\ &\qquad \times\big[1 + \tfrac{1}{6}\Phi_k^{(3)}(x_0)(\mathrm{j}y)^3 + \cdots\big]\,\mathrm{d}y \\ &= \frac{\mathrm{e}^{\Phi_k(x_0)}}{\sqrt{2\pi\,\Phi_k^{(2)}(x_0)}}\,[1+R] \\ &= \frac{x_0^{-k}\mathrm{e}^{v_k - \alpha - \beta}}{\sqrt{2\pi\, v_k}}\,[1+R] \end{aligned} \tag{21}$$

where

$$R = \frac{\Phi_k^{(4)}(x_0)}{8\left[\Phi_k^{(2)}(x_0)\right]^2} + \cdots.$$

Using the algorithm by Rice (1968), the residuum R can be written as

$$\begin{aligned} R &= \frac{1}{24(k+1)}\Big[\frac{-5 + 12\sqrt{1+\eta} - 9(1+\eta)}{(1+\eta)^{3/2}}\Big] \\ &\qquad + \mathrm{O}\left[\left(\frac{1}{k+1}\right)^2\right] \end{aligned}$$

where $\eta = 4\alpha\beta/(k+1)^2$. The residuum goes asymptotically to zero as $k \to \infty$ and more importantly we have observed empirically that the approximation error is negligibly small even for very small values of k. Neglecting R in Equation (21) results in our saddle-point approximation for the PMF $P_Y(k)$ as

$$P_Y(k) \simeq P_Y^s(k) = \frac{x_0^{-k}\mathrm{e}^{v_k - \alpha - \beta}}{\sqrt{2\pi\, v_k}}, \qquad k \geq 0. \tag{22}$$

For $k < 0$ the integrand in Equation (18) is not guaranteed to be convex for $z > 0$. Moreover, the integrand does not vanish along the semicircle around the left half-plane at infinity. Thus we use the change of variables $w = 1/z$ in

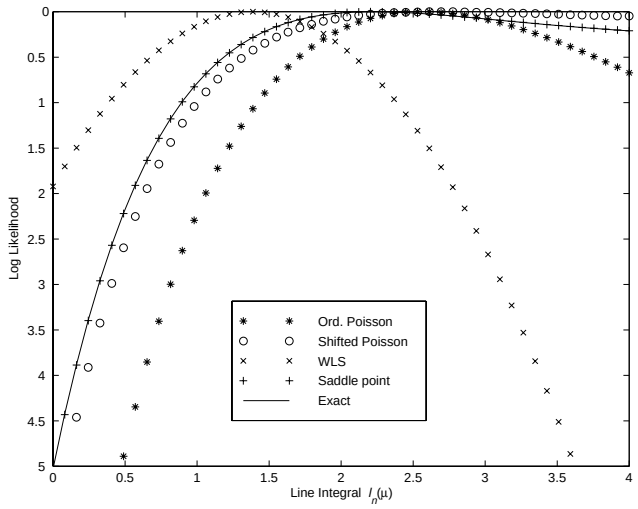

Figure 2. Representative comparison of the exact log-likelihood function with objective functions of different models as a function of line integral $l_n(\mu)$. The randoms rate is 5%. The proposed saddle-point approximation agrees with the exact log-likelihood method significantly better than the other models.

Equation (18), so that

$$\begin{aligned} P_Y(k) &= \frac{1}{2\pi \mathrm{j}} \oint_{C^+} w^{k-1} G_Y(w^{-1}) \, \mathrm{d}w \\ &= \frac{1}{2\pi \mathrm{j}} \oint_{C^+} \mathrm{e}^{\bar{\Phi}_k(w)} \, \mathrm{d}w \end{aligned} \tag{23}$$

where

$$\bar{\Phi}_k(w) = (k-1)\log(w) + \alpha(w^{-1} - 1) + \beta(w - 1).$$

Following similar steps as for the case where $k \geq 0$, the saddle-point approximation for $k < 0$ can be shown to be

$$P_Y(k) \simeq P_Y^s(k) = \frac{w_0^k \mathrm{e}^{v_k - \alpha - \beta}}{\sqrt{2\pi v_k}}, \qquad k < 0 \tag{24}$$

where

$$w_0 = \frac{-(k-1) + v_k}{2\beta} = \frac{2\alpha}{(k-1) + v_k}.$$

Thus, combining Equations (22) and (24) the saddle-point (SD) approximation for the log-likelihood (7) is

$$\begin{aligned} L_{\mathrm{SD}}(\mu) &= \sum_{n=1}^{N} \log P_Y^s(y_n; \bar{y}_n(\mu)) \\ &= \sum_{n=1}^{N} h_n^s(\mu) \end{aligned} \tag{25}$$

where

$$h_n^s(\mu) = \begin{cases} y_n \log\left(\dfrac{\bar{y}_n(\mu) + r_n}{y_n + 1 + u_n(\mu)}\right) - t_n(\mu), & y_n \geq 0 \\ y_n \log\left(\dfrac{\bar{y}_n(\mu) + r_n}{y_n - 1 + u_n(\mu)}\right) - t_n(\mu), & y_n < 0 \end{cases} \tag{26}$$

with

$$\begin{aligned} t_n(\mu) &= \bar{y}_n(\mu) - u_n(\mu) + \tfrac{1}{2}\log u_n(\mu), \\ u_n(\mu) &= \sqrt{(|y_n| + 1)^2 + 4(\bar{y}_n(\mu) + r_n) r_n}, \end{aligned}$$

and disregarding constants independent of μ.

Note that the approximation Equation (25) is considerably simpler than the exact log-likelihood Equation (7), since no infinite sums or factorials are needed. Nevertheless, it is remarkably accurate as shown below. Also, one can observe that as $r_n \to 0$, $h_n^s(\mu) \to [y_n \log \bar{y}_n(\mu) - \bar{y}_n(\mu)] = L_{\mathrm{OP}}(\mu)$ (to within constants independent of μ), which is expected, because for $r_n = 0$ the ordinary Poisson model is appropriate.

Figure 2 shows a representative comparison of the exact log-likelihood function and the approximations for noiseless data as a function of μ. $L_{\mathrm{WLS}}(\mu)$ is particularly poor, in part because of the condition $y_n > 0$ in Equation (10). Although $L_{\mathrm{SP}}(\mu)$ fits the exact log-likelihood better than $L_{\mathrm{WLS}}(\mu)$ and $L_{\mathrm{OP}}(\mu)$, clearly $L_{\mathrm{SD}}(\mu)$ has the best agreement with the exact log-likelihood $L(\mu)$. In a large number of additional comparisons not shown due to space considerations, we have observed that $L_{\mathrm{SD}}(\mu)$ agrees remarkably well with the exact log-likelihood $L(\mu)$ and clearly better than the other models.

6. TWO-DIMENSIONAL SIMULATIONS

To study bias and variance properties of the estimators based on the above approximations, we performed 2-D simulations. For μ we used the synthetic attenuation map shown in Figure 3, which represents a human abdomen with a linear attenuation coefficient of 0.0096 mm^{-1}. The image was a 128×128 array of 4.7 mm pixels. We simulated a PET transmission scan with 192 radial bins and 256 angles spaced uniformly over 180°. The a_{nj} factors correspond to 3.1 mm wide strip integrals on 3.1 mm center-to-center spacing. The b_n factors were generated using pseudo-random log-normal variates with a standard deviation of 0.3 to simulate detector pairs with non-uniform detector efficiencies and scaled so that $\sum_n \bar{y}_n$ was 3.6×10^6 counts. The r_n factors corresponded to a uniform field of 10% random coincidences. Pseudo-random transmission measurements were generated according to Equations (2) and (3). For regularization, we used the modified quadratic penalty (Fessler and Rogers, 1996).

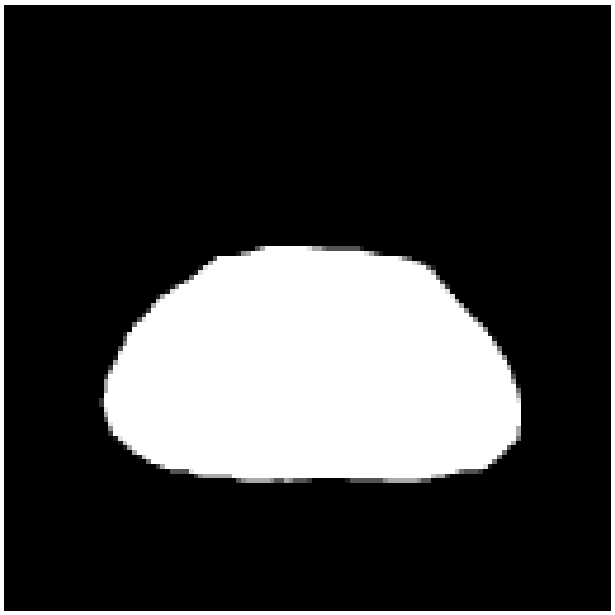

Figure 3. Simulated abdomen attenuation map.

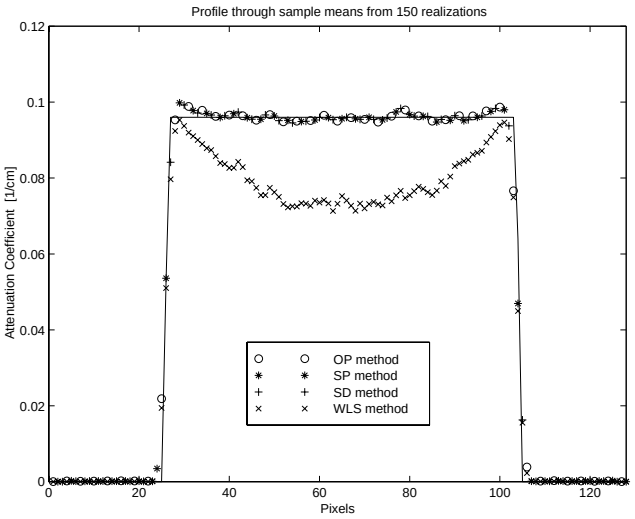

Figure 4. Horizontal profile through the sample mean images for abdomen phantom. The WLS method has a systematic negative bias. The OP, SP and SD methods are free of this systematic negative bias.

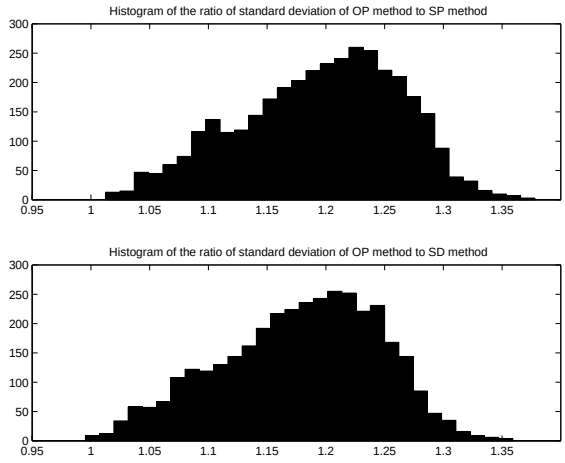

Figure 5. Histogram of the ratio of standard deviations in reconstructions of the abdomen phantom. The OP method yields, on average, 19% higher standard deviation than the proposed SP and SD methods.

We generated 150 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 algorithms (Fessler *et al.*, 1997) applied to the objective functions (10), (12), (13) and (25). Although $L_{\mathrm{OP}}(\mu)$ is globally convex, $L_{\mathrm{SP}}(\mu)$ is only locally convex (Fessler, 1995). In our simulations, we initialized the iterations with a FBP image and always observed monotonic increase in the log-likelihood for all methods. Nevertheless, further investigation of global convergence properties is necessary for $L_{\mathrm{SP}}(\mu)$ and $L_{\mathrm{SD}}(\mu)$.

We computed both the sample mean and sample standard deviation images for all methods. Figure 4 shows horizontal profiles through the sample mean images. These profiles show that WLS is systematically negatively biased (Fessler, 1995), whereas the OP, SP and SD models are free of systematic bias.

To study the variance, we computed the ratio of sample standard deviation images of different estimators, over all interior pixels. Figure 5 shows the histogram of the standard deviation ratios. The OP model yields, on the average, 19% higher standard deviation than the both SP and SD models. In other words, to achieve the same noise level, the OP method would require about 40% greater scan time.

Although the standard deviation values could be decreased by using higher count rates, the ratio of standard deviations of different estimators will remain approximately the same for higher count rates (Fessler, 1996). This follows from the fact that analytic approximations (15) and (16) will be more accurate with increasing count rates, and these approximations show that for a set of fixed system parameters, the ratio of the standard deviation of different estimators remains constant independent of the count rate.

We performed additional simulations using a digital thorax phantom with non-uniform attenuation (Yavuz and Fessler,

Table 1. Local impulse response and the local sample standard deviation for the central pixel

Estimator	FWHM (pixels) Horizontal	Vertical	Average	% Std dev.
FBP	2.66	2.68	2.67	18.20 ± 1.05
OP	2.13	3.22	2.67	9.94 ± 0.57
SP	1.94	3.40	2.67	7.70 ± 0.44
SD	1.93	3.41	2.67	7.94 ± 0.45

1996). The reductions in noise with the proposed methods were comparable.

It is well known in tomographic image reconstruction that one can compromise between the resolution and noise in the reconstructed image. In the simulations reported here, we have used the modified quadratic penalty (Fessler and Rogers, 1996), which matches the spatial resolution of both least-squares-based and Poisson-based estimators. In order to show that the noise reduction with the proposed SP and SD methods does not come with the price of lower resolution (compared to the OP method), we have investigated the local resolution and the standard deviation of a pixel at the center of the abdomen phantom.

We computed the linearized local impulse response (Fessler and Rogers, 1996) of different estimators at the central pixel of the abdomen phantom. Table 1 shows the full width at half maximum (FWHM) values of local impulse response functions and the local sample standard deviation for the central pixel estimates. The table also reports the standard errors for the sample standard deviation estimates. These results show that the reductions in the standard deviations are truly due to the improved statistical modeling rather than resolution differences.

Although the local impulse response functions are asymmetric with respect to the horizontal and vertical axis, the 'average' resolution of each method is matched. As expected the non-statistical FBP method yields a much higher standard deviation than statistical methods. The standard deviations of the proposed SP and SD estimators are about 27% lower than the OP method.

The asymmetry of the local impulse responses is caused partly by the eccentricity of the abdomen phantom (Figure 3), (Fessler and Rogers, 1996) and we plan to include a new penalty with a symmetric impulse response in our future work. In Table 1 the resolution of SP and SD models are observed to be more asymmetric than the OP model. In order to investigate this effect we performed additional simulations using a circularly symmetric disk phantom which yields a symmetric impulse response at the center. For the central pixel (where all methods have the same impulse response) the reductions in standard deviation with the proposed SP and SD methods were around 24% compared to the OP method.

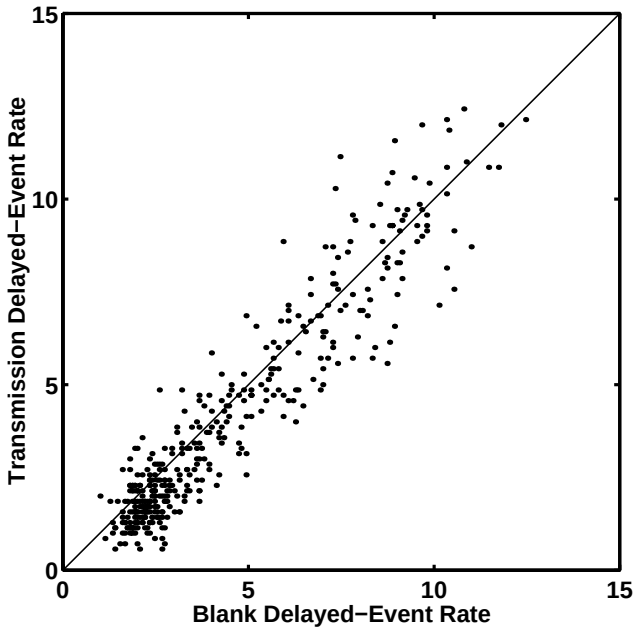

Figure 6. Scatter plot of the delayed coincidence event of blank and transmission scans.

7. ESTIMATES OF THE AC RATES

One must know the mean of the AC events (r_n) to compute $L_{\mathrm{SP}}(\mu)$ and $L_{\mathrm{SD}}(\mu)$. Since the r_n terms are not readily available from the real (precorrected) data, some estimates of the randoms must be used. Fortunately, for PET transmission scans we can obtain reasonable estimates from a blank scan.

Figure 6 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 acquire the delayed coincidence events during the blank scan and use them (after properly normalizing for different scan durations) as an estimate of the AC rates for transmission scans performed on the same PET system. The above technique is only applicable to transmission scans. For emission scans some other method will be necessary. However as described below the suggested methods are robust to errors in the estimates of AC events. We performed additional simulations (not shown) in which we substituted a simple constant for r_n rather

than the true values into the SP and SD objective functions. This approximation resulted in only a slight increase in the standard deviation (around 2%) of the SP and SD estimates without any systematic bias. These results demonstrate that both the SP and SD approximations are robust to errors in the r_n estimates.

8. DISCUSSION

AC events are a primary source of background noise in positron emission tomography. After the AC events are precorrected, the measurement statistics are no longer Poisson. For transmission scans, the WLS method and the ML method based on the OP model lead to a systematic bias and a higher variance, respectively, compared to our proposed SP model for measurement statistics which matches both the first- and second-order moments.

We proposed a new approximation for the exact log-likelihood which is derived using a saddle-point approximation to the PMF of precorrected measurements. Both the analysis of the error term and the log-likelihood plots and one-dimensional (1-D) simulations (not shown due to space considerations) show that the new approximation agrees very closely with the exact log-likelihood compared to previous approximations.

Our 2-D simulations yielded comparable performance for the SP and SD models. They were both free of systematic bias and yielded a reduced standard deviation ($\sim$19%) compared to the OP model. We further analyzed the resolution and sample standard deviation at a central pixel of the phantom and showed that with our proposed methods the standard deviation is still reduced when the resolution of all estimators are matched. We observed very close agreement between the exact log-likelihood and the SD approximation both from the log-likelihood plots and 1-D simulations. Therefore we were expecting the SD method to perform better than the SP method. However, for the 2-D simulations reported here, the SP method performed as well as the SD method. Thus the SP method is particularly attractive since it requires comparable computation to the OP method but has reduced variance. We plan to compare the SD and SP methods to the uniform Cramer–Rao bounds (Hero *et al.*, 1996).

The high correlation between delayed coincidence events of blank and transmission scans suggest that one can use AC rates estimated from blank scans. We have seen that even using constant AC rates in 2-D simulations resulted in only a slight increase in the standard deviation without any systematic bias. Thus the proposed SP and SD methods are robust enough for practical use.

We plan to apply the proposed method to emission tomography, where even higher AC rates than the transmission tomography are common, particularly in 3-D PET. Moreover, in 3-D PET, very large data sets are likely to preclude separate acquisition of random coincidences, so the real-time subtraction methods are usually used for emission scans. So the potential benefit of the proposed models should be even greater.

ACKNOWLEDGEMENTS

This work was supported in part by NIH grants CA-60711 and CA-54362. We gratefully acknowledge Edward Ficaro for assistance with the phantom measurements shown in Figure 6.

REFERENCES

Casey, M. E. and Hoffman, E. J. (1986) Quantitation in positron emission computed tomography: 7. A technique to reduce noise in accidental coincidence measurements and coincidence efficiency calibration. *J. Comput. Assis. Tomogr.*, 10, 845–850.

Fessler, J. A. (1995) Hybrid Poisson/polynomial objective functions for tomographic image reconstruction from transmission scans. *IEEE Trans. Image Proc.*, 4, 1439–1450.

Fessler, J. A. (1996) Mean and variance of implicitly defined biased estimators (such as penalized maximum likelihood): applications to tomography. *IEEE Trans. Image Proc.*, 5, 493–506.

Fessler, J. A. and Rogers, W. L. (1996) Spatial resolution properties of penalized-likelihood image reconstruction methods: space-invariant tomographs. *IEEE Trans. Image Proc.*, 5, 1346–1358.

Fessler, J. A., Ficaro, E. P., Clinthorne, N. H. and Lange, K. (1997) Grouped-coordinate ascent algorithms for penalized-likelihood transmission image reconstruction. *IEEE Trans. Med. Imag.*, 16, 166–175.

Helstrom, C. (1978) Approximate evaluation of detection probabilities in radar and optical communications. *IEEE Trans. Aero. Electron. Syst.*, 14, 630–640.

Hero, A. O., Fessler, J. A. and Usman, M. (1996) Exploring estimator bias–variance tradeoffs using the uniform CR bound. *IEEE Trans. Signal Proc.*, 44, 2026–2041.

Hoffman, E. J., Huang, S. C., Phelps, M. E. and Kuhl, D. E. (1981) Quantitation in positron emission computed tomography: 4. Effect of accidental coincidences. *J. Comput. Assis. Tomogr.*, 5, 391–400.

Mumcuoğlu, E. Ü., Leahy, R. M. and Cherry, S. R. (1996) Bayesian reconstruction of PET images: methodology and performance analysis. *Phys. Med. Biol.*, 41, 1777–1807.

Ollinger, M. and Fessler, J. A. (1997) Positron emission tomography. *IEEE Signal Proc. Mag.*, 14, 43–55.

Politte, D. G and Snyder, D. L. (1991) Corrections for accidental coincidences and attenuation in maximum-likelihood image reconstruction for positron-emission tomography. *IEEE Trans. Med. Imag.*, 10, 82–89.

Rice, S. O. (1968) Uniform asymptotic expansions for saddle point integrals—application to a probability distribution occurring in noise theory. *Bell Syst. Tech. J.*, 47, 1971–2013.

Sauer, K. and Bouman, C. (1993) A local update strategy for iterative reconstruction from projections. *IEEE Trans. Signal Proc.*, 41, 534–548.

Snyder, D. L., Helstrom, C. W., Lanterman, A. D., Faisal, M. and White, L. (1995) Compensation for readout noise in CCD images. *J. Opt. Soc. Am.* A, 12, 272–283.

Yavuz, M. and Fessler, J. A. (1996) Objective functions for tomographic reconstruction from randoms-precorrected PET scans. *Proc. IEEE Nuclear Science Symp. on Medical Imaging*, vol. 2, pp. 1067–1071.