Penalized-Likelihood Estimators and Noise Analysis for Randoms-Precorrected PET Transmission Scans

Mehmet Yavuz,* *Member, IEEE*, and Jeffrey A. Fessler, *Member, IEEE*

***Abstract*—This paper analyzes and compares image reconstruction methods based on practical approximations to the exact log likelihood of randoms-precorrected positron emission tomography (PET) measurements. The methods apply to both emission and transmission tomography, however, in this paper we focus on transmission tomography. The results of experimental PET transmission scans and variance approximations demonstrate that the shifted Poisson (SP) method avoids the systematic bias of the conventional data-weighted least squares (WLS) method and leads to significantly lower variance than conventional statistical methods based on the log likelihood of the ordinary Poisson (OP) model. We develop covariance approximations to analyze the propagation of noise from attenuation maps into emission images via the attenuation correction factors (ACF's). Empirical pixel and region variances from real transmission data agree closely with the analytical predictions. Both the approximations and the empirical results show that the performance differences between the OP model and SP model are even larger, when considering noise propagation from the transmission images into the final emission images, than the differences in the attenuation maps themselves.**

***Index Terms*—Covariance approximations, penalized maximum likelihood image reconstruction, randoms-precorrected PET, statistical approximations.**

I. INTRODUCTION

To obtain accurate images of the radioactivity distribution within a patient, using emission computed tomography, one must correct for the effects of attenuation [1] and accidental coincidences [2]. The measured attenuation correction method is routinely performed in positron emission tomography (PET) centers, where transmission scans are used to measure the unique attenuation characteristic of each patient over the slices of interest [3]. Since PET transmission scans are performed, essentially, to compute attenuation correction factors, rather than being the primary medical interest, it is desirable to minimize their durations. Short scans suffer from a limited number of counts, especially for the projections passing through high attenuation regions of the patient, which results in propagation of unwanted errors and artifacts into emission images [4], [5]. Smoothing of the transmission data before computing the attenuation correction factors leads to resolution mismatch between transmission and emission data [6], [7]. Reconstructing images of attenuation distributions (attenuation maps) from noisy transmission scans has desirable properties, such as better noise performance [4], [5] and anatomical localization [8]. These attenuation maps are usually reprojected to form attenuation correction factors. Using the conventional filtered back projection (FBP) method for reconstruction of attenuation maps results in biased estimates when the transmission counts are small [9]. Penalized-likelihood methods, which incorporate the measurement statistics appropriately, avoid this bias problem.

In PET, accidental coincidence (AC) events occur when two photons that originate in separate annihilations are detected within the coincidence timing window [2], [10]. In transmission scans, photons that originate from different transmission sources (rod or ring sources) cause most AC events. The effect of AC events is most severe for rays with low true-coincidence rates [11], such as those traversing the thorax. Most PET scans are compensated for AC events by real-time subtraction of delayed-window coincidences. This precorrection yields the proper ensemble mean, but destroys the Poisson measurement statistics [10]. To preserve Poisson measurements, one should acquire separate transmission and randoms sinograms [5], [12]. However, because of hardware and data storage space limitations (and historical momentum), most PET centers currently collect and archive only the randoms-precorrected data.

The exact log likelihood for randoms-precorrected data contains infinite summations, so we have previously developed a few practical approximations [13]–[15]. The three simplest of these approximations are the focus of this paper: the data-weighted least squares (WLS) method, the ordinary Poisson (OP) model (which ignores the AC corrections), and the shifted Poisson (SP) model, which matches both the first- and second-order moments of the model to the underlying statistics of the precorrected data [13]–[15]. Previous two-dimensional (2-D) computer simulation studies of penalized-likelihood estimators, based on these approximations, showed that the WLS method leads to a large systematic negative bias and the SP method yields attenuation maps with lower reconstructed image variance than the OP method at matched spatial resolutions [15].

In this paper, we compare the WLS, OP, and SP methods, using 100 experimental PET transmission scans of an anthropomorphic thorax phantom acquired with a Siemens/CTI 931 PET scanner. The empirical results from this study are consistent with the previous simulation results: a large bias for

Manuscript received July 17, 1998; revised August 3, 1999. This work was supported in part by the NIH under Grants CA-60711 and CA-54362. The work of M. Yavuz was supported by a TUBITAK-NATO Science Fellowship for doctoral studies. The Associate Editor responsible for coordinating the review of this paper and recommending its publication was D. Townsend. *Asterisk indicates corresponding author.*

*M. Yavuz is with the GE Research and Development Center, KWB 611, Niskayuna, NY 12309 USA.

J. A. Fessler is with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109 USA

Publisher Item Identifier S 0278-0062(99)08512-2.

the WLS method, and lower variance for the SP method. We also implemented and evaluated an extremely precise approximation to the exact statistical model, based on truncating the infinite summations in the log likelihood. Although the individual images reconstructed by the SP method and the truncated exact log-likelihood method differed slightly, the differences in the ensemble means and variances were insignificant. We also investigated the reconstruction results for the saddle-point (SD) approximation that we introduced previously [14], [15]. We observed very close agreement between the SD method and truncated exact log likelihood for each reconstruction in our simulations. However, for the transmission scans reported here the difference between the SD and the SP method were not statistically significant, so we did not include the SD reconstruction results in this paper. Thus, the simple and practical SP approximation appears adequate for routine use on randoms-precorrected PET transmission data.

In addition to evaluating the attenuation maps themselves, we also investigated the propagation of noise from the reconstructed attenuation maps into emission images reconstructed using the FBP method. Again, the SP method leads to lower variance than the OP method. Interestingly, the difference in variances in the emission images was even greater than in the attenuation maps.

To corroborate the empirical studies described above, we have also developed analytical approximations to the reconstructed image covariance based on the techniques developed in [16]. The covariance approximations for the OP and SP methods, which had previously been evaluated only with simulated data in [16], were found to agree well with the empirical variance computed from the experimental PET transmission scans.

In this paper we also develop analytical approximations for the propagation of noise from attenuation maps into reconstructed emission images. To isolate the effect of transmission noise in the resultant emission image, we consider noise-free emission measurements and develop approximations for the covariance of emission images reconstructed using ACF's computed from noisy attenuation maps. These approximations describe the propagation of noise from attenuation maps into emission reconstruction. We also show that the predicted variances agree with the empirical results from the experimental PET transmission scans.

Section II reviews the measurement model and exact log likelihood, and Section III describes the log-likelihood approximations. Section IV describes the empirical results for attenuation map reconstruction. Sections V and VI describe the covariance approximations and examine the propagation of noise into the emission images.

II. MEASUREMENT MODEL AND EXACT LOG LIKELIHOOD

In conventional PET scans, the system detects coincidence events during two time windows, the prompt window and the delayed window, and the data are precorrected for AC events by real-time subtraction of delayed window coincidences [10]. Each such precorrected measurement is the difference of two independent Poisson random variables, which compensates in mean for AC events, but which also increases the measurement variance.

Let $Y = [Y_1, \cdots, Y_N]'$ denote the vector of precorrected measurements, where $'$ denotes the vector and matrix transpose. 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 $\mu = [\mu_1, \cdots, \mu_M]'$ 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 $\bar{y}^{\text{p}}$ and $\bar{y}^{\text{d}}$, respectively, as

$$\begin{aligned} \mathrm{E}\{Y_n^{\text{prompt}}\} &= \bar{y}^{\text{p}}(\mu) = b_n e^{-l_n(\mu)} + r_n && (2)\\ \mathrm{E}\{Y_n^{\text{delay}}\} &= \bar{y}^{\text{d}} = r_n && (3) \end{aligned}$$

where $l_n(\mu) = \Sigma_{j=1}^{M} a_{nj}\mu_j$ is the total attenuation between 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

$$\begin{aligned} \mathrm{E}\{Y_n\} &= \bar{y}^{\text{p}}(\mu) - \bar{y}^{\text{d}} = b_n e^{-l_n(\mu)} \triangleq \bar{y}(\mu) && (4)\\ \mathrm{Var}\{Y_n\} &= \bar{y}^{\text{p}}(\mu) + \bar{y}^{\text{d}} = b_n e^{-l_n(\mu)} + 2r_n. && (5) \end{aligned}$$

Let $y = [y_1, \cdots, y_N]'$ be an observed realization of Y in (1). Since the measurements are independent, one can express the exact log likelihood as follows [15]:

$$L(\mu) = \sum_{n=1}^{N} h_n(l_n(\mu), y_n) \tag{6}$$

where, ignoring constants independent of μ throughout

$$\begin{aligned} h_n(l_n(\mu), y_n) \triangleq\ & \log\left(\sum_{m=\lfloor -y_n \rfloor_+}^{\infty} \frac{[\bar{y}^{\text{p}}(\mu)]^{y_n+m}}{(y_n+m)!}\frac{r_n^m}{m!}\right)\\ & - (\bar{y}^{\text{p}}(\mu) + r_n) \end{aligned} \tag{7}$$

where $\lfloor x \rfloor_+ = x$ if $x > 0$ and is zero otherwise.

Since image reconstruction is ill conditioned, we combine a roughness penalty $R(\mu)$ with the log likelihood to form 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 nonnegative cone

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

Since the exact log-likelihood function (7) is complicated, we describe approximations to the exact log likelihood.

III. APPROXIMATIONS TO THE EXACT LOG LIKELIHOOD

In this section, we review three practical approximations to $L(\mu)$: the WLS model, the conventional OP model, and the SP model approximation that we introduced previously [13]–[15]. All three log-likelihood approximations have the form (6) for different choices for $h_n(l, y_n)$.

A. Quadratic Approximations

A quadratic approximation to the exact log-likelihood function [9], [17] leads to the WLS objective function $L^{\rm WLS}(\mu)$ of the form (6) with

$$h_n^{\rm WLS}(l, y_n) = \begin{cases} -\frac{1}{2}(l - \hat{l}_n)^2 \frac{1}{\hat{\sigma}_n^2}, & y_n > 0 \\ 0, & y_n \le 0 \end{cases} \tag{10}$$

where $\hat{l}_n = \log(b_n/y_n)$ is the method-of-moments estimate of $l_n(\mu)$. The weighting factor $\hat{\sigma}_n^2 = \left((y_n + 2r_n)/y_n^2\right)$ is an estimate of the variance of $\hat{l}_n(y_n)$ based on a second-order Taylor expansion of $\hat{l}(\cdot)$ around $\overline{y}$. The residuals corresponding to projections with large values of y_n are weighted more heavily in (10). These rays pass through less dense objects and, consequently, have higher SNR values. Alternatively, the choice of $\hat{\sigma}_n^2 = 1$ would result in the unweighted least-squares (ULS) approach, which would lead to estimates with much higher variance.

B. 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 $\overline{y}(\mu)$ given by (6). The log likelihood $L^{\rm OP}(\mu)$ corresponding to this OP approximation is of the form (6) with

$$h_n^{\rm OP}(l, y_n) = y_n \log(b_n e^{-l}) - b_n e^{-l}. \tag{11}$$

C. 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 $\{\overline{y}(\mu) + 2r_n\}$. This idea leads to the SP approximation $L^{\rm SP}(\mu)$ [13]–[15] of the form (6) with

$$h_n^{\rm SP}(l, y_n) = (y_n + 2r_n)\log(b_n e^{-l} + 2r_n) - (b_n e^{-l} + 2r_n). \tag{12}$$

Although both $L^{\rm WLS}$ and $L^{\rm SP}$ 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 $\overline{y}(\mu)$ appropriately.

We have previously shown, both analytically and empirically in simulations, that the SP model better agrees with the exact log likelihood than either the WLS or OP model [13] and results in lower variance [14], [15]. Next, we present experimental results, using estimators based on the above approximations.

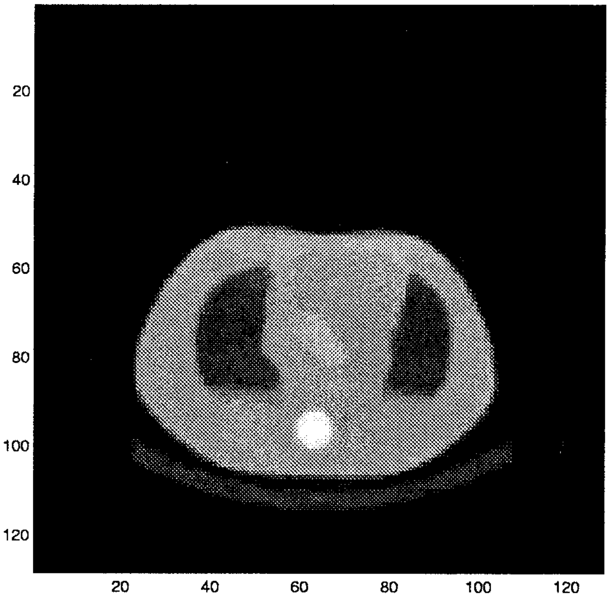

Fig. 1. Reconstruction of attenuation map for the slice of interest from 5-h transmission scan.

IV. EXPERIMENTAL RESULTS

We applied penalized-likelihood estimators, based on the above approximations, to reconstruct attenuation maps from transmission scans acquired with a Siemens/CTI 931 PET scanner. To study the bias and variance properties of these estimators, we collected 100 2-min transmission scans of an anthropomorphic thorax phantom (Data Spectrum, North Carolina). Fig. 1 shows the reconstructed attenuation map of the slice of interest from a 5-h transmission scan. In each 2-min scan, there were about 4.5M prompt coincidence events and 0.7M delayed events for the slice of interest. The sinograms had 192 radial bins and 256 angles uniformly sampled over 180°. We approximated the system geometry with 3.1-mm-wide strip integrals and 3.1-mm ray spacing. The reconstructed images were 128 by 128 with 4.7-mm pixels. For regularization, we used the modified quadratic penalty, as described by (30) and (35) in [18]. This penalty improves the resolution uniformity and enables matching of the spatial resolutions of different methods. We matched the resolution of the reconstructed transmission images for all methods to 2.65-pixels FWHM.

Previously, we have shown that a time-scaled version of delayed-coincidence events acquired during the blank scan is a good estimate for the r_n factors. (Even using a single scalar constant works fairly well [15].) Note that these estimates of the r_n factors are used, essentially, for estimating the variance of the randoms precorrected data in (5), not for performing randoms precorrection. In our experiments the r_n factors were unavailable for both the transmission and blank scans because the data was precorrected for randoms. Thus, to estimate the r_n factors for use in (12), we simply scaled the blank scan so that its sum corresponded to the total number of AC events (this scalar is available in the transmission scan file header) with no additional processing. Despite this possibly being

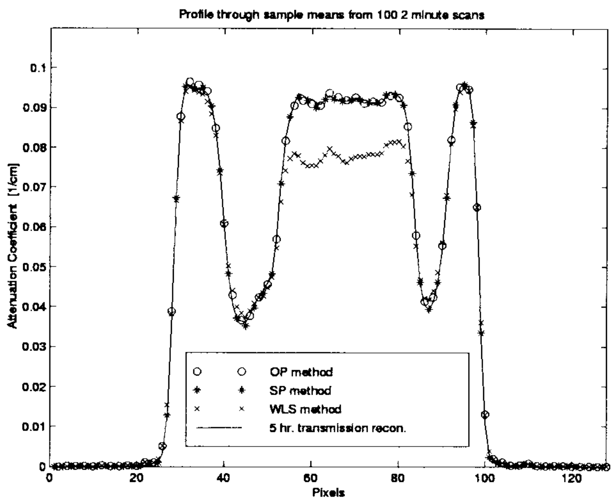

Fig. 2. Horizontal profile 66 through the sample mean images for an abdomen phantom. The WLS method has a systematic negative bias. The OP and SP methods appear free of this systematic negative bias.

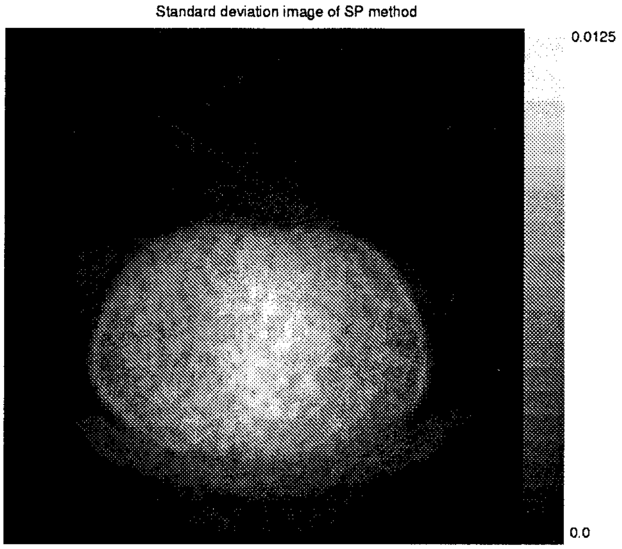

Fig. 3. Sample standard deviation image of the SP method from 100 transmission scans.

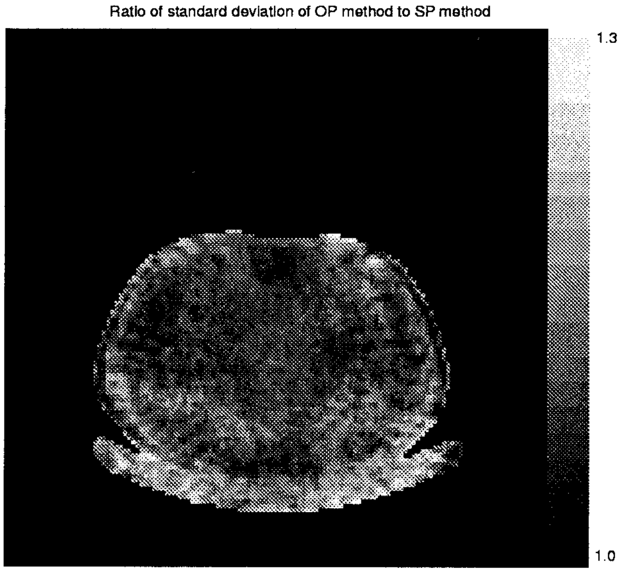

Fig. 4. Ratio of sample standard deviation images of the OP method to the SP method from 100 transmission scans.

a suboptimal approach, the SP method still yielded lower variance attenuation maps than the OP method.

For each transmission scan, an estimate of the attenuation map was reconstructed, using 20 iterations of the grouped-coordinate ascent algorithms [15], [19] applied to the objective functions (10)–(12). Although $L^{\mathrm{OP}}(\mu)$ is globally convex, $L^{\mathrm{SP}}(\mu)$ is only locally convex [9]. This problem is not unique to the SP method; it is a general problem with transmission reconstruction (with nonzero background), even when prompt and delayed events are available separately [9]. Recently, a monotonic algorithm has been developed that is also suitable for the SP objective function [20]. In our simulations, we initialized the iterations with a resolution-matched FBP image and always observed monotonic increase in Φ for all cases.

We computed both the sample mean and sample standard deviation images for all methods. Fig. 2 shows horizontal profiles of the sample mean images. These profiles show that WLS is systematically negatively biased [9], whereas the OP and SP models appear free of such systematic bias. As explained in [9, App. A], the logarithm required by the WLS method negatively biases the reconstructed transmission images, and this bias increases as counts decrease. Since the rays traversing the center of the transmission phantom have the lowest counts, these regions show the largest negative bias.

Fig. 3 shows the sample standard deviation image for the SP method. To study the variance, we computed the ratio of the sample standard deviation image of the OP method to the SP method, shown in Fig. 4. Fig. 5 shows the histogram of the standard deviation ratios over all interior pixels. The OP model yields, on the average, about 11% higher standard deviation than the SP model. Although the absolute standard deviation values could be decreased by using longer scan durations, we expect relative standard deviations of OP and SP estimators to remain approximately constant for higher counts [15], [16]. This follows from the fact that analytic approximations (15) and (16) in [15] become more accurate with increasing counts, and these approximations predict that the ratio of standard deviation of different estimators remains constant, independent of total counts.

V. Covariance Approximations for Transmission Tomography

One can use analytic approximations proposed in [16] to predict the covariance of penalized-likelihood reconstruction methods, without exhaustive simulations. In [16], these approximations were shown to agree with empirical results from computer-simulated PET scans (without randoms precorrection), even for the highly nonlinear transmission reconstruction methods. Here, we apply the covariance approximation presented in [16] to the OP and SP methods and compare the results with experimental randoms precorrected transmission data.

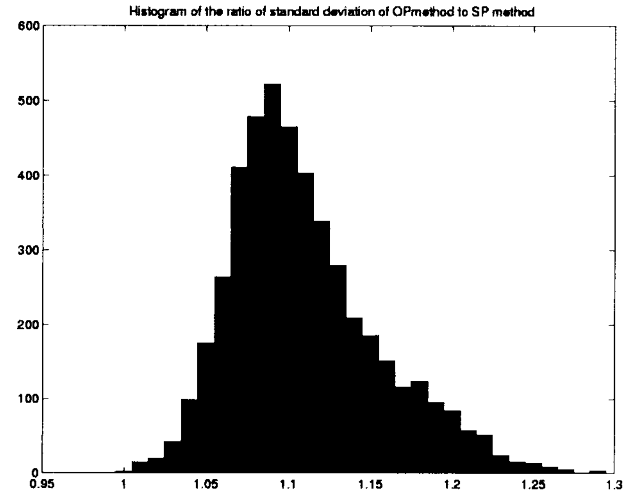

Fig. 5. Histogram of the ratio of standard deviations in reconstructed attenuation maps. The OP method yields, on the average, about 11% higher standard deviation than the proposed SP method.

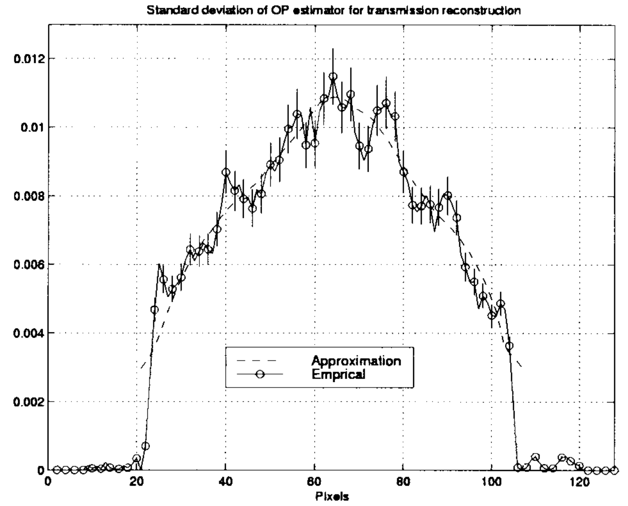

Fig. 6. Empirical standard deviation (with error bars) and the approximate standard deviation of the OP method for pixels along horizontal profile 90 through the attenuation map.

We can express both the OP (11) and SP (12) log-likelihood approximations in the form (6) with

$$h_n(l, y_n) = (y_n + d_n)\log(b_n e^{-l} + d_n) - (b_n e^{-l} + d_n) \tag{13}$$

and

$$d_n \triangleq \begin{cases} 0, & \text{OP} \\ 2r_n, & \text{SP.} \end{cases} \tag{14}$$

Combining the log-likelihood approximation with a roughness penalty forms the penalized log-likelihood objective function $\Phi(\mu)$, as in (8).

A first-order Taylor expansion of $\hat{\mu}(Y) = \arg\max_{\mu \geq 0} \Phi(\mu, Y)$ around $\bar{Y} \triangleq \mathrm{E}\{Y\}$ leads to the following approximation for the covariance of $\hat{\mu}$ [16]:

$$\begin{aligned} \mathrm{Cov}\{\hat{\mu}\} \approx & [-\nabla^{20}\Phi(\check{\mu}, \bar{Y})]^{-1} \nabla^{11}\Phi(\check{\mu}, \bar{Y}) \mathrm{Cov}\{Y\} \\ & \cdot [\nabla^{11}\Phi(\check{\mu}, \bar{Y})]' [-\nabla^{20}\Phi(\check{\mu}, \bar{Y})]^{-1} \end{aligned} \tag{15}$$

where

$$\check{\mu} \triangleq \arg\max_{\mu} \Phi(\mu, \bar{Y}). \tag{16}$$

Following [16]

$$-\nabla^{20}\Phi(\check{\mu}, \bar{Y}) = \boldsymbol{H} \triangleq \boldsymbol{A}' \operatorname{diag}\{u_n\} \boldsymbol{A} + \beta \boldsymbol{R}(\check{\mu}) \tag{17}$$

$$\nabla^{11}\Phi(\check{\mu}, \bar{Y}) = -\boldsymbol{A}' \operatorname{diag}\{c_n\} \tag{18}$$

where $\boldsymbol{A} = \{a_{nj}\}$ is the sparse system matrix and

$$u_n \triangleq \left(1 - \frac{d_n(\bar{y}(\mu^{\text{true}}) + d_n)}{(\bar{y}(\check{\mu}) + d_n)^2}\right) \bar{y}(\check{\mu}), \tag{19}$$

$$c_n \triangleq \frac{\bar{y}(\check{\mu})}{\bar{y}(\check{\mu}) + d_n} \tag{20}$$

and $\boldsymbol{R}(\mu) = \nabla^2 R(\mu)$. Substituting (17), (18), and $\mathrm{Cov}\{Y\} = \operatorname{diag}\{\bar{y}(\mu^{\text{true}}) + 2r_n\}$ into (15) yields the following approximation for the estimator covariance:

$$\mathrm{Cov}\{\hat{\mu}\} \approx \boldsymbol{H}^{-1} \boldsymbol{A}' \operatorname{diag}\{v_n\} \boldsymbol{A} \boldsymbol{H}^{-1} \tag{21}$$

with

$$v_n \triangleq \frac{\bar{y}^2(\check{\mu})(\bar{y}(\mu^{\text{true}}) + 2r_n)}{(\bar{y}(\check{\mu}) + d_n)^2}.$$

For the experimental transmission data we predicted the variance of $\hat{\mu}^{\text{OP}}$ and $\hat{\mu}^{\text{SP}}$ using the above approximations. Following the plug-in approach of [16], we replaced each $\bar{y}(\check{\mu})$ and $\bar{y}(\mu^{\text{true}})$ in (21) with the corresponding sample mean of the 100 transmission sinograms.[1] We used the preconditioned conjugate gradient method [21], [22] to compute selected diagonal elements of (21). Overall computation for computing the variance of each pixel was roughly equivalent to one maximization of $\Phi(\mu)$. Figs. 6 and 7 show the comparison of the empirical standard deviation and the approximate standard deviation of pixels through a horizontal cross section through the attenuation map for the OP method and the SP method. The predicted variance agrees well with the empirical results both for the OP and SP methods. These results show that, even for two-minute transmission scans, analytical approximations can be used reliably. For longer scans, with higher transmission counts, the agreement should be even better [16].

VI. NOISE PROPAGATION INTO EMISSION RECONSTRUCTION

In this section we derive approximate expressions to analyze the propagation of noise from the attenuation maps through the ACF's into reconstructed emission images. This analysis describes the effects of transmission noise on the final emission images, which may assist studies of the tradeoff between emission and transmission scan times, e.g., [23] and [24]. Dahlbom and Hoffman [11] have analyzed emission image noise for the special case of uniform density disk phantom (assuming both

---

[1] Although replacing $\bar{y}(\check{\mu})$ and $\bar{y}(\mu^{\text{true}})$ in (21) with the sample mean of the transmission sinograms is impractical, it enables us to compute quickly the approximations for many pixels in the reconstructed image. In Section VI we present the results of variance approximations for a set of pixels for noise propagation into emission images, using the true plug-in approach (where we replace $\bar{y}(\check{\mu})$ and $\bar{y}(\mu^{\text{true}})$ with noisy measurements). There, we show that the predictions again agree well with empirical standard deviation values.

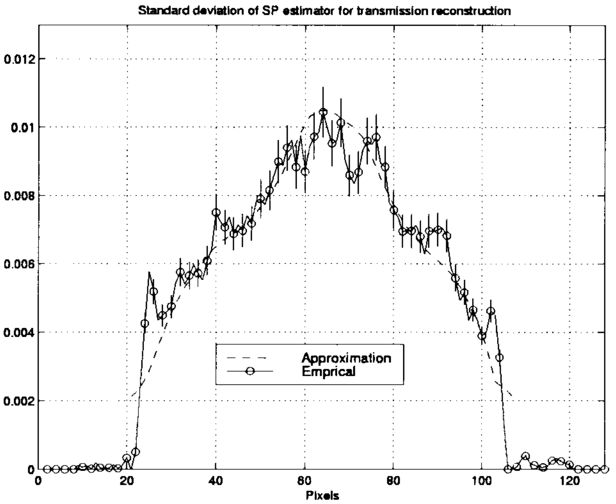

Fig. 7. Empirical standard deviation (with error bars) and the approximate standard deviation of the SP method for pixels along horizontal profile 90 through the attenuation map.

emission and transmission images are reconstructed using the FBP method). The covariance approximations presented here apply to arbitrary objects for attenuation maps reconstructed by penalized-likelihood estimators with quadratic regularization.

A. Theory

To isolate the effects of transmission noise on the resultant emission images, we consider noiseless emission measurements and we consider the FBP method for reconstructing emission images after correcting for attenuation using noisy attenuation maps. We assume the noiseless emission measurements are

$$z_n = e^{-l_n(\mu^{\mathrm{true}})} p_n \tag{22}$$

where

$$p_n = \sum_{k=1}^{M} g_{nk}\lambda_k$$

is the attenuation-free projection of the emission image and where $\lambda = [\lambda_1 \cdots \lambda_P]'$ denotes the vector of radio-isotope concentration. $\boldsymbol{G} = \{g_{nk}\}$ represents the tomographic system response, including the geometric system model, ray-dependent factors (e.g., detector efficiency factors, dead time, radio-isotope decay), and pixel-dependent factors such as spatial variations in sensitivity. And $e^{-l_n(\mu^{\mathrm{true}})}$ (with $l(\mu^{\mathrm{true}}) = \boldsymbol{A}\mu^{\mathrm{true}}$) represents the survival probability for the nth ray. The noiseless emission measurements z_n (22) are corrected for attenuation, using ACF's based on the attenuation-map estimates $\hat{\mu}$. If one directly corrects the emission measurements for attenuation by multiplication, the resultant images have some artifacts because of the resolution mismatch between emission and transmission sinograms [6], [7]. Thus, one needs to smooth the emission sinogram to the same resolution as the survival probabilities. We can write the attenuation-corrected emission sinogram as follows:

$$\hat{z}_n = e^{l_n(\hat{\mu})}\ \mathrm{smooth}\{e^{-l_n(\mu^{\mathrm{true}})} p_n\}. \tag{23}$$

For FBP reconstruction of the emission images, we consider the constrained least-squares (CLS) window corresponding to (50) of [25]

$$\frac{\mathrm{sinc}\{ku\}/\mathrm{sinc}\{u\}}{\mathrm{sinc}^2\{ku\} + \alpha u^3}, \qquad u \in \left[0, \frac{1}{2}\right] \tag{24}$$

where u denotes spatial frequency in cycles per radial sample, k is the ratio of the strip width to the pixel size of the system model, and α is linearly related to β below [25]. (The detector response is a rectangular function with frequency response $\mathrm{sinc}(ku)$.) Dividing by $\mathrm{sinc}(u)$ in the numerator compensates for the linear interpolation step of the FBP method. The FBP algorithm with the above smoothing window (24) is essentially equivalent to a quadratically penalized unweighted least-squares (QPULS) estimator without the nonnegativity constraint [25]. The QPULS estimator is defined as [26]

$$\begin{aligned}\hat{\lambda}^{\mathrm{QPULS}} &= \arg\min_{\lambda} \|\hat{z} - \boldsymbol{G}\lambda\|^2 + \beta\lambda'\boldsymbol{R}_o\lambda \\ &= [\boldsymbol{G}'\boldsymbol{G} + \beta\boldsymbol{R}_o]^{-1}\boldsymbol{G}'\hat{z}\end{aligned} \tag{25}$$

with

$$\boldsymbol{R}_o[j,k] = \begin{cases} \sum_l w_{jl}, & k = j \\ -w_{jk}, & k \neq j \end{cases} \tag{26}$$

where $w_{jk} = 1$ for horizontal and vertical neighboring pixels and zero otherwise. Since this estimator is linear, its covariance is

$$\begin{aligned}\mathrm{Cov}\{\hat{\lambda}^{\mathrm{QPULS}}\} = {}& [\boldsymbol{G}'\boldsymbol{G} + \beta\boldsymbol{R}_o]^{-1}\boldsymbol{G}'\,\mathrm{Cov}\{\hat{z}\} \\ & \cdot \boldsymbol{G}[\boldsymbol{G}'\boldsymbol{G} + \beta\boldsymbol{R}_o]^{-1}.\end{aligned} \tag{27}$$

We must find $\mathrm{Cov}\{\hat{z}\}$ to complete the above approximation. For simplicity we first make the following approximation:

$$\mathrm{smooth}\{e^{-l_n(\mu^{\mathrm{true}})} p_n\} \approx e^{-l_n(\check{\mu})}\ \mathrm{smooth}\{p_n\} \tag{28}$$

where $\check{\mu}$ is defined in (16). We plug this into (23)

$$\hat{z}_n \approx e^{l_n(\hat{\mu})} e^{-l_n(\check{\mu})}\ \mathrm{smooth}\{p_n\} \tag{29}$$

and approximate $\mathrm{Cov}\{\hat{z}\}$ as

$$\mathrm{Cov}\{\hat{z}\} \approx \boldsymbol{D}\,\mathrm{Cov}\{t(\hat{\mu})\}\boldsymbol{D}' \tag{30}$$

where $t(\hat{\mu}) \triangleq [t_1(\hat{\mu}) \cdots t_N(\hat{\mu})]'$ with $t_n(\hat{\mu}) = e^{l_n(\hat{\mu}) - l_n(\overline{\mu})}$ and $\overline{\mu} \triangleq \mathrm{E}\{\hat{\mu}\}$ and $\boldsymbol{D} \triangleq \mathrm{diag}\{\ \mathrm{smooth}\{p_n\}\}$.

Using first-order Taylor expansion around $\overline{\mu}$, we approximate $\mathrm{Cov}\{t(\hat{\mu})\}$ as

$$\mathrm{Cov}\{t(\hat{\mu})\} \approx \boldsymbol{A}\,\mathrm{Cov}\{\hat{\mu}\}\boldsymbol{A}'. \tag{31}$$

Finally, plugging (30) and (31) into (27) yields

$$\begin{aligned}\mathrm{Cov}\{\hat{\lambda}^{\mathrm{QPULS}}\} \approx {}& [\boldsymbol{G}'\boldsymbol{G} + \beta\boldsymbol{R}_o]^{-1}\boldsymbol{G}'\boldsymbol{D}\boldsymbol{A}\mathrm{Cov}\{\hat{\mu}\} \\ & \cdot \boldsymbol{A}'\boldsymbol{D}\boldsymbol{G}[\boldsymbol{G}'\boldsymbol{G} + \beta\boldsymbol{R}_o]^{-1}.\end{aligned} \tag{32}$$

The variance of the estimated total activity within a region of interest (ROI), i.e., $\hat{\theta}_e = e'\hat{\lambda}^{\mathrm{QPULS}}$, is simply

$$\mathrm{Var}\{\hat{\theta}_e\} = e'\,\mathrm{Cov}\{\hat{\lambda}^{\mathrm{QPULS}}\}e \tag{33}$$

where e is a column vector of length M that equals unity for the pixels in the region of interest and zero elsewhere.

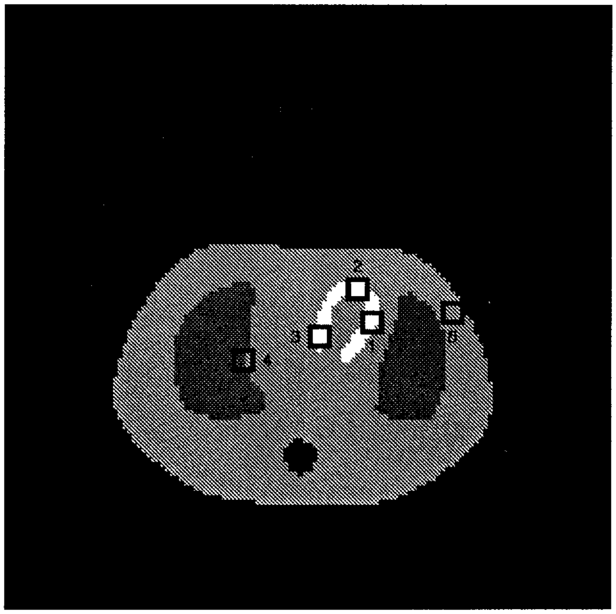

Fig. 8. Emission phantom with several rectangular regions for noise computation.

To within the accuracy of the preceding approximations, (32) shows the first-order propagation of the noise from the attenuation map $\hat{\mu}$ into the emission reconstruction and (30) also shows that $\mathrm{Cov}\{\hat{\mu}\}$ is scaled quadratically by the attenuation-free emission projections p_n (22) before propagating into emission image covariance (since it is sandwiched between $\boldsymbol{D}$ matrices).

B. Results

We simulated noiseless emission measurements (22) for the emission phantom shown in Fig. 8, using the same system specifications as the experimental transmission data. (The rectangular regions numbered 1–5 are regions of interest used at the end of this section). The spine, lungs, soft tissue, and heart had relative radioactivity concentrations of zero, one, two, and four, respectively. The effects of attenuation were included (22) by calculating survival probabilities from an attenuation map reconstructed from a 5-h transmission scan. To reconstruct this attenuation map, we used the very precise saddle point (SD) approximation that we introduced previously [14], [15], along with an edge-preserving penalty function [19]. This attenuation map was also used to generate the simulated emission phantom in Fig. 8 (by segmentation), so that the computer simulated emission activity and experimental attenuation maps were well aligned.

After smoothing the noiseless emission measurements to match the resolution of the transmission data [6], [7], we applied ACF's computed from the noisy attenuation map estimates $\hat{\mu}^{\mathrm{OP}}$ and $\hat{\mu}^{\mathrm{SP}}$ that were reconstructed from each experimental transmission scan. We reconstructed emission images, using FBP with the CLS window (24). Fig. 9 shows the sample mean emission image with ACF's based on the SP method computed from 100 2-min transmission scans, as explained in Section V. (The mean image of the OP method is not shown, since it was very similar to that of the SP method.)

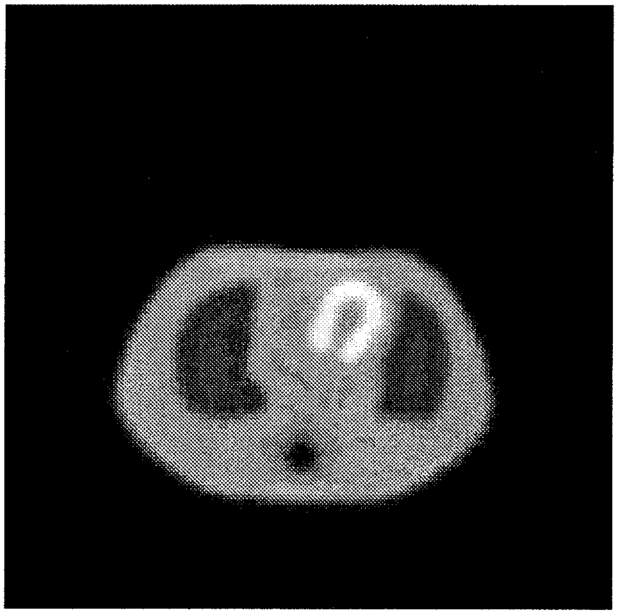

Fig. 9. Empirical sample mean of emission images reconstructed with ACF's based on 100 different estimates of $\hat{\mu}^{\mathrm{SP}}$.

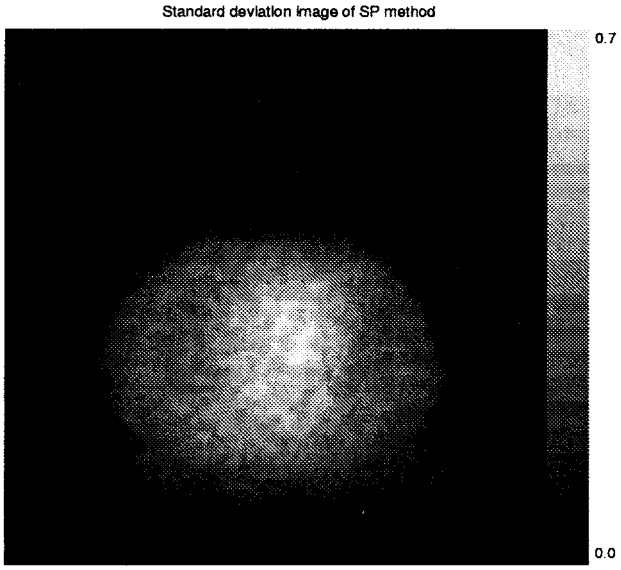

Fig. 10. Sample standard deviation image of emission reconstruction with ACF's based on the SP method.

Fig. 10 shows the sample standard deviation image of the 100 emission reconstructions with ACF's based on the SP method. To study the noise due to different methods, we computed the ratio of sample standard deviation images of emission reconstruction with ACF's, based on the OP method and SP method shown in Fig. 11. Fig. 12 shows the histogram of the standard deviation ratios over all interior pixels. Attenuation correction based on the OP model yielded about 20% higher standard deviation than the SP model, on average.

To assess the accuracy of our analytical approximations, we compared the empirical variances described above to the variances predicted by (33). We used the preconditioned

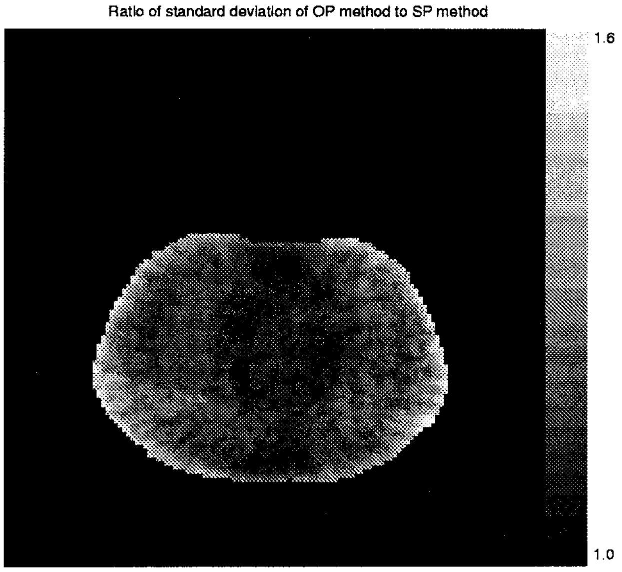

Fig. 11. Ratio of sample standard deviation images of emission reconstruction with ACF's based on the OP method and the SP method.

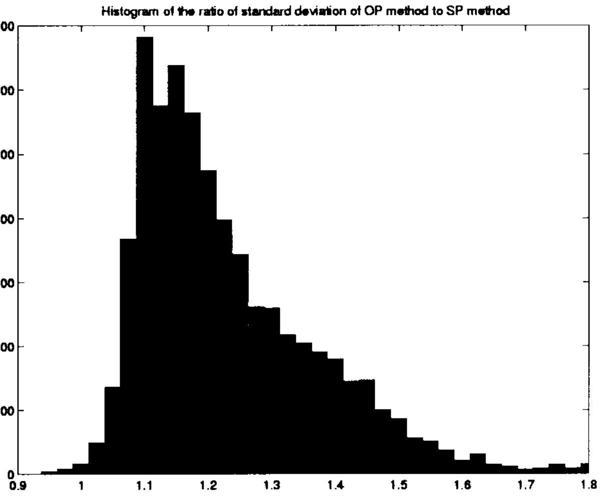

Fig. 12. Histogram of the ratio of standard deviations in the reconstructed emission images with ACF's based on the OP model and the SP model. Attenuation correction factors based on the OP model yielded about 20% higher standard deviation than the SP model, on average.

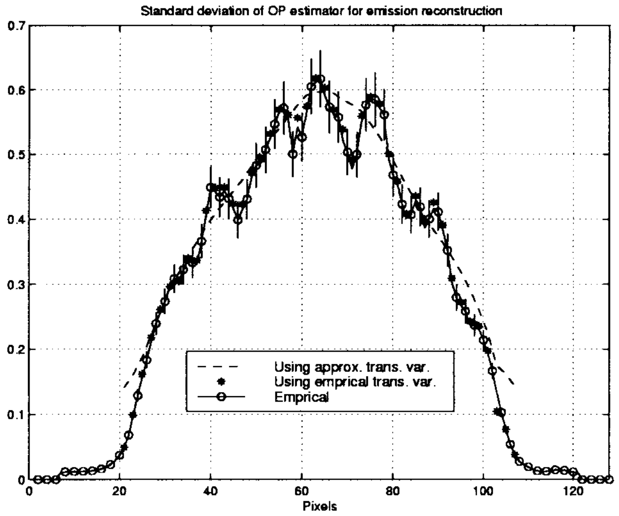

Fig. 13. Empirical standard deviation (with error bars) and the approximate standard deviation of the OP method (using both empirical transmission variance and approximate transmission variance) for pixels along horizontal profile 90 through the reconstructed emission images.

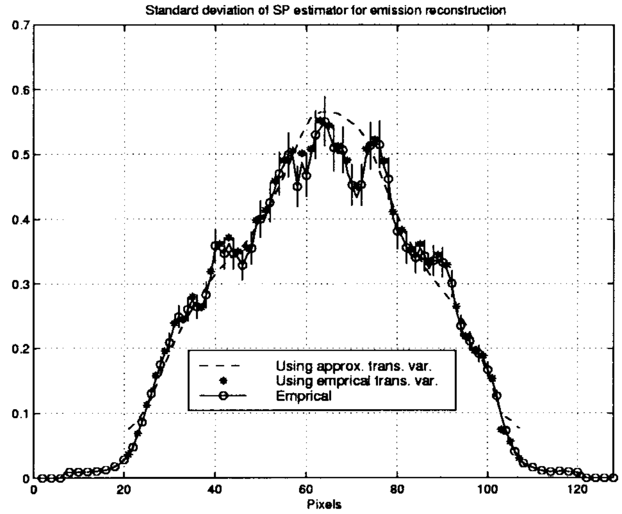

Fig. 14. Empirical standard deviation (with error bars) and the approximate standard deviation of the SP method (using both empirical transmission variance and approximate transmission variance) for pixels along a horizontal profile 90 through the reconstructed emission images.

conjugate gradient method to compute (33) for a set of pixels in the reconstructed emission image. We determined the elements of $\mathrm{Cov}\{\hat{\mu}\}$ in (32) two different ways. One way used the approximation (21) and the other used the empirical covariance of the 100 independent attenuation map reconstructions.[2] Although replacing $\mathrm{Cov}\{\hat{\mu}\}$ with an empirical covariance is impractical for routine use, it helps establish the accuracy of approximation (33). Figs. 13 and 14 compare the empirical standard deviations and the approximate standard deviations of pixels along a horizontal profile through the emission images. The analytical approximations for transmission noise propagation agree well with the empirical results, and confirm the reduction in noise for the SP method compared to the OP method.

Table I shows the percent standard deviation of the activity within the five different 3 × 3-pixel ROI's shown in Fig. 8 for the reconstructed images, with ACF's based on the OP method and SP method. For each ROI, we also implemented the practical plug-in approach for computing (21) in which we replaced each $\overline{y}(\check{\mu})$ and $\overline{y}(\mu^{\mathrm{true}})$ in (21) with the corresponding noisy sinogram element y_n. We computed variance approximation (33) for each of the 100 sinograms. Table I shows the

---

[2] Instead of computing the empirical covariance directly from the independent attenuation map reconstructions, we used the following computationally more efficient method. It can be seen from (32) and (33) that $\mathrm{Var}\{\hat{\theta}_e\} = \mathrm{Var}\{\mathrm{S}\}$ where $S \triangleq c'\hat{\mu}$ and $c' \triangleq e'[\boldsymbol{G}'\boldsymbol{G} + \beta \boldsymbol{R}_o]^{-1}\boldsymbol{G}'\boldsymbol{D}\boldsymbol{A}$. Using the preconditioned conjugate gradient method [21], [22], we precompute the row vector c' only once, and then compute the scalar S for each independent attenuation map reconstruction. Then, finally, the sample variance of S is computed.

TABLE I

EMPIRICAL PERCENT STANDARD DEVIATION AND THE APPROXIMATE ANALYTICAL PERCENT STANDARD DEVIATION OF EMISSION RECONSTRUCTION USING ACF'S BASED ON THE OP METHOD AND THE SP METHOD (USING BOTH EMPIRICAL TRANSMISSION VARIANCE AND APPROXIMATE TRANSMISSION VARIANCE AND PLUG-IN TRANSMISSION VARIANCE) FOR DIFFERENT REGIONS SHOWN IN FIG. 8. LAST COLUMN SHOWS THE EMPIRICAL PERCENT NOISE OF THE REGIONS DUE TO ONLY EMISSION NOISE FOR TWO MILLION COUNTS PER EMISSION SCAN

Region	OP Method				SP Method				Emission Noise
	Empr. Std.	App. Std. (emp tr var)	App. Std. (app tr var)	App. Std. (plug-in)	Empr. Std.	App Std. (emp tr var)	App. Std. (app tr var)	App. Std. (plug-in)	
1	11.35	11.56	12.28	12.23 ±0.14	10.20	10.39	11.34	10.88 ±0.10	2.60
2	12.04	12.14	10.82	10.74 ±0.14	10.93	10.98	9.80	9.39 ±0.09	2.12
3	16.87	17.09	14.74	15.07 ±0.22	15.68	15.91	14.32	13.99 ±0.16	2.79
4	25.55	25.72	23.86	23.59 ±0.27	24.85	25.34	23.53	22.54 ±0.20	4.66
5	8.89	8.89	9.63	9.74 ±0.10	7.30	7.35	7.61	7.49 ±0.07	2.47

sample means (and standard errors) of the plug-in predicted variances for each ROI. The OP model yields 8–23% higher standard deviation than the SP model and all the analytical approximations agree well with empirical standard deviation values.

For comparison purposes, we simulated 100 noisy emission sinograms having an average of 2M counts per scan, and performed FBP reconstruction of the emission images. For the ACF's, we used the empirical mean of the transmission scans, to ensure that only emission noise affected the reconstructions. (Since the emission noise is inversely proportional to the square root of the total counts per scan, one could also predict emission noise for other count levels.) Table I shows the empirical standard deviations for different ROI's due to emission noise. These simulations illustrate the relative effects of emission and transmission noise.

VII. CONCLUSIONS

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 and the exact log likelihood is complicated. We compared different approximations for the exact log likelihood, using experimental PET transmission scans. The WLS method leads to systematic bias and penalized-likelihood methods, based on the OP model, lead to higher standard deviation (on the average about 11%) than our SP model, which matches both the first and second moments of the measurement statistics.

We also investigated the reconstruction results for the SD approximation that we introduced previously [14], [15]. We observed very close agreement between the SD method and truncated exact log likelihood for each reconstruction in our simulations. However, for the transmission scans reported here, the difference between the SD and the SP method was not statistically significant. Thus, we did not include the reconstruction results from the SD method in this paper. 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 [27].

We applied the covariance approximations to the attenuation map estimates from the OP method and SP method and demonstrated that these approximations agree with the empirical results from experimental PET transmission scans. These approximations can be used to determine the variance of transmission reconstruction to investigate parameters of interest (e.g., regularization parameters) and can supplement simulations.

We also developed approximations to analyze the propagation of noise from attenuation maps into emission reconstruction. For this purpose, we assumed noiseless emission measurements and developed approximations for the covariance of emission reconstruction with ACF's computed from noisy attenuation maps. The approximations agree with the empirical results and describe the propagation of noise from attenuation maps into emission reconstruction.

Both approximations and empirical results showed the interesting property that when the transmission scan noise was propagated into the emission images, the relative differences in variances between the OP model and the SP model can be even greater than when one considers the noise in the attenuation maps alone. The ACF's computed from the OP model yielded, on the average, about 20% higher standard deviation than the SP model in the reconstructed emission images, compared with 11% differences in the attenuation maps.

We plan to apply the proposed methods to emission tomography, where even higher AC rates than the transmission tomography are common, particularly in three-dimensional (3-D) PET. Moreover, in 3-D PET, very large data sets may deter separate acquisition of prompt and delayed coincidences, so the real-time subtraction methods are usually used. The potential benefit of the proposed models may be even greater in emission studies.

REFERENCES

[1] S. C. Huang, E. J. Hoffman, M. E. Phelps, and D. E. Kuhl, "Quantitation in positron emission computed tomography: 2 effects of inaccurate attenuation correction," *J. Comput. Assisted Tomogr.*, vol. 3, no. 6, pp. 804–814, Dec. 1979.

[2] 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. Comput. Assisted Tomogr.*, vol. 10, no. 5, pp. 845–850, 1986.

[3] R. H. Huesman, S. E. Derenzo, J. L. Cahoon, A. B. Geyer, W. W. Moses, D. C. Uber, T. Vuletich, and T. F. Budinger, "Orbiting transmission source for positron tomography," *IEEE Trans Nucl. Sci.*, vol. 35, pp. 735–739, Feb. 1988.

[4] S. R. Meikle, M. Dahlbom, and S. R. Cherry, "Accuracy of attenuation correction PET due to transmission processing," *J. Nucl. Med.*, vol. 33, no. 5, p. 862, May 1992.

[5] E. Ü. Mumcuoğlu, R. M. Leahy, and S. R. Cherry, "Bayesian reconstruction of PET images: Methodology and performance analysis," *Phys. Med. Biol.*, vol. 41, no. 9, pp. 1777–1807, Sept. 1996.

[6] A. Chatziioannou and M. Dahlbom, “Detailed investigation of transmission and emission data smoothing protocols and their effects on emission images,” *IEEE Trans. Nucl. Sci.*, vol. 43, pp. 290–294, Feb. 1996.

[7] ———, “Detailed investigation of transmission and emission data smoothing protocols and their effects on emission images,” in *Proc. IEEE Nuclear Science Symp. Medical Imaging Conf.*, 1994, vol. 4, pp. 1568–1572.

[8] S. L. Bacharach, M. A. Douglas, R. E. Carson, P. J. Kalkowski, N. M. T. Freedman, P. Perrone-Filardi, and R. O. Bonow, “Three dimensional registration of cardiac PET attenuation scans,” *J. Nucl. Med.*, vol. 33, no. 5, p. 881, May 1992.

[9] J. A. Fessler, “Hybrid Poisson/polynomial objective functions for tomographic image reconstruction from transmission scans,” *IEEE Trans. Image Processing*, vol. 4, no. 10, pp. 1439–1450, Oct. 1995.

[10] 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. Comput. Assisted Tomogr.*, vol. 5, no. 3, pp. 391–400, 1981.

[11] M. Dahlbom and E. J. Hoffman, “Problems in signal-to-noise ratio for attenuation correction in high resolution PET,” *IEEE Trans. Nucl. Sci.*, vol. 34, pp. 288–293, Feb. 1987.

[12] D. G. Politte and D. L. Snyder, “Corrections for accidental coincidences and attenuation in maximum-likelihood image reconstruction for positron-emission tomography,” *IEEE Trans. Med. Imag.*, vol. 10, pp. 82–89, Mar. 1991.

[13] M. Yavuz and J. A. Fessler, “Objective functions for tomographic reconstruction from randoms-precorrected PET scans,” in *Proc. IEEE Nuclear Science Symp. Medical Imaging Conf.*, 1996, vol. 2, pp. 1067–1071.

[14] ———,“New statistical models for randoms-precorrected PET scans,” in *Information Processing in Medical Imaging*, (Lecture Notes in Computer Science), J. Duncan and G. Gindi, Eds. Berlin, Germany: Springer-Verlag, 1997, vol. 1230, pp. 190–203.

[15] ———, “Statistical image reconstruction methods for randoms-precorrected PET scans,” *Med. Imag. Anal.*, vol. 2, no. 4, pp. 369–378, 1998.

[16] J. A. Fessler, “Mean and variance of implicitly defined biased estimators (such as penalized maximum likelihood): Applications to tomography,” *IEEE Trans. Image Processing*, vol. 5, pp. 493–506, Mar. 1996.

[17] K. Sauer and C. Bouman, “A local update strategy for iterative reconstruction from projections,” *IEEE Trans. Signal Processing*, vol. 41, pp. 534–548, Feb. 1993.

[18] J. A. Fessler and W. L. Rogers, “Spatial resolution properties of penalized-likelihood image reconstruction methods: Space-invariant tomographs,” *IEEE Trans. Image Processing*, vol. 5, pp. 1346–1358, Sept. 1996.

[19] J. A. Fessler, E. P. Ficaro, N. H. Clinthorne, and K. Lange, “Grouped-coordinate ascent algorithms for penalized-likelihood transmission image reconstruction,” *IEEE Trans. Med. Imag.*, vol. 16, pp. 166–175, Apr. 1997.

[20] H. Erdoğan and J. A. Fessler, “Fast monotonic algorithms for transmission tomography,” *IEEE Trans. Med. Imag.*, to be published.

[21] N. H. Clinthorne, T. S. Pan, P. C. Chiao, W. L. Rogers, and J. A. Stamos, “Preconditioning methods for improved convergence rates in iterative reconstructions,” *IEEE Trans. Med. Imag.*, vol. 12, pp. 78–83, Mar. 1993.

[22] J. A. Fessler and S. D. Booth, “Conjugate-gradient preconditioning methods for shift-variant PET image reconstruction,” *IEEE Trans. Image Processing*, vol. 8, pp. 688–699, May 1999.

[23] T. Beyer, P. E. Kinahan, and D. W. Townsend, “Optimization of transmission and emission scan duration in 3D whole-body PET,” *IEEE Trans. Nucl. Sci.*, vol. 44, pp. 2400–2407, Dec. 1997.

[24] H. Erdoǵan and J. A. Fessler, “Scan time optimization for post-injection PET scans,” in *Proc. IEEE Nuclear Science Symp. Medical Imaging Conf.*, 1998, vol. 3, pp. 1842–1846.

[25] J. A. Fessler, “Resolution properties of regularized image reconstruction methods,” Comm. and Sign. Proc. Lab., Dept. of EECS, Univ. Michigan, Ann Arbor, Tech. Rep. 297, Aug. 1995.

[26] ———, “Approximate variance images for penalized-likelihood image reconstruction,” in *Proc. IEEE Nuclear Science Symp. Medical Imaging Conf.*, 1997, vol. 2, pp. 949–952.

[27] A. O. Hero, J. A. Fessler, and M. Usman, “Exploring estimator bias-variance tradeoffs using the uniform CR bound,” *IEEE Trans. Signal Processing*, vol. 44, pp. 2026–2041, Aug. 1996.