Quantitative comparison of FBP, EM, and Bayesian reconstruction algorithms, including the impact of accurate system modeling, for the IndyPET scanner *†

Thomas Frese*,[1], Ned C. Rouze[2], Charles A. Bouman[1], Ken Sauer[3], and Gary D. Hutchins[2]

[1]School of Electrical and Computer Engineering, Purdue University, 1285 ECE Building, West Lafayette, IN 47907-1285, USA, Phone: (765) 494-0340, Fax: (765) 494-3358, Email: {frese, bouman}@ecn.purdue.edu

[2] Indiana University School of Medicine, Imaging Science, Department of Radiology, 541 Clinical Drive, Indianapolis, IN 46202-5111, USA, Phone: (317) 278-7430, Fax: (317) 274-4074, Email: {nrouze,gdhutchi}@iupui.edu

[3]University of Notre Dame, Department of Electrical Engineering, 270 Fitzpatrick Hall, Notre Dame, IN 46556-5637, USA, Phone: (219) 631-6999, Fax: (219) 631-4393, Email: sauer@nd.edu.

Abstract

We quantitatively compare filtered backprojection (FBP), expectation maximization (EM), and Bayesian reconstruction algorithms as applied to the IndyPET scanner - a dedicated research scanner which has been developed for small and intermediate field of view imaging applications. In contrast to previous approaches that rely on Monte Carlo simulations, a key feature of our investigation is the use of an empirical system kernel determined from scans of line source phantoms. This kernel is incorporated into the forward model of the EM and Bayesian algorithms to achieve resolution recovery.

Three data sets are used, data collected on the IndyPET scanner using a bar phantom and a Hoffman 3-D brain phantom, and simulated data containing a hot lesion added to a uniform background. Reconstruction quality is analyzed quantitatively in terms of bias-variance measures (bar phantom) and mean square error (lesion phantom).

We observe that without use of the empirical system kernel, the FBP, EM, and Bayesian algorithms give similar performance. However, with the inclusion of the empirical kernel, the iterative algorithms provide superior reconstructions compared to FBP, both in terms of visual quality and quantitative measures. Furthermore, Bayesian methods outperform EM. We conclude that significant improvements in reconstruction quality can be realized by combining accurate models of the system response with Bayesian reconstruction algorithms.

Keywords: *positron emission tomography, iterative methods, image reconstruction, Bayesian methods*

Correspondence to: Thomas Frese
School of ECE, Purdue University
1285 Electrical and Computer Engineering Building
West Lafayette, IN 47907-1285
Office: (312) 787-0668, Fax: (765) 494-3358

*This work was supported by the National Science Foundation under Grant MIP97-07763, the National Cancer Institute under Grant ICMIC P20-CA86350-2, and the State of Indiana 21st Century Research and Technology Fund "Indiana Center of Excellence in Biomedical Imaging".

†To appear in the IEEE Trans. on Medical Imaging.

Email: frese@ecn.purdue.edu

1 Introduction

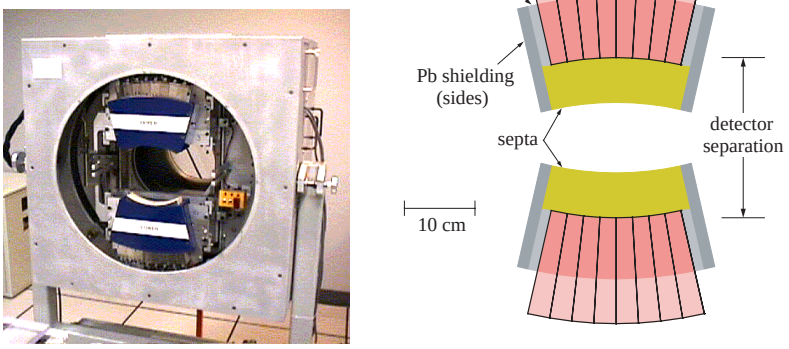
Iterative methods for positron emission tomography (PET) reconstruction can theoretically improve reconstruction quality by accurately modeling the data acquisition process and by regularizing the ill-posed inverse problem. Importantly, iterative methods can achieve significant resolution recovery by incorporating an accurate model of the tomography scanner accounting for sinogram blurring due to detector crystal penetration, inter-crystal scatter, depth of interaction and other effects[1, 2, 3]. However, this advantage is often not fully exploited in practical implementations of iterative reconstruction algorithms.

Early investigations for the expectation maximization (EM) reconstruction algorithm[4] have found that a more accurate system kernel can reduce bias in region-of-interest (ROI) quantitation[5, 1]. However, the more accurate kernel increased the variance of the ROI estimates obtained from EM reconstructions, even for simulated projection data that had been generated with the same kernel. This effect is likely due to the fact that a more accurate system kernel results in a more ill-posed reconstruction problem[5], which is consistent with the observation of ringing artifacts[6]. The increased ill-posedness adversely affects the quality of EM reconstructions, particularly, since EM does not use any explicit regularization.

Bayesian reconstruction techniques[7, 8, 9, 10, 11, 12, 13, 14] on the other hand, explicitly address the ill-posedness by regularizing the reconstruction through use of a prior model such as a Markov random field (MRF)[15, 7, 8, 16, 17, 18, 19]. Consequently, we would expect Bayesian methods to provide better reconstruction quality as compared to EM, especially when a more accurate system kernel is used. Qi, *et al.*, [6, 20] have successfully applied 3D Bayesian MAP reconstruction combined with an advanced response model[3] to data from the microPET[21] and the Siemens/CTI HR+[22] scanners. The results demonstrate that this approach can give superior resolution and ROI quantitation at lower variance as compared to FBP and OSEM[20, 23].

In this study, we quantitatively compare FBP, EM, and Bayesian reconstruction algorithms and evaluate the impact of accurately modeling the scanner for iterative reconstruction. Our results are based on simulated data and phantom data acquired on the IndyPET tomography scanner[24], a high resolution dedicated research scanner for use in small animal, intermediate sized animal and small field of view human (brain, breast) imaging applications. The scanner has a non-circular detector geometry consisting of two, approximately planar detector banks mounted on a rotatable gantry as shown in Fig. 1.

In contrast to previous studies which have relied on analytical modeling of detector properties[25, 26]



(a) Front view of open scanner (b) Detector geometry

Figure 1: (a) Photograph of the IndyPET scanner with the front cover removed. (b) Layout of the scanner showing the detector modules, septa, and shielding. The detector separation is adjustable.

or on Monte Carlo simulations[27, 28, 29, 3], we have obtained a system kernel for the IndyPET scanner empirically using scans of line source phantoms. This empirical approach is simplified by the non-circular detector geometry of the IndyPET scanner which allows for measuring the full range of incident photon angles for any pair of detectors using only sources positioned near the center of the field of view. While our approach is also applicable to a circular scanner geometry, more measurements would be required since the incident photon angle is a strong function of radial source position. We note that the empirical approach implemented in this study ignores the depth dependence along the lines of response, though it could be added as an additional model parameter. As shown in Fig. 2, our results indicate that even for a centered source and normal photon incidence, our empirical system kernel is very different from the triangular shape expected for idealized detectors.

For EM and the Bayesian reconstruction methods, we compare reconstruction quality for our empirical (2-D) kernel and the simple triangular kernel shown in Fig. 2. The objective of this comparison is to demonstrate the importance of accurate system modeling for iterative reconstruction, and, more importantly, to investigate the relative improvement in reconstruction quality separately for each algorithm. We note that the use of the triangular system kernel is not intended to reflect performance of the empirical system kernel compared to existing accurate system modeling techniques such as Monte Carlo methods[3]. To further investigate the importance of accurate system modeling for iterative reconstruction, we use parametric system kernels of different widths to study the impact of deviations in the kernel.

In addition to FBP and EM, we compare three different prior models for Bayesian reconstruction; a Gaussian Markov random field (GMRF), a generalized Gaussian Markov Random field (GGMRF)[18] and

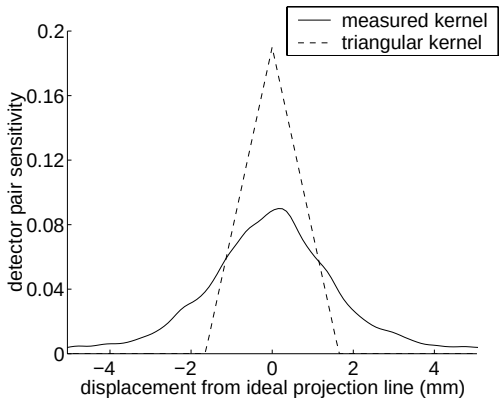


Figure 2: Plots of triangular and empirical detector kernels for the IndyPET scanner which characterize the response of a single detector pair as a function of source offset from the ideal projection line. The triangular kernel is the analytical result, based on solid angle, for a source located midway between idealized detectors of width 3.3 mm. The empirical kernel shown here was obtained as a slice from Fig. 7 and corresponds to a projection with normal detector incidence for a source centered between the detectors.

a recently proposed nonlinear wavelet prior model[30, 31] that performs space-adaptive regularization. Furthermore, we evaluate a modification of Bayesian methods proposed by Fessler and Rogers[32] that produces reconstructions with approximately uniform resolution. All Bayesian methods are based on maximum *a posteriori* (MAP) estimation using iterative coordinate descent (ICD) optimization[9, 14].

To compare reconstruction methods, three data sets are used: The first data set was acquired on the IndyPET scanner using a bar phantom. This phantom is similar to an X-ray line-pair phantom and generates square-wave emission profiles with different periodicities. Reconstructions are analyzed in terms of spatial bias and variance which are calculated by fitting harmonic series to the reconstructed wave forms. Different combinations of reconstruction algorithms and system kernels are compared in terms of the bias-variance envelopes obtained by varying a smoothing parameter for each method. In addition, we compare the convergence speed of EM and Bayesian GMRF MAP. The second data set was collected on the IndyPET scanner using a Hoffman 3-D brain phantom. Reconstruction images are compared in terms of visual quality. The third data set consists of simulated projection data containing a circular lesion added to an approximately uniform background. The reconstructions are analyzed quantitatively in terms of the expected mean square error (MSE) for the lesion and the background regions. While mean square error is not a direct indicator of clinical performance, it provides an objective measure of the quantitative accuracy of the reconstruction.

2 IndyPET Scanner

As illustrated in Fig. 1, the distinguishing feature of the IndyPET scanner is the use of two, approximately planar detector banks with adjustable separation which are mounted on a rotatable gantry. In comparison

to using a full ring of detectors, this geometry results in reduced parallax and approximately uniform resolution throughout the field of view. The detector separation can be adjusted for each subject, giving increased sensitivity for small subjects. The use of two detector banks is also economical since fewer detectors are required compared to a circular geometry. However, because the detectors do not form a full ring, data must be collected at different gantry angles, resulting in increased imaging time.

The two detector banks of the scanner consist of 8 CTI-HR[33] BGO detector modules each. The radius of curvature of the detector banks is 41.2 cm, the same as in the original CTI-HR scanner. The detector crystal in each of the 16 detector modules is cut into a 7 (transaxial) by 8 (axial) array of crystal segments with a transaxial spacing of 3.3 mm and an axial spacing of 6.3 mm. The crystal depth is 30 mm.

As indicated in Fig. 1(b), removable septa can be mounted directly in front of the detectors. These septa limit the oblique angles between axial slices and significantly reduce the influence of scatter[24]. The septa were used for all data sets collected in this study.

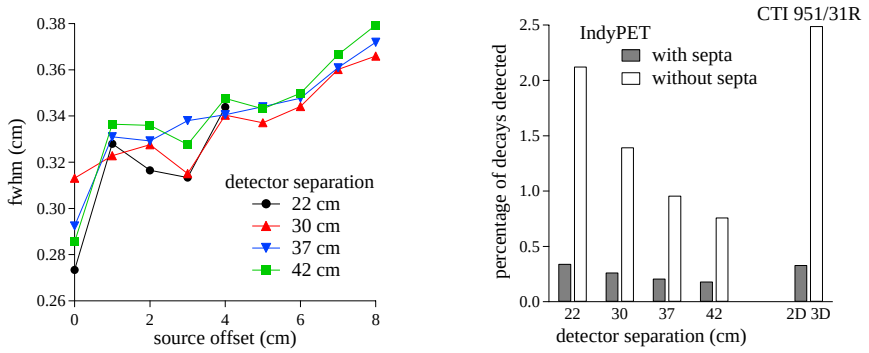
The scanner uses CTI bucket controllers and coincidence electronics. Coincidence events are sorted into sinogram bins corresponding to angle and displacement using the CTI ACS-II real-time sorter. For all results presented here, automatic randoms correction was performed. The bucket electronics report the singles count rates for each bucket with and without deadtime correction. The ratios of these count rates were used for deadtime correction.

Figure 3(a) shows the resolution of the IndyPET scanner as a function of source position relative to the center of the field of view for different detector separations. The resolution throughout the field of view is within approximately 20% of the resolution at the center and is nearly independent of detector separation. Figure 3(b) shows the sensitivity of the scanner for typical values of detector separation in comparison to that of a ECAT 951/31R scanner (Siemens/CTI Inc., Knoxville, TN, see [34, 35]).

Data collection on the IndyPET scanner is performed in a 'step and shoot' mode where multiple frames of data are collected with equal durations at uniformly spaced gantry angles. Because the detectors do not form a complete ring, the data collected for each frame form a diamond shaped sinogram segment as illustrated in Fig. 4. In cases where radioactive decay is significant, data can be collected using multiple gantry sweeps to avoid imaging different angular positions at significantly different activity levels.

2.1 Non-uniform Sinogram Spacing

The geometry of the IndyPET scanner results in projection measurements that are non-uniformly spaced in angle and displacement. Iterative reconstruction algorithms such as EM and Bayesian techniques can account for the non-uniform spacing, however, FBP requires a sinogram with a full set of evenly spaced



(a) Resolution (b) Sensitivity

Figure 3: Resolution and sensitivity of the IndyPET scanner. The plots in (a) show the resolution of the IndyPET scanner as a function of source offset from the center for different detector separations. The resolution was measured as geometrical mean of the radial and tangential FWHM measurements obtained from line source reconstructions using FBP with a ramp filter. The sinogram for the resolution measurement was generated as described in Section 2.1 using only the 12 central rows of the sinogram segments for each frame. The limit of 12 rows was chosen to minimize the angular error in the generation of a full sinogram to less than 0.01 degrees to characterize the capabilities of the scanner. The barplot in (b) shows the absolute sensitivity of the scanner as a function of detector separation in comparison to that of a CTI 951/31R. The sensitivity was obtained by calculating the percentage of detected decays for a 1 ml spherical source in the center of the field of view using all 56 rows of the sinogram segment. Note that the axial field of view of the CTI 951/31R scanner is 10.8 cm as compared to 5.0 cm for the IndyPET scanner.

angles and uniform displacements.

For the FBP reconstructions in this study, an approximate sinogram was generated as follows: First, the gantry rotation steps for all data sets were chosen to be integer multiples of the 1.0 degree angular span of a detector segment in the center of the detector banks. Specifically, the detector separation was set to 37.815 cm, and the data sets were acquired using 45 frames with gantry rotation steps of 4 degrees between frames. The sinogram segments were individually corrected for relative detector sensitivity, deadtime, and decay. The normalization for detector sensitivity was performed using measurements collected with a plane phantom centered in the field of view. A full sinogram was then obtained by summing the segments after they had been shifted by the appropriate number of rows to account for the gantry rotation. This procedure avoids interpolation which would smooth the sinogram data. The resulting sinogram is approximate because the center of curvature of the detector banks does not coincide with the center of gantry rotation. To limit the effect of this approximation, we only summed the central angular region (32 out of 56 rows) of each segment as illustrated in Fig. 4. For 32 rows, the maximum error in the angular position of the lines of response relative to a uniform angular spacing is 0.14 degrees as compared to the 1 degree angular span of one detector segment. In contrast, at 56 rows, the error in angular position is 0.9 degrees which would result in significant loss of resolution for FBP reconstruction.

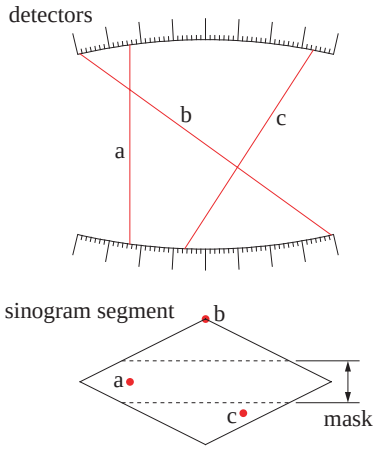


Figure 4: Diamond shaped sinogram segment acquired by the IndyPET scanner for a single gantry position. The points a, b, and c illustrate the positions of sample lines of response relative to the detectors and the sinogram. For FBP reconstruction, a full sinogram is generated by combining the central angular portions of the sinogram segments collected at different gantry positions.

Since the 32 central rows contain 81% of the total number of elements in each sinogram segment, we consider this limit an acceptable tradeoff between FBP resolution and sensitivity.

We emphasize that the procedure above was only used for FBP reconstruction. All iterative reconstructions used the original, uncorrected sinogram segments including all 56 rows. For iterative reconstruction, the projection angle and displacement for each element of each sinogram segment were calculated using the exact scanner geometry.

2.2 Modeling the System Kernel

Iterative reconstruction algorithms can achieve resolution recovery by incorporating an accurate system kernel that accounts for sinogram blurring due to the width of the detector segments and effects such as crystal penetration and inter-crystal scatter into the forward operator of the reconstruction algorithm. For the IndyPET scanner, we obtained an empirical system kernel using scans of line sources. We compare iterative reconstructions using the empirical kernel and the triangular kernel shown in Fig. 2.

2.2.1 Parameterization of System Kernel

Let x be the N-dimensional vector of emission rates of the pixels in raster order, and let y denote the vector of photon counts for all M projections, where for a single reconstruction plane $M = 45 \times 3136$, corresponding to 45 gantry positions, each with a sinogram segment containing 3136 measurements. Define P as the $M \times N$ projection matrix such that P_{ij} is proportional to the probability that an emission from

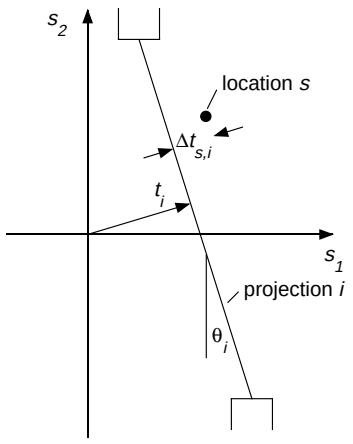


Figure 5: The detector kernel $h_i(s)$ is parameterized by the displacement difference $\Delta t_{s,i}$ between projection i and location s, $h_i(s) = \tilde{h}_i(\Delta t_{s,i})$.

pixel j is registered by the i^{th} detector pair. The expectation $E[y]$ is then given by

$$E[y] = Px \ . \tag{1}$$

Let $s = (s_1, s_2)$ denote the Cartesian coordinates in the field of view of the scanner and define $h_i(s)$ as the kernel of the i^{th} detector pair. Then P_{ij} can be computed as the 2-D integral

$$P_{ij} = \int_{s \in S_j} h_i(s) \ ds \tag{2}$$

where S_j is the set of coordinates s in pixel j. We now assume that pixels in the reconstruction image are considerably smaller than the detectors so that $h_i(s)$ is approximately constant over the support of a single pixel. Then

$$P_{ij} \approx h_i(s_{c_j}) A_j \tag{3}$$

where A_j is the area and s_{c_j} is the center coordinate of pixel j. For larger pixels, this assumption can be enforced by computing P for a sub-sampled pixel grid and summing the contributions for each pixel[6].

As illustrated in Fig. 5, define the ideal projection i as the line connecting the surface centers of the i^{th} detector pair. Let θ_i be the projection angle, and let t_i be the displacement of projection i from the coordinate origin. The displacement difference $\Delta t_{s,i}$ between an arbitrary location $s = (s_1, s_2)$ and projection i is $\Delta t_{s,i} = s_1 \cos(\theta_i) + s_2 \sin(\theta_i) - t_i$. We model the detector kernel $h_i(s)$ as having the form

$$\begin{aligned} h_i(s) &= \tilde{h}_i(\Delta t_{s,i}) & (4) \\ &= \tilde{h}_i(s_1 \cos(\theta_i) + s_2 \sin(\theta_i) - t_i) & (5) \end{aligned}$$

where $\tilde{h}_i$ is a 1-D function of $\Delta t_{s,i}$. This parameterization ignores the depth dependence of the detector kernel, i.e., the coordinate parallel to the projection line; however, this dependence could be added as an additional parameter.

Considering effects such as detector penetration, we expect the differences between the $\tilde{h}_i(\Delta t_{s,i})$ for different detector pairs i to be predominantly due to the incident photon angles of the different projections. For the IndyPET scanner, the detector banks are approximately planar, and it is a good approximation to assume that the incident photon angle is equal to the projection angle. Therefore, we assume that the detector kernel is only a function of the projection angle θ_i

$$h_i(s) = \tilde{h}_i(\Delta t_{s,i}) = h(\theta_i, \Delta t_{s,i}) \tag{6}$$

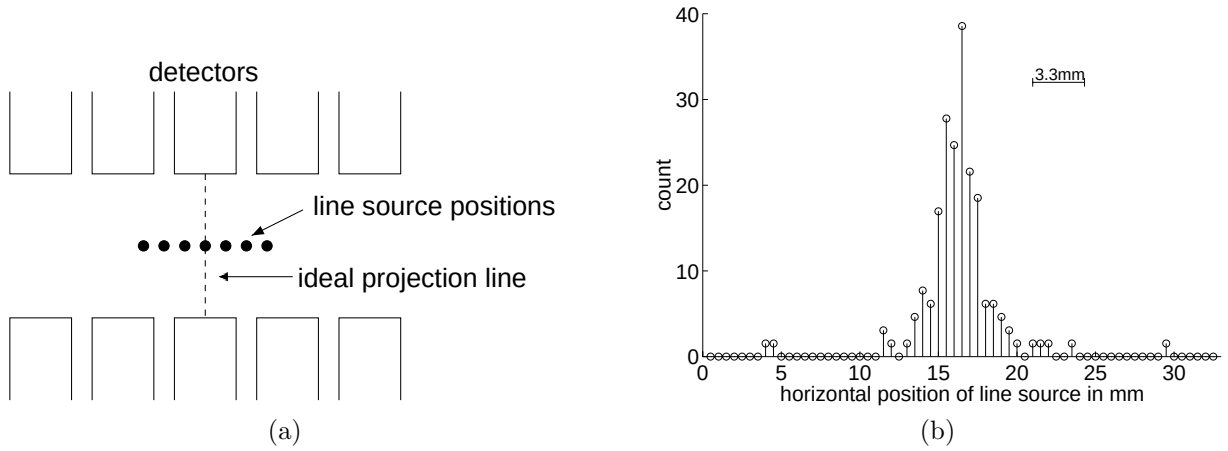
where we call the 2-D function h the *system kernel*. We note that our assumption of planar detector banks is used only to simplify the parameterization in (6), and the exact shape of the detector banks is used to determine the positions of lines of response in the measurement of the system kernel in Section 2.2.2 and for all iterative reconstructions. The validity of this assumption is confirmed by the weak dependence of the measured system kernel in Fig. 7 on projection angle. The parameterization in (6) does not account for the boundaries between different detector blocks, an effect that has typically been ignored[3, 6, 20].

While the parameterization in (6) is applicable to the geometry of the IndyPET scanner, we note that an analogous approach is applicable to a circular geometry. In that case, the parameterization can be expressed in terms of the displacement t of the lines of response, i.e. $h_i(s) = h(t, \Delta t_{s,i})$. This parameterization is possible because the angle of incidence is a function only of the displacement when depth dependence of the detector kernel is ignored, as assumed for the IndyPET geometry above. For a circular geometry, an empirical system kernel can be obtained as described in the next section by extending the measurements to the radius of the field of view.

2.2.2 Empirical Model Based on Line Source Scans

We measured the system kernel $h(\theta_i, \Delta t_{s,i})$ using scans of an axially oriented line source phantom, resulting in an approximate point source in the transaxial 2-D reconstruction plane. The line source was a stainless steel tube with an inside diameter of 0.5 mm filled with F-18 FDG. As illustrated in Fig. 6(a), the line source was mounted on a translation stage positioned near the center of the field of view and 65 data sets were collected as the source was moved horizontally in steps of 0.5 mm. Figure 6(b) shows the signal measured for a single detector pair as a function of the horizontal position of the line source.

The measured counts for all detector pairs and all needle positions were corrected for detector efficiency and parameterized by projection angle θ and displacement difference $\Delta t_{s,i}$ between the needle position



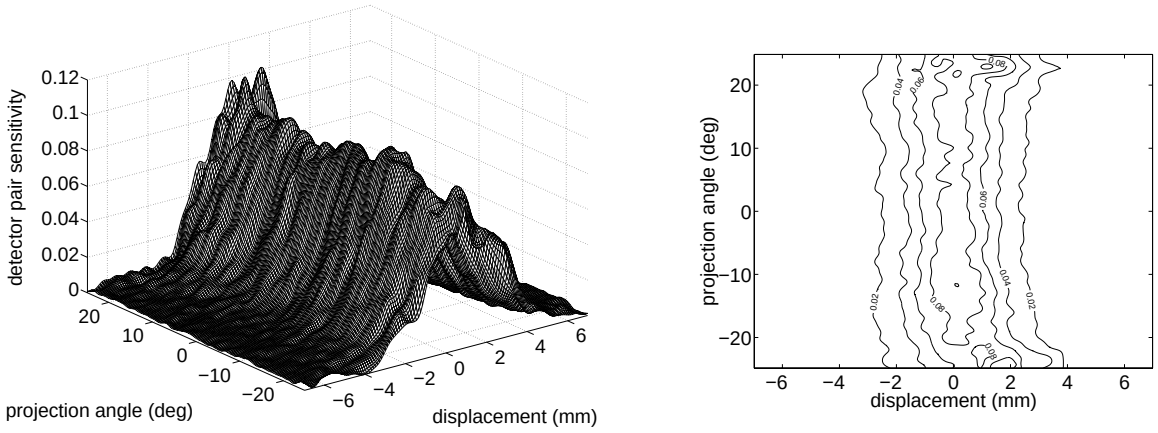
(a) (b)

Figure 6: (a) Illustration of the procedure used to determine the empirical system kernel using line source scans. (b) Signal measured by one detector pair as a function of line source position. The FWHM of the detector pair response is roughly equal to the detector crystal spacing indicated by the scale bar.

s and the ideal projection line. All counts were then accumulated onto a 180×200 point meshgrid with dimensions -24.87 deg $\leq \theta \leq -24.87$ deg and -10 mm $\leq \Delta t \leq 10$ mm using nearest neighbor assignment, followed by normalization of each grid value by the number of assigned measurements. Note that $\theta_{max} = 24.87$ deg is the maximum projection angle for the zero gantry position using our detector separation of 37.815 cm. To reduce noise, the gridded data were filtered with a 2-D Gaussian lowpass kernel with a standard deviation of 2 samples, corresponding to 0.2 mm in displacement and 0.5 degrees in angle. The result is an estimate for the system kernel $h(\theta, \Delta t)$.

A critical component in performing the procedure described above is to estimate precisely the physical positions s of the line source. While the horizontal position increments are accurately known by virtue of the translation stage, the initial source position s_0 is unknown. However, estimating $h(\theta, \Delta t)$ with a constant offset in the assumed source positions implies that reconstructions of data collected at a zero gantry angle will be shifted by the same offset. For reconstruction of data collected at multiple gantry angles, this results in circular blurring artifacts since projections at non-zero gantry angles α are modeled using the rotated kernel $h(\theta - \alpha, \Delta t)$. Thus, to achieve resolution recovery, it is critical to estimate the initial source position with an uncertainty much smaller than the width of a detector crystal segment.

Interestingly, the circular blurring effect can be exploited to iteratively estimate s_0 and $h(\theta, \Delta t)$ by using an approach similar to that proposed in [36] for center of rotation calibration in computed tomography (CT). Below, we summarize our approach, the details can be found in the Appendix. We first calculate an initial kernel $h(\theta, \Delta t)$ using an estimate of s_0 obtained from the sinogram. Given a second set of projection data containing a point source taken at different gantry angles, we use $h(\theta - \alpha, \Delta t)$



(a) (b)

Figure 7: Meshplot (a) and contour plot (b) showing the empirical system kernel $h(\theta, \Delta t)$ for the IndyPET scanner. The kernel is parameterized by projection angle θ and the displacement Δt of the source from the ideal projection line.

to separately reconstruct the projection data collected at each gantry angle. The positions of the point source in the reconstruction images for each gantry angle allow us to estimate the phase and radius of the circular blurring, which in turn give an improved estimate of s_0. Based on this improved estimate, we calculate a new $h(\theta, \Delta t)$ and iteratively repeat the calibration procedure until the radius of the circular blurring is below a specified threshold, in our case 0.57 mm. Note that this procedure calibrates the offset relative to the center of rotation of the gantry, independent of errors in the detector mounting.

Figure 7 shows our empirical system kernel for the IndyPET scanner estimated using the iterative procedure described above. The result does not indicate a significant dependence on projection angle. This effect is due to the limited range of projection angles acquired with the IndyPET scanner. Furthermore, our measurements become noisier towards larger angles since, due to the scanner geometry, the number of projections decreases with increasing angle, and fewer data points are available in this range. For iterative reconstruction, we truncate the system kernel to the range $|\Delta t| < 5$ mm.

We note that even though our line source measurements are performed only in the center of the field of view, these measurements capture the full range of incident angles for any detector pair and therefore provide a sufficient characterization for any line of response within the field of view (ignoring depth dependence). For reconstruction, this characterization of the lines of response is used throughout the field of view, taking into account the exact scanner geometry including the curvature of the detector banks.

3 Reconstruction Algorithms

We compare reconstructions obtained using FBP, EM, and Bayesian techniques with three different prior models. For EM and the Bayesian methods, we use the exact Poisson counting statistics to model the data acquisition. Given the notation of section 2.2.1, the shifted Poisson log-likelihood[37] is

$$\log p_{y|x}(y|x) = \sum_{i=1}^{M} \Big(-(P_{i*}x + s + 2r) \; + (y_i + 2r)\log(P_{i*}x + s + 2r) - \log((y_i + 2r)!) \Big) \tag{7}$$

where P_{i*} the vector formed by the i^{th} row of the projection matrix P and s and r denote contributions due to scatter and random coincidences respectively.

In general, scatter and randoms are a function of the projection index i; however, for this study, we assumed identical scatter and randoms values for all projections. We only performed a simplistic scatter correction by manually setting the scatter value s to the average sinogram activity near the edge of the imaged object. Randoms were not modeled for the reconstruction; the projection counts y_i were randoms pre-corrected by the scanner acquisition hardware and the parameter r was set to a small constant ($r \leq 1.5$, which in all cases was less than 2% of the average projection value) to compensate for negative sinogram entries y_i resulting from the pre-correction. Importantly, for a given set of projection data, the same values of s and r were used for all reconstruction algorithms to ensure a fair comparison.

The system matrix P was calculated from $h(\theta, \Delta t)$ using (6) and (3). In our implementation, we speed up computation by storing P. To reduce storage space, we only store the submatrix for the zero gantry position. The elements P_{ij} for non-zero gantry angles are obtained by rotating the coordinates of pixel j by the negative gantry angle and mapping the rotated coordinates into the pixel grid at the zero gantry position. Corrections for attenuation, deadtime, and decay were performed by multiplying the entries P_{ij} by the appropriate factors. The attenuation coefficients were calculated by forward projecting the estimated support of the imaged object using a constant attenuation coefficient. This is appropriate for our phantoms since the attenuation of acrylic and FDG are approximately equal.

3.1 Bayesian Reconstruction Algorithms

Bayesian reconstruction methods regularize the tomographic reconstruction by using a prior probability distribution $p(x)$ that incorporates assumptions about typical images such as local smoothness. The image reconstruction is computed as the approximate maximum *a posteriori* (MAP) estimate

$$\hat{x}_{MAP} = \arg\max_{x} (\log p(y|x) + \log p(x)) \; . \tag{8}$$

We compare three Bayesian techniques using a Gaussian Markov random field prior (GMRF), a generalized Gaussian Markov random field (GGMRF) prior, and an adaptive wavelet graph model (WGM) prior respectively. The GMRF prior can be written as

$$\log p(x) = -\frac{1}{2\sigma^2} \sum_{\{k,j\}\in\mathcal{N}} b_{k-j}(x_k - x_j)^2 + \text{constant} \tag{9}$$

where σ is a smoothing parameter and $\mathcal{N}$ is the set of all neighboring pixel pairs in the image x. The b_{k-j} are the 8-point MRF neighborhood coefficients with $b_{k-j} = \sqrt{2}/(4(\sqrt{2}+1))$ for horizontal and vertical neighbors, and $b_{k-j} = 1/(4(\sqrt{2}+1))$ for diagonal neighbors.

A disadvantage of the GMRF is that its squared potential function tends to oversmooth edges. To address this problem, several edge-preserving potential functions have been proposed[15, 7, 8, 16, 17, 18, 19]. One such model is the GGMRF[18] which can be written as

$$\log p(x) = -\frac{1}{p\sigma^p} \sum_{\{k,j\}\in\mathcal{N}} b_{k-j}(x_k - x_j)^p + \text{constant} \tag{10}$$

where p is a shape parameter. We chose $p = 1.5$ because we have found that this significantly enhances edge sharpness without producing the segmented appearance that can result with smaller values of p.

The adaptive wavelet graph model (WGM)[30, 31] is a nonlinear prior model formulated in the wavelet domain. The WGM exploits dependencies between wavelet coefficients at different resolutions to adapt the degree of regularization to the local image statistics. This allows for strong regularization of uniform image regions while edges and textures are preserved. Due to the adaptation, the prior is data dependent and strictly speaking, does not result in a MAP reconstruction, however, for simplicity we refer to it as a MAP technique. The WGM requires a set of training data to estimate the model parameters. For this study, the WGM was trained on the same set of MRI reconstructions as previously described[30, 31]. Since the training data serves to provide information about the interpolation of textures and edges, it is not necessary to specifically train the model on PET imagery.

The MAP optimization for all three prior models was performed using the iterative coordinate descent (ICD) algorithm[18, 14]. The ICD algorithm sequentially optimizes the MAP equation (8) with respect to single pixel values. In comparison to gradient methods[38, 10, 12, 39], the ICD algorithm allows for simple enforcement of pixel positivity, resulting in more robust convergence. However, a disadvantage of ICD is that the computational complexity cannot be reduced by using a factored matrix model as in [12, 3, 20].

While previous ICD implementations rely on FBP initialization, in this work, we propose a simple multiresolution initialization strategy. Starting at a reconstruction resolution of 16×16 pixels with constant initialization, we compute a sequence of coarse-to-fine scale reconstructions where each reconstruction is

initialized with the interpolated result from the next coarser scale. This strategy eliminates the need for a preliminary FBP reconstruction while providing good computational efficiency. Specifically, the coarse scale solutions are fast to compute and provide a good initialization, thereby reducing the number of computationally expensive iterations required at the finest scale. Note that the assumption of constant system kernel per pixel (3) does not hold at coarse scales, however, since the coarse scale solutions are only used for initialization, the final fine scale reconstructions are not affected. The choice of MRF smoothing parameters at coarse scales is not critical, specifically, we have used $\sigma_n = 2^{-n}\sigma_0$ where σ_n denotes the smoothing parameter at scale n and $n = 0$ denotes the finest scale. For all reconstructions, 25 iterations were performed at each scale to ensure sufficient convergence. Note, that the convergence could be measured by using a stopping criterion based on the change in successive values of the log-posterior at each scale. However, since for MAP reconstruction in general, very small changes of the log-posterior can result in large visual differences, we prefer to run a fixed number of iterations at each scale, chosen sufficiently large such that no more visible changes occur in the reconstruction. Note, that for the optimization of the MAP equation (8), there is in general no closed-form update equation at each iteration similar to (11) for EM. The detailed ICD optimization algorithm is described in [18, 14].

3.2 Expectation Maximization Reconstruction Algorithm

The EM algorithm used in this study is the basic algorithm[4] modified as in [20] to include corrections for scatter and random coincidences. The EM update equation used was

$$x_j^{n+1} = \frac{x_j^n}{\sum_{i=1}^{M} P_{ij}} \sum_{i=1}^{M} P_{ij} \frac{y_i + 2r}{2r + s + \sum_{k=1}^{N} P_{ik} x_k^n} \tag{11}$$

where x^n is the reconstruction estimate after the n^{th} iteration. The smoothness of the reconstructions was controlled by the number of iterations performed. For all reconstructions, the EM algorithm was initialized with a constant image whose forward projection matched the total number of counts in the measured projection data. To ensure a fair comparison, our implementation of EM was based on the same programming code as the Bayesian ICD-MAP algorithm, and thus, both methods used exactly the same system matrix computation and corrections.

3.3 Filtered Backprojection

For filtered backprojection reconstruction, an approximate complete sinogram was generated as described in section 2.1. The frequency response of the FBP reconstruction filter was

$$H(f) = H_{ramp}(f) \left(0.5 + 0.5 \cos\left(\frac{\pi f}{\alpha f_N} \right) \right) \quad \text{for} \quad |f| \leq \min(\alpha f_N, f_N) \tag{12}$$

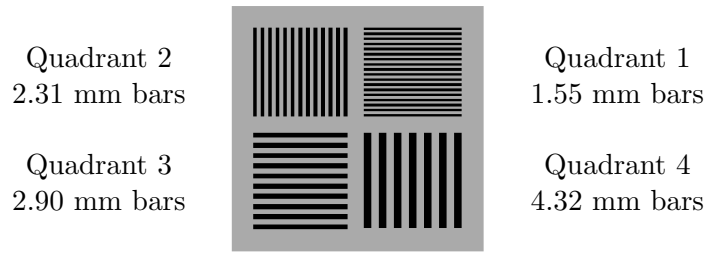


Figure 8: Illustration of the bar phantom. The phantom consists of four arrays of parallel bars with thicknesses of 1.55, 2.31, 2.90, and 4.32 mm separated by gaps of equal widths.

and zero elsewhere, where α is a smoothing parameter and f_N is the Nyquist frequency. For $\alpha = 1$, we obtain a Hanning filter, and for $\alpha > 1$, the Hanning response transitions to a ramp filter. For the results presented, we varied α in the range from 0.3 to 8.0 and used a ramp filter for minimal smoothing.

4 Experimental Results

Phantom data acquired on the IndyPET scanner as well as simulated data were used to evaluate reconstruction quality for different reconstruction algorithms and for different system kernels. Three data sets were used, a bar phantom, a Hoffman 3-D brain phantom, and a simulated data set containing a circular lesion on an approximately uniform background. Fluorine-18 in the form of FDG was used as the radioisotope for the phantom studies. Quantitative analyses of reconstruction quality for FBP, EM, and the Bayesian techniques were performed for the bar phantom and the simulated data set. For the bar phantom, we also compare reconstruction quality for the empirical and the triangular system kernels, and we evaluate reconstruction quality for a set of parametric system kernels with different widths. The triangular and parameterized kernels are modeled as a function of $\Delta t_{s,i}$ only, whereas the empirical kernel includes the dependence on θ_i as shown in Fig. 7. Finally, we compare the computational efficiency of different initialization strategies for the ICD-MAP algorithm.

To ensure fair comparison of the images presented for each data set, all system kernels used in the iterative algorithms were normalized for the same area and the same pixel intensity scale is used to display the results. The FBP images were scaled such that their summed intensity was equal to that of the Bayesian GMRF MAP reconstructions.

4.1 Bar Phantom

As illustrated in Fig. 8, the bar phantom consists of four arrays of parallel acrylic bars with thicknesses of 1.55, 2.31, 2.90, and 4.32 mm separated by gaps of equal widths, resulting in square wave emission profiles

with different periods for each quadrant. Fourteen different data sets were collected sequentially on the IndyPET scanner, each corresponding to a full 180 degree gantry sweep using 45 frames. The acquisition time per frame was identical for all 14 data sets. By summing different subsets of these data sets frame by frame, we generated 4 combined data sets with a total of 10.4M, 4.4M, 2M, and 586,000 counts in the reconstructed plane. For these combined data sets, the randoms rates reported by the delayed window estimation were 17.7%, 9.0%, 4.6%, and 2.6% respectively. The average deadtime fractions per detector pair were 31%, 16%, 9%, and 5% respectively.

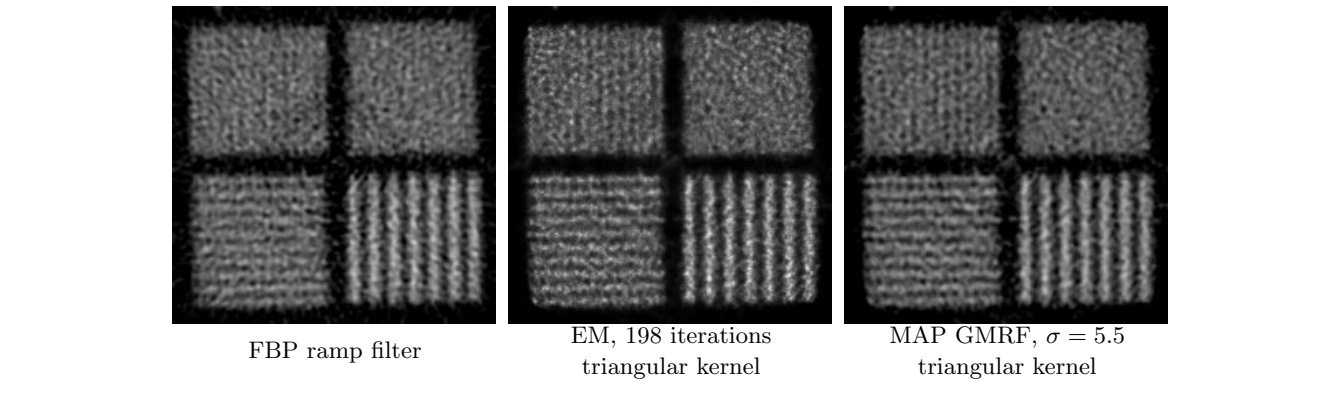
FBP, EM, and Bayesian reconstructions of the four combined data sets were computed using a wide range of smoothing parameters for each algorithm. The reconstruction resolution was 256×256 pixels with a field of view of 18 cm, resulting in a pixel size of 0.7 mm. For EM and Bayesian MAP with a GMRF prior, we compare reconstructions for both the empirical system kernel shown in Fig. 7 and the triangular kernel shown in Fig. 2. For Bayesian GGMRF MAP and the Bayesian wavelet graph model (WGM), only the empirical system kernel was used. In a separate experiment described in section 4.1.2, we compare Bayesian GMRF reconstructions for a set of parametric system kernels of different widths.

Figures 9 and 10 show sample reconstructions for the data set with 10.4M counts. The images were cropped to 192×196 pixels, corresponding to a physical size of 13.50×13.78 cm. The smoothing parameters for the reconstructions shown in Fig. 9 were chosen to minimize the expected spatial reconstruction error (bias^2+variance, see section 4.1.1) for the 4.32 mm bars in quadrant 4. Similarly, Fig. 10 shows reconstructions with minimum error for the 2.90 mm bars in quadrant 3. The images of the EM reconstructions are slightly clipped due to the high amplitude of the noise peaks in the reconstructions.

Based on visual inspection, the empirical system kernel significantly improves the quality of the iterative reconstruction techniques. Also, when the empirical system kernel is used, the iterative techniques provide improved resolution over FBP. For the Bayesian reconstructions, the non-Gaussian GGMRF and WGM prior models provide sharper edges as compared to the GMRF prior. However, the GMRF prior appears to have a wider frequency response, allowing for good reconstruction quality of all quadrants using a single regularization parameter. Note that none of the reconstruction methods can resolve the horizontal 1.55 mm bars in quadrant 1, thus our quantitative analysis below is limited to quadrants 2-4.

4.1.1 Bar Phantom Bias and Variance Analysis

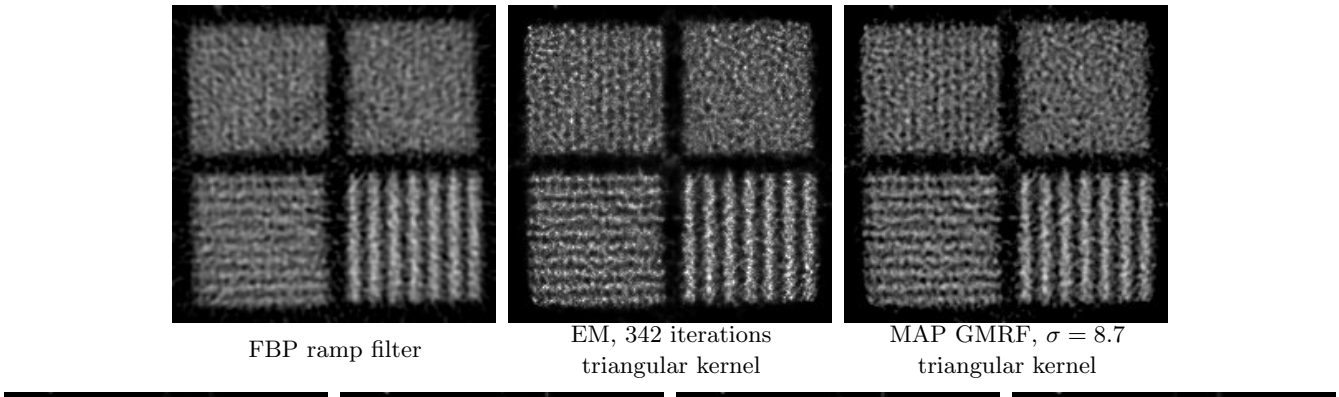
The bar phantom reconstructions were analyzed by comparing the reconstructed signal to the expected square wave profile. Bias and variance measures were computed separately for each quadrant by fitting harmonic series to the reconstructed bar profiles. Let $x_Q(r, n)$ denote the reconstruction values in a 50



FBP ramp filter | EM, 198 iterations triangular kernel | MAP GMRF, $\sigma = 5.5$ triangular kernel

EM, 348 iterations empirical kernel | MAP GMRF, $\sigma = 6.9$ empirical kernel | MAP GGMRF, $\sigma = 6.9$ empirical kernel | Bayesian WGM, $\omega = 2.5$ empirical kernel

Figure 9: Bar phantom reconstructions for 10.4M total counts. The images correspond to a physical size of 13.50×13.78 cm with a pixel size of 0.70 mm. The smoothing parameter for each method was selected to minimize the expected error for the bar profile in quadrant 4 (4.32 mm bars).



FBP ramp filter | EM, 342 iterations triangular kernel | MAP GMRF, $\sigma = 8.7$ triangular kernel

EM, 1068 iterations empirical kernel | MAP GMRF, $\sigma = 13.0$ empirical kernel | MAP GGMRF, $\sigma = 15.0$ empirical kernel | Bayesian WGM, $\omega = 6.0$ empirical kernel

Figure 10: Bar phantom reconstructions for 10.4M total counts with smoothing parameters selected to minimize the expected error for the bar profiles in quadrant 3 (2.90 mm bars).

by 50 pixel window centered over quadrant Q. Let n denote the coordinate running perpendicular to the bars (e.g., the horizontal coordinate in quadrants 2 and 4) and let r be the coordinate parallel to the bars. The 1-D average reconstruction bar profile $\bar{x}_Q(n)$ is obtained as

$$\bar{x}_Q(n) = \frac{1}{R}\sum_{r=1}^{R} x_Q(n,r) \tag{13}$$

where $R = 50$ is the size of the window and $1 \leq n \leq R$. We then determine the coefficients a_k and b_k of a harmonic series $\tilde{x}_Q(n)$

$$\tilde{x}_Q(n) = b_0 + \sum_{k=1}^{K}(a_k \sin(2\pi k f_Q n) + b_k \cos(2\pi k f_Q n)) \tag{14}$$

by using a least squares optimization to fit $\tilde{x}_Q(n)$ to $\bar{x}_Q(n)$. The frequency f_Q is a fixed frequency appropriate for each quadrant. The order of the fitted series was $K = 7$.

The reconstruction bias is calculated by comparing the fit $\tilde{x}_Q(n)$ to an ideal square wave. Let α_k and β_k be the Fourier coefficients of an ideal square wave with unit amplitude, a DC value of one, and a phase equal to the phase of the fundamental of $\tilde{x}_Q(n)$. We then calculate the squared reconstruction bias for quadrant Q as

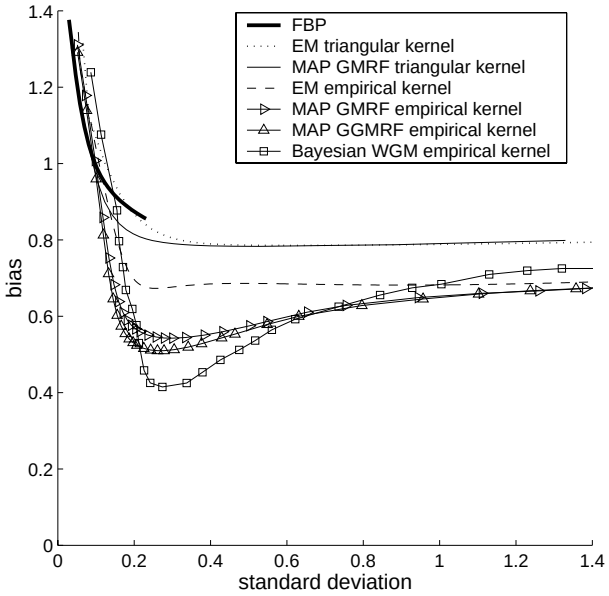
$$\text{bias}_Q^2 = \min_c \left\{ \sum_{k=1}^{K} \left(\frac{a_k}{c} - \alpha_k \right)^2 + \sum_{k=0}^{K} \left(\frac{b_k}{c} - \beta_k \right)^2 \right\} . \tag{15}$$

The minimization over c accounts for the fact that the absolute scaling of the reconstruction algorithms is not calibrated. Thus, we scale the reconstructed signal to the ideal square wave such that bias is minimized. The minimization in (15) is performed analytically. The bias is affected by the choice of the scatter and randoms parameters r and s in (7) used for the reconstruction. Specifically, a smaller value for s results in increased bias due to an increased DC offset in the reconstructed wave profile. Thus, we emphasize that for a given set of projection data, we use the same scatter and randoms values for all reconstruction methods to ensure a fair comparison.

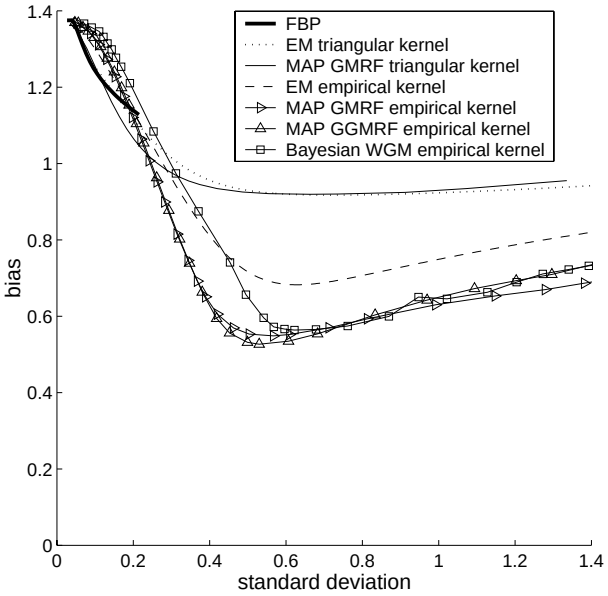
While the bias characterizes the systematic deviation of the reconstruction from the true signal, the variance characterizes the random variation of the reconstruction around its average. The variance for quadrant Q is calculated as the combined variance over all rows r around the fit $\tilde{x}_Q(n)$

$$\text{var}_Q = \text{std}_Q^2 = \frac{1}{c^2(R^2-1)} \sum_{r=1}^{R}\sum_{n=1}^{R} (x_Q(n,r) - \tilde{x}_Q(n))^2 \tag{16}$$

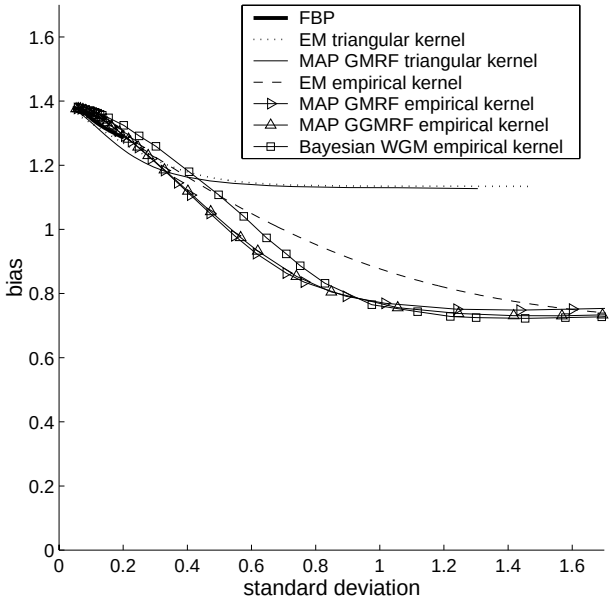
where c is the scaling constant determined in (15). Note that the variance defined here is the spatial variance characterizing the noisiness of a single reconstruction as compared to the variance of the reconstruction estimator.



(a) quadrant 4, 4.32 mm bars



(b) quadrant 3, 2.90 mm bars



(c) quadrant 2, 2.31 mm bars

Figure 11: Plots of bias versus standard deviation for the bar phantom reconstructions of the data set containing 10.4M total counts. Bias and standard deviation were calculated using the Fourier analysis described in section 4.1. The curve for each reconstruction algorithm was obtained by varying the smoothing parameter.

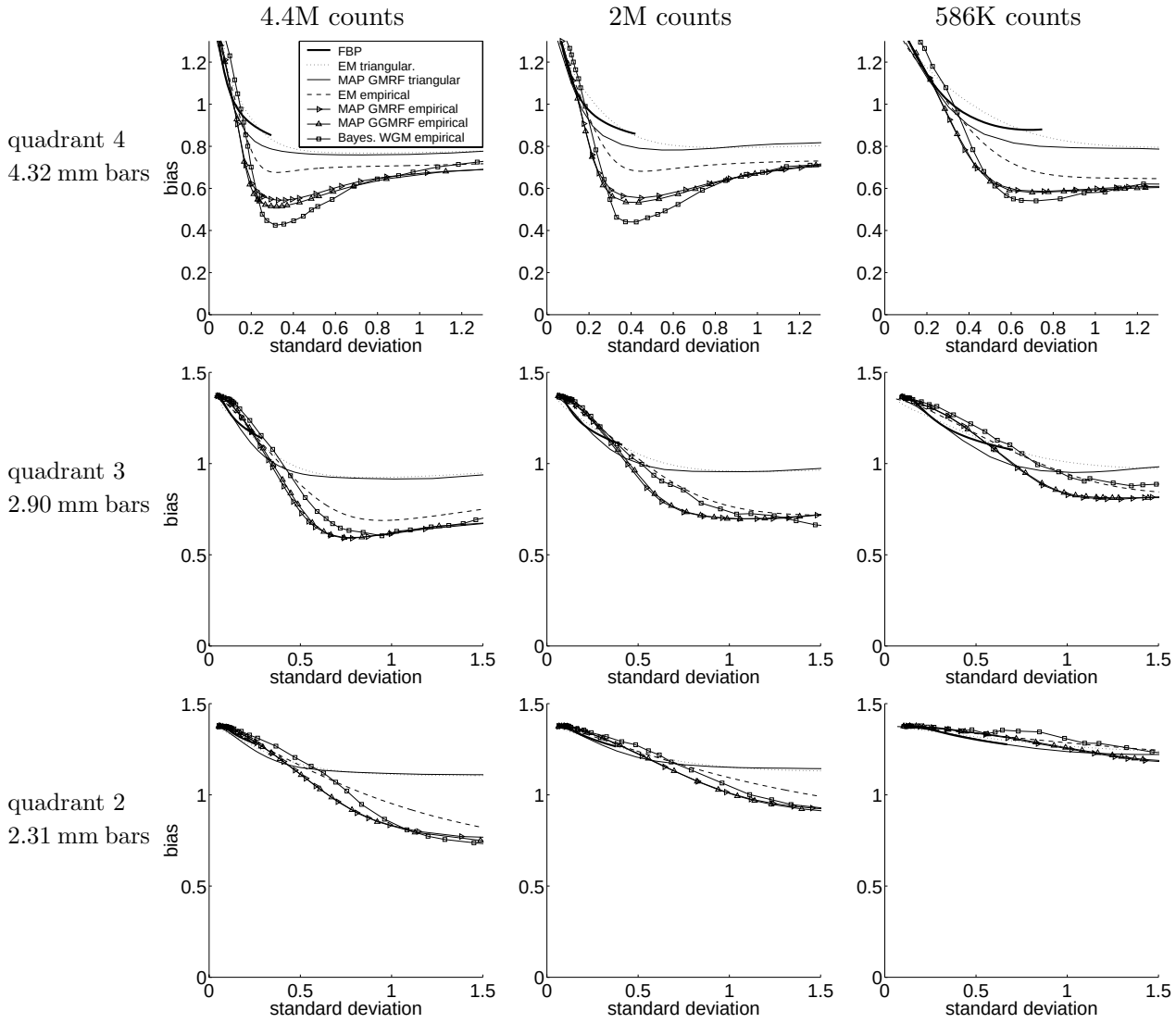


Figure 12: Plots of bias versus standard deviation for reconstructions of the bar phantom data. The three rows correspond to quadrants 4, 3, and 2 with 4.32, 2.90, and 2.31 mm bars, respectively. The three columns correspond to data sets containing 4.4M, 2.0M, and 586K counts, respectively.

Figure 11 shows reconstruction bias versus standard deviation for the 10.4M count case separately for the three quadrants. The curves were obtained by varying the smoothing parameter for each reconstruction method. Note that the points closest to the coordinate origin in Fig. 11(a) and (b) correspond to the reconstruction images shown in Fig. 9 and Fig. 10 respectively.

The bias versus standard deviation plots indicate that using the empirical system kernel as compared to the triangular kernel significantly lowers the bias at equal variance for EM and Bayesian GMRF MAP. Interestingly, when the triangular kernel is used, EM and GMRF MAP perform similarly. However, when the empirical kernel is used, GMRF MAP achieves significantly lower bias at equal variance as compared to EM. For the empirical kernel, the GGMRF prior performs slightly better than the GMRF model. The

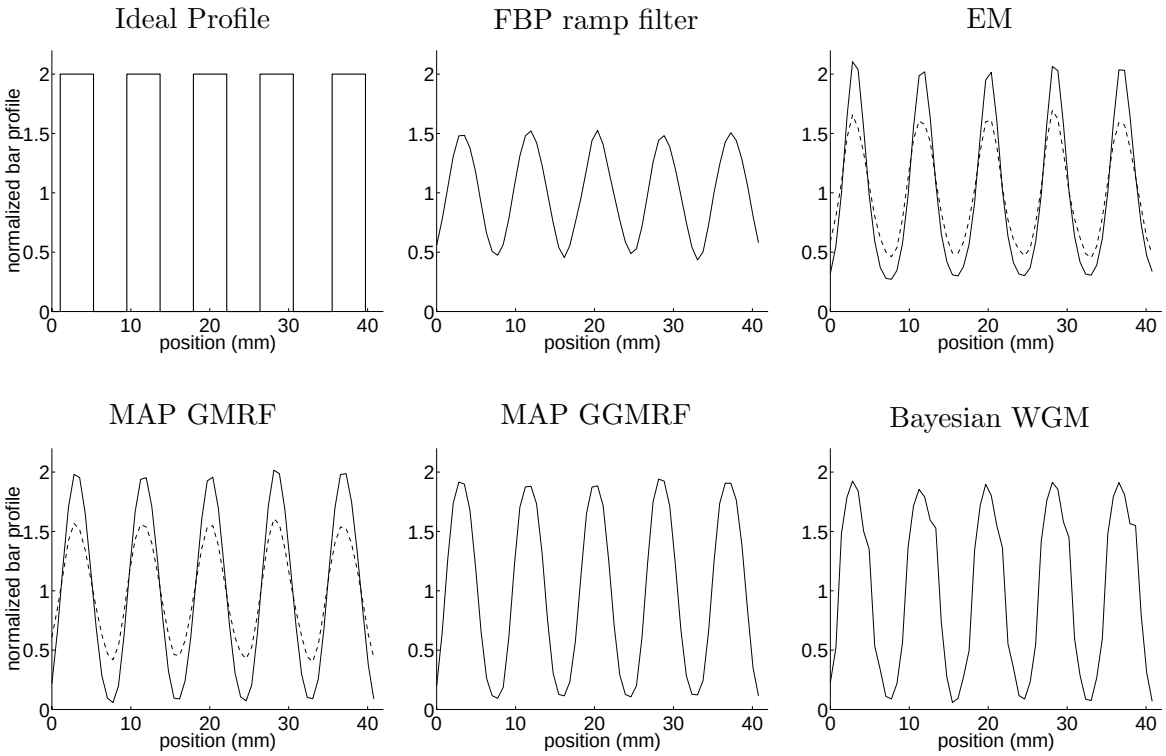


Figure 13: Bar profile reconstruction averages $\bar{x}_4(n)$ for quadrant 4 (4.32 mm bars). The profiles were calculated from the reconstruction images shown in Fig. 9, i.e., for smoothing parameters minimizing the error in quadrant 4. The dashed lines for EM and MAP GMRF correspond to reconstructions using the triangular system kernel.

WGM model performs well in quadrant 4 but not as well in quadrants 2 and 3. The FBP reconstructions have low variance and high bias even for a ramp filter since no resolution recovery is performed. FBP only partly resolves the bars in quadrant 3 and does not resolve quadrant 2. For quadrant 4, the FBP bias versus standard deviation is similar to the low variance range of the curves for the iterative algorithms using a triangular system kernel. However, when the empirical system kernel is used, the iterative techniques achieve significantly lower bias at equal variance as compared to FBP.

Figure 12 shows the bias versus standard deviation curves for the bar phantom data sets with 4.4M, 2M, and 586K total counts. The results are consistent with those for the 10.4M count data set. Due to the lower signal to noise ratio at lower count levels, the advantage of using the empirical system kernel decreases slightly. This is expected, as the potential for resolution recovery is best at high signal to noise ratios. For the 586K data set, the bars in quadrant 2 cannot be resolved in the reconstructed images (not shown) and the bias versus standard deviation curves are essentially flat for all methods.

For an ideal system, or reconstruction of simulated projection data, we would expect the reconstruction bias to approach zero as the standard deviation increases, i.e., as less smoothing is performed. Since our experiments are based on phantom data, our measurements have limited accuracy due to partial volume

effects and limited accuracies of the system kernel and estimated ground truth. These systematic errors result in a non-zero bias even when no smoothing is performed in the reconstruction. Moreover, most of the curves in Figs. 11 and 12 indicate that after reaching a minimum, the bias increases for larger values of the standard deviation. We attribute the increase to two factors. First, our reconstructions have slight circular artifacts resulting from an imperfect normalization scan that was applied to the data collected at each gantry position. Thus, the reconstruction profile deviates systematically from an ideal square wave as less smoothing is performed. Secondly, the bias is calculated using a limited window of 50 by 50 pixels for each quadrant which may result in noisy estimates of the higher order Fourier coefficients. As expected, the bias curves for EM and the Bayesian methods approach each other for reconstructions performed with the same system kernel.

Figure 13 shows the reconstructed bar profiles in quadrant 4 for the reconstruction images shown in Fig. 9. The plotted profiles are the averages $\bar{x}_4(n)$ as defined in (13) normalized by their DC value. For EM and GMRF MAP, the dashed lines correspond to the reconstructions using the triangular system kernel. Again, the iterative reconstructions using the empirical system kernel achieve much larger values of the amplitude to DC ratio. For the EM results, using the empirical kernel produces a slightly asymmetric profile with peaks overshooting the ideal profile. The GMRF MAP profile is symmetric but has very smooth transitions, indicating blurred edges. The GGMRF MAP method produces slightly sharper transitions. The WGM prior results in much sharper transitions but has some artifacts in the regions of high amplitude. We note that in general, the reduction in bias achieved by the empirical system kernel as shown in Figs. 11 and 12 is largely due to the increased amplitude of the reconstructed bar profiles.

4.1.2 Comparison of Parametric System Kernels of Different Widths

To evaluate the impact of systematic changes in the width of the system kernel, we compare Bayesian GMRF reconstructions of the bar phantom for a set of parametric kernels. The parametric kernels were obtained by convolving the triangular kernel of Fig. 2 with Gaussians of different standard deviations β. To obtain a baseline 1-D kernel that is similar to the 2-D empirical system kernel, we first computed an average of the empirical kernel in Fig. 7 over the 80% central portion of the angular range. The β value for the baseline parametric kernel was then estimated to obtain a least-squares fit to the averaged empirical kernel. For the reconstructions, we consider parametric kernels with parameters from $\beta = 0$ (triangular kernel) to $\beta = 2\beta_f$, where β_f is the fitted value. Figure 14 shows examples for $\beta = 0.5\,\beta_f$ and $\beta = 1.5\,\beta_f$. All parametric kernels were normalized such that the area under the curve was equal to that of the 1-D averaged empirical kernel.

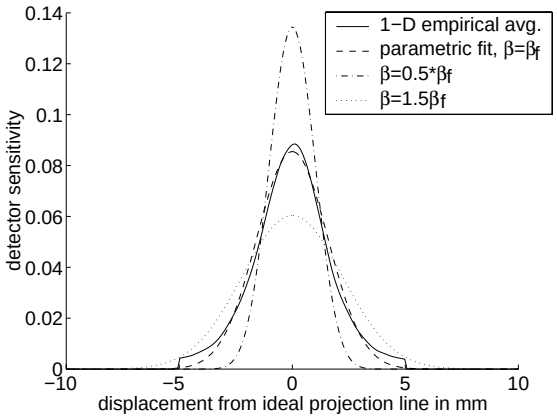
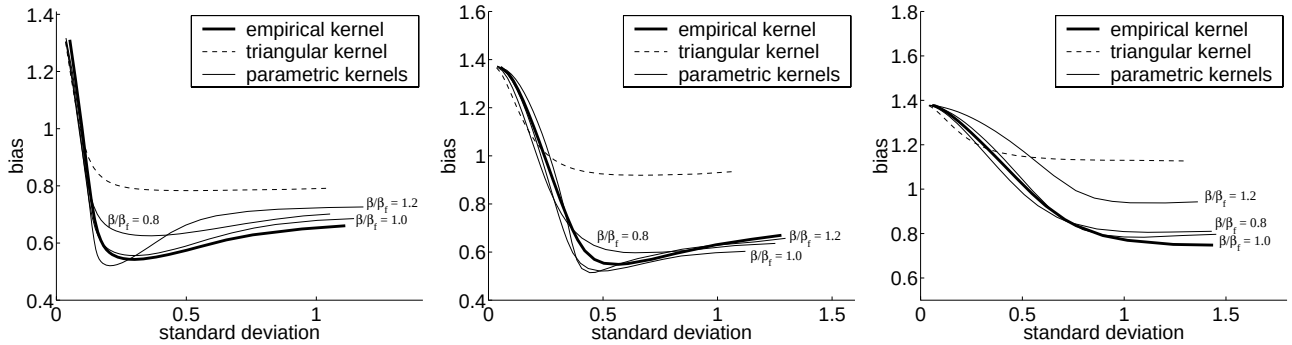


Figure 14: Examples of system kernels used to evaluate reconstruction quality as a function of kernel width. The empirical 2-D system kernel $h(\theta, \Delta t)$ was averaged in the angular coordinate θ to obtain a 1-D average kernel (solid line). A parametric fit (dashed line) to the averaged empirical kernel was obtained by convolving the triangular kernel with a Gaussian. Parametric kernels of different widths are then generated by varying the standard deviation of the convolving Gaussian relative to the value obtained for the fit. The 1-D kernels are normalized for equal area.



Quadrant 4 Quadrant 3 Quadrant 2

Figure 15: Plots of bias versus standard deviation for Bayesian GMRF reconstructions of the bar phantom data (10.4M counts) for different system kernels. The β/β_f values indicate the width of the parametric kernel where β_f is the value obtained by fitting the 1-D averaged empirical kernel.

Figure 15 compares the bias versus standard deviation curves of reconstructions using parametric kernels of different widths and the original 2-D empirical kernel shown in Fig. 7. We observed that some improvements in reconstruction quality can be achieved by increasing the width of the parametric kernel as compared to the triangular kernel ($\beta = 0$). However, the result for $\beta = \beta_f$ deviates noticeably from that obtained using the original 2-D empirical system kernel. This difference indicates that small deviations in the system kernel can significantly affect reconstruction quality. For very high values of the reconstruction standard deviation, i.e., as the influence of the Bayesian prior model is minimal, the empirical system kernel results in the lowest bias for quadrants 2 and 4. This effect may indicate that the empirical kernel more accurately describes the scanner as compared to the parametric models.

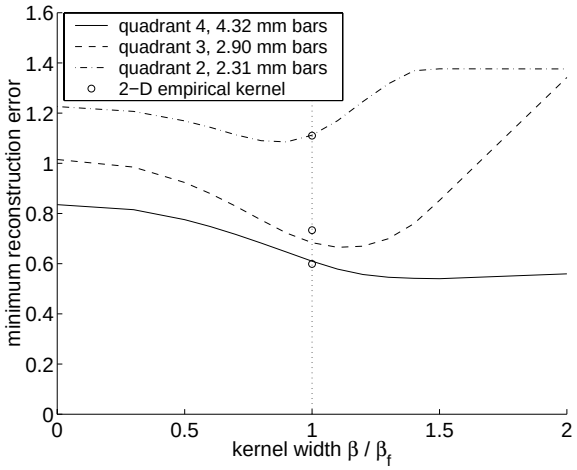


Figure 16: Plots of minimum value, taken over all smoothing parameters σ, of the reconstruction error $\sqrt{(\text{bias}^2 + \text{variance})}$ as a function of the width of the parametric system kernel. The discrete points correspond to the minimum reconstruction error using the 2-D empirical system kernel. Data are shown for Bayesian GMRF reconstructions of the bar phantom data with 10.4M counts.

Figure 16 shows the minimum value, taken over all smoothing parameters σ, of the reconstruction error $\sqrt{(\text{bias}^2 + \text{variance})}$ as a function of the width of the parametric system kernels. The results indicate that the fitted width $\beta = \beta_f$ provides a good compromise for all quadrants. Using a wider system kernel increases the error in quadrant 2 but improves the result for quadrants 3 and 4. However, as can be seen from the curve for $\beta = 1.2\,\beta_f$ in Fig.15(a), a wider kernel improves the result only for a very small range of standard deviations. This effect may be due to interactions with the GMRF prior, i.e., over-sharpening the reconstruction through the forward model may counteract the blurring of the GMRF prior.

4.2 Hoffman 3-D Brain Phantom

The second data set used in this study was acquired on the IndyPET scanner using a Hoffman 3-D brain phantom. A single plane of the data set containing 2M counts was used for the reconstruction. The randoms rate and deadtime per detector pair were 13% and 14% respectively. The reconstruction was performed at a resolution of 256×256 pixels with a field of view of 20 cm, resulting in square pixels of size 0.78 mm. For EM and Bayesian GMRF MAP, we compare the empirical and the triangular system kernels. Only the empirical system kernel was used for the other Bayesian techniques.

Figures 17 and 18 show reconstructions that have been cropped to 177×148 pixels, corresponding to a physical size of 13.8×11.5 cm. For each row, the figures correspond to a single reconstruction algorithm with three manually selected smoothing parameters. Note that due to the planar structure of the phantom, the reconstructions correspond to overlapping physical planes in locations where the shapes

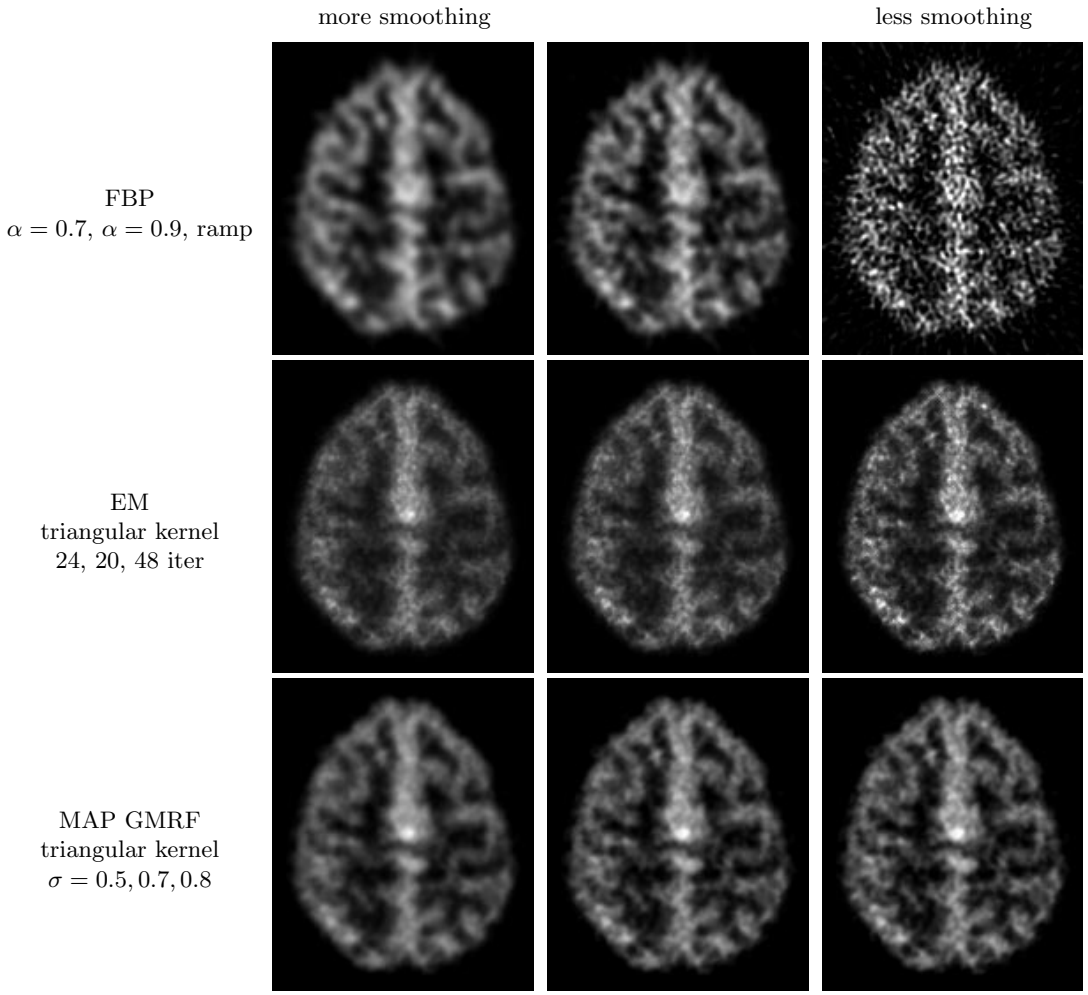


Figure 17: Reconstructions of the Hoffman 3-D brain phantom using FBP and iterative techniques with the triangular system kernel. The three images in each row represent reconstructions from a single algorithm with three manually selected degrees of smoothing. For comparison purposes, Fig. 18 shows reconstructions obatined using the empirical system kernel.

of the adjacent planes in the phantom differ. The reconstructions in Fig. 17 correspond to FBP and iterative algorithms using the triangular system kernel. Figure 18 shows EM and Bayesian reconstructions using the empirical system kernel. We observe that using the empirical kernel improves the quality of the iterative reconstructions. The EM reconstructions deteriorate with increasing numbers of iterations before some of the detail is converged and do not reproduce the interior contours as well as the Bayesian techniques. As compared to the GMRF, the GGMRF prior produces sharper contours with few artifacts. The WGM produces a much sharper image, however, the result has significant artifacts.

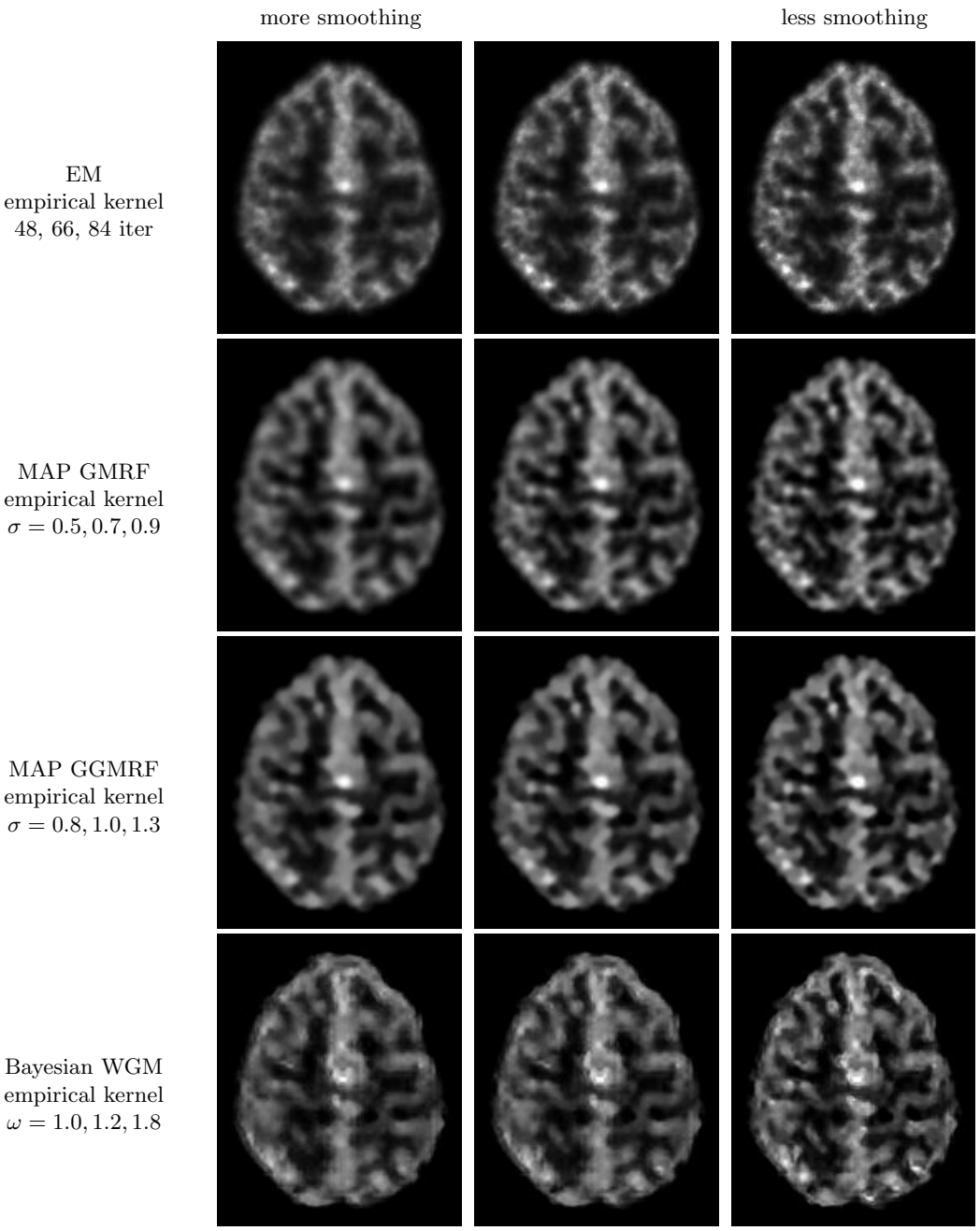


Figure 18: Reconstructions of the Hoffman 3-D brain phantom using the EM algorithm and Bayesian methods using the empirical system kernel.

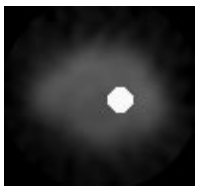


Figure 19: Simulated lesion phantom obtained by adding a circular lesion to an approximately uniform background. The lesion has a diameter of 4 mm and an intensity equal to 5 times the average background level.

4.3 Simulated Phantom

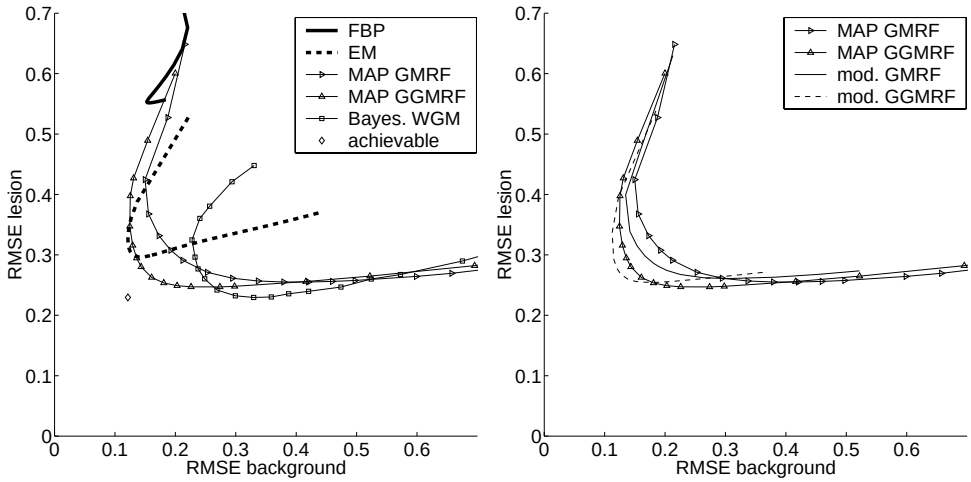
The third data set used in this study consists of simulated data generated by adding a constant circular lesion of 4 mm diameter and 5 times background activity to an approximately uniform background as shown in Fig. 19. The background image was an FBP reconstruction of an FDG mouse study on the IndyPET scanner. Simulated projection data were generated by forward projecting the image in Fig. 19 using the empirical system kernel for the IndyPET scanner. The data samples were formed by Poisson random variables with the appropriate means. Deadtime, decay, and relative detector sensitivity were not modeled. Forward projection and reconstruction were performed at a resolution of 256×256 pixels with a field of view of 8 cm. This results in a pixel size of 0.31 mm square which is about one tenth of the detector resolution. The total number of counts in the reconstruction plane was 130,000. Twenty separate noisy realizations of projection data were generated.

For reconstruction, only the empirical system kernel was used for EM and the Bayesian techniques. Even though the use of the same system kernel for forward projection and reconstruction will bias the results towards the iterative reconstruction techniques[40], we still obtain fair comparisons between different iterative techniques and obtain an approximate idea about their performance relative to FBP.

The reconstructions were evaluated quantitatively in terms of root mean square error (RMSE) estimator performance as a function of smoothing parameter. We do not use the clinically more relevant contrast recovery ratio (CRC) since we have found the CRC values and the relative performance of the reconstruction algorithms to be a strong function of the exact procedure used to estimate the ROI activities for the CRC calculation, an observation that is consistent with earlier reports[1]. Thus, for clinical evaluation, we consider it necessary to specify precisely a task with an associated ROI estimator or to perform human observer studies[41]. For this study, we limit ourselves to MSE analysis to decouple the reconstruction problem from the ROI estimation problem.

Let x be the ground truth image and let $\hat{x}$ be a single reconstruction. Define L and B as the known sets of lesion pixels and background pixels. The set B was obtained by thresholding the ground truth image, eroding the resulting mask, and excluding the lesion plus a 1 mm thick annulus around it. The set L was obtained as the non-eroded lesion region as added to the background image. Let μ_L and μ_B be the true mean activity values in the lesion and background. We then calculate the relative MSE errors e_L^2 and e_B^2 for the lesion and background regions as

$$e_L^2 \;=\; \frac{1}{\mu_L^2 N_L} \sum_{j \in L} (\hat{x}_j - x_j)^2 \tag{17}$$



(a) (b)

Figure 20: Plots of lesion RMSE versus background RMSE for reconstructions of the simulated lesion phantom. The curves were obtained by varying the smoothing parameter for each reconstruction method. The errors are averages over reconstructions of 20 noisy realizations of the projection data. Shown in (a) is a comparison for the different iterative methods and FBP. Shown in (b) is the effect of using the modification of the MAP prior proposed in [32].

$$e_B^2 \quad = \quad \frac{1}{\mu_B^2 N_B} \sum_{j \in B} (\hat{x}_j - x_j)^2 \tag{18}$$

where N_L and N_B are the number of pixels in the lesion and background regions L and B respectively. We note that e_L^2 and e_B^2 are normalized relative to the true signal amplitudes since one is typically more interested in the ratio of lesion activity to background activity as compared to their absolute values. The errors e_L^2 and e_B^2 are averaged over the 20 reconstructions of the different noise realizations to give estimates $\hat{E}[e_L^2]$ and $\hat{E}[e_B^2]$ for the expected relative errors for each reconstruction method as a function of regularization parameter.

Figure 20(a) shows a plot of the expected reconstruction error $\sqrt{\hat{E}[e_L^2]}$ for the lesion versus the expected reconstruction error $\sqrt{\hat{E}[e_B^2]}$ for the background for different algorithms. The curve for each algorithm was obtained by varying the smoothing parameter. We observe that for the lesion region, the Bayesian methods achieve lower error compared to EM. Also, the GGMRF prior model performs significantly better than the GMRF prior since the GGMRF curve reaches a point closer to the coordinate origin, signifying lower combined error.

An interesting characteristic of the Bayesian methods is that the minimum error in the lesion region and the minimum error in the background region are achieved at different points on the curve, i.e., for two different values of the smoothing parameter. In other words, there is no single smoothing parameter that optimizes algorithm performance for both lesion and background. This behavior may be a result

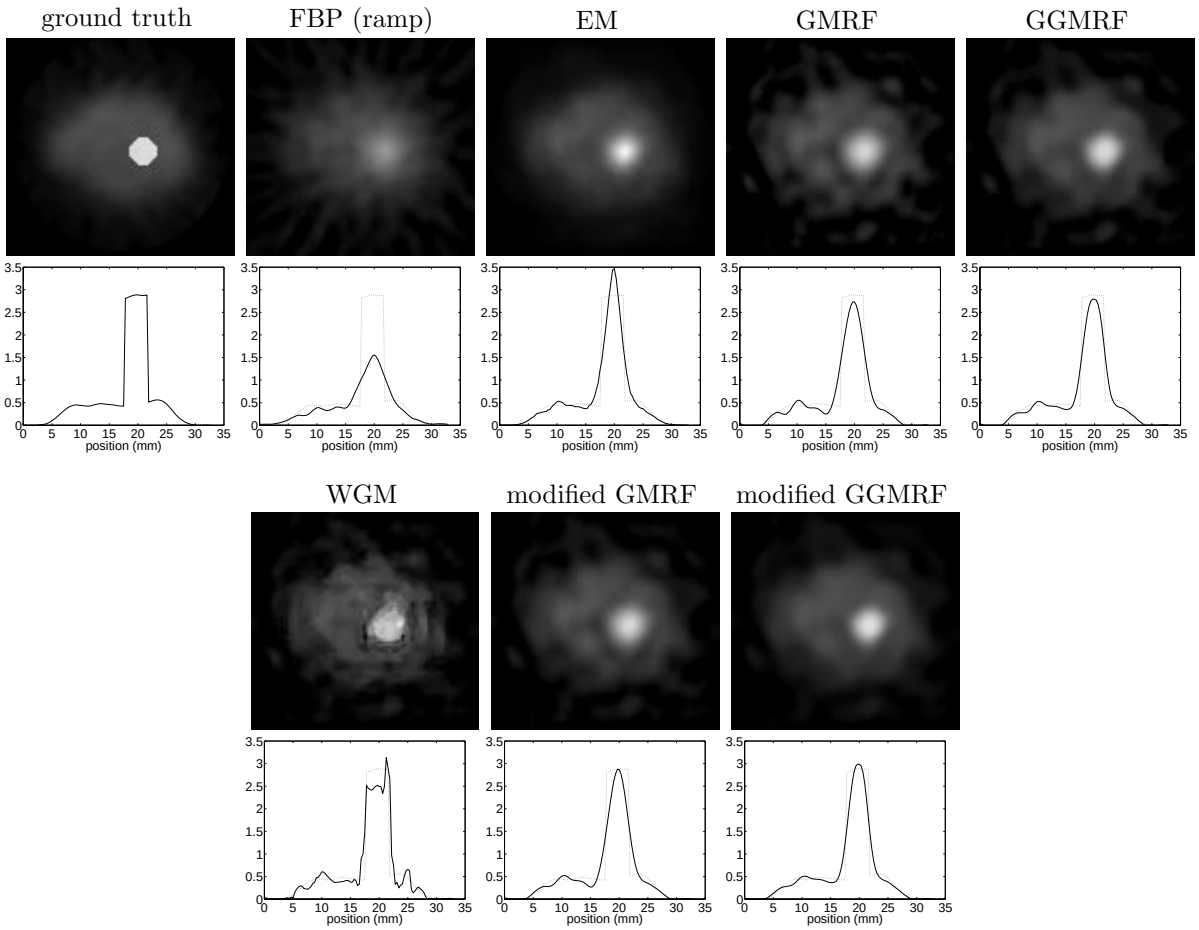


Figure 21: Reconstruction images and profiles for the simulated tumor phantom. Except for FBP, the smoothing parameter for each method was selected to minimize the sum of the relative errors for the tumor and the background region. For FBP, a ramp filter reconstruction is shown to demonstrate the limited resolution of the algorithm. The 1-D profiles correspond to a single row of pixels intersecting the center of the lesion. The ringing artifacts are due to the high reconstruction resolution of 0.31 mm per pixel compared to the detector crystal spacing of 3.3 mm on the IndyPET scanner.

of the activity dependency of the regularization of Bayesian methods[32]. Note that as indicated by the single diamond in Fig. 20(a), a lower bound for achievable performance for this problem is given by the combination of the lowest value of lesion error over all methods and lowest value of background error over all methods. Specifically, since we assume known size and location of the lesion, we can trivially achieve this bound by using one reconstruction method for the background and another one for the lesion. The distance of the bound to the measured curves indicates that there is significant room for further improvement.

Figure 20(b) shows the reconstruction error for Bayesian GMRF and GGMRF reconstructions using the modification proposed by Fessler and Rogers[32] to achieve approximately space-invariant regularization. This modification is effective in reducing the combined lesion and background errors since the curves

move closer to the coordinate origin. For the GGMRF prior, the modification results in a much smaller compromise between choosing the optimal smoothing parameter for background or lesion. However, the minimum error in the lesion region is slightly higher for the modified priors.

Figure 21 shows reconstruction images as well as 1-D profiles through the reconstructed lesion for the different algorithms. The images were cropped to a size of 101×106 pixels corresponding to a simulated physical size of 3.16×3.31 cm. The reconstruction pixel size is 0.31 mm which is about one tenth of the 3.3 mm nominal detector resolution of the IndyPET scanner. For all algorithms except for FBP, the images in Fig. 21 correspond to the value of smoothing parameter that minimizes the combined error ($\hat{E}[e_L^2] + \hat{E}[e_B^2]$) in Fig. 20. For FBP, a ramp filter reconstruction is shown to demonstrate the limited resolution that can be achieved using this algorithm. We observed that the iterative algorithms resolve the lesion substantially better than FBP. In comparison to EM, the Bayesian GMRF and GGMRF reconstructions more accurately quantify the lesion. The WGM prior model segments the lesion surprisingly accurately, however, it slightly underestimates lesion activity and has some artifacts. Overall, the modified GGMRF model appears to achieve the most accurate result. Note that the reconstructions have ringing artifacts in the background which is consistent with observations in earlier studies[6]. This ringing is due to the extremely high reconstruction resolution, which in combination with the wide detector kernel results in a highly ill-posed reconstruction problem.

4.4 Computational efficiency

As expected, the computational cost of the EM algorithm is significantly higher than for the ICD MAP algorithm. Depending on the desired level of smoothing, a factor of 10 to 20 in computation time between EM and MAP was not uncommon. We note, that ordered subset techniques can significantly improve the convergence speed of EM.

Due to the large dimensionality of the problem, even the MAP reconstructions take about 40 minutes for the reconstruction of a single 2-D plane. Specifically, the acquisition of 45 frames on the IndyPET scanner results in 45×3136 separate projection measurements. Furthermore, the sparsity of the projection matrix is reduced by using the relatively wide empirical system kernel. Storage of the complete projection matrix for all 45 frames for a reconstruction resolution of 256×256 pixels would require on the order of 5 GB while only 120 MB are needed to store the submatrix for a single gantry position.[1]

Figure 22 compares the convergence speed of ICD-MAP for different initializations of the algorithm. The graphs are based on reconstructions of the bar phantom data set with 10.4M counts using a GMRF

[1] As described in Section 3, we only store the 120 MB submatrix for the zero gantry position which can be kept in memory.

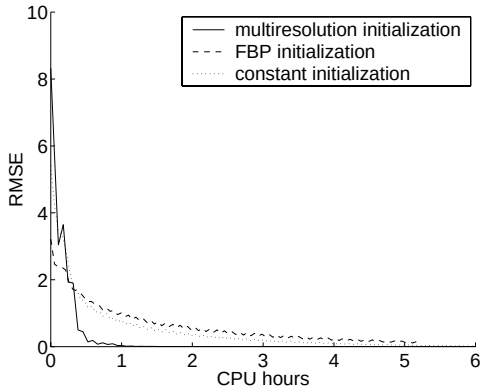


Figure 22: Plot of pixel-wise RMSE reconstruction error as a function of CPU-time for different initializations of the Bayesian ICD-MAP algorithm. The graphs are based on the reconstruction of the 10.4M count bar phantom data using a GMRF prior with smoothing parameters selected to minimize error in quadrant 4. The graph for the multiresolution initialization includes the CPU-time for the reconstructions at all resolutions.

prior. We compare the multiresolution strategy (Section 3.1) to initialization by a constant and by an FBP ramp reconstruction image. Both the constant and the FBP image were scaled such that their forward projection matched the total number of counts in the measured sinogram. The RMSE values were calculated relative to the final reconstruction images with smoothing parameters corresponding to the images shown in Fig. 9. The RMSE curves for the multiresolution initialization were obtained by separately running the complete multiresolution strategy for different numbers of iterations, using the same number of iterations at each scale. The CPU-times were measured on a 700 MHz Pentium III CPU and do not include the time needed for pre-processing such as the computation of the tomographic system matrices that can be performed off-line.

The results indicate that the multiresolution initialization significantly improves the convergence speed of the algorithm. Based on the known convergence properties of ICD[9], we do not expect a constant initialization to perform well. However, it is surprising that the convergence for FBP initialization is actually worse than that for constant initialization. This effect may be due to local or global scaling differences of the FBP reconstruction in comparison to the MAP estimate. While the convergence for FBP initialization can potentially be improved through more careful scaling, the multiresolution initialization is uncomplicated, effective, and does not require separate implementation of FBP.

5 Conclusions

We have presented a comparison of reconstruction quality for FBP, EM, and Bayesian algorithms as implemented for the IndyPET scanner. A key feature of this investigation is the determination of an

empirical system kernel based on scans of line source phantoms. For the data sets and evaluation methods used in this study, we found that without an accurate system kernel, the reconstruction quality of the iterative methods was similar to that of FBP. However, when an accurate system kernel was incorporated, Bayesian MAP techniques were superior to EM and EM was superior to FBP. In terms of computational efficiency, Bayesian MAP using ICD optimization with multiresolution initialization was also dramatically faster than EM. For the Bayesian methods, the generalized Gaussian Markov random field prior model improved quantitative accuracy and produced sharper edges compared to the Gaussian Markov random field. The recently proposed wavelet graph prior model produced sharper edges than the MRF priors but resulted in artifacts. For a simulated phantom containing a lesion on an approximately uniform background, the modifications proposed by Fessler and Rogers[32] improved the performance of the Bayesian GMRF and GGMRF techniques.

Acknowledgment

We thank Robert Koeppe, Ph.D., of the University of Michigan for supplying the Hoffman 3-D brain phantom.

Appendix A: Precise Estimation of Line Source Position

An accurate estimate of the line source position s_0 is obtained as follows. We first obtain an approximate estimate $\hat{s}_0$ directly from the projection data and use this estimate to compute an initial empirical system kernel $h(\theta, \Delta t)$. This kernel is then used to compute single-frame reconstructions of a second data set of a point source taken for a full set of $F = 45$ frames with gantry angles $\alpha_f = f * 180/F$, $0 \leq f \leq F-1$. The reconstruction images x_f for each frame are computed using only the projection data collected at frame f and using the system kernel $h(\theta - \alpha_f, \Delta t)$. Let $\tilde{s}^{(f)}$ denote the position of the point source of the second data set in the reconstruction image for frame f such that $\tilde{s}_1^{(f)}$ and $\tilde{s}_2^{(f)}$ are the horizontal and vertical positions, respectively. We compute $\tilde{s}^{(f)}$ as the center of gravity in x_f in a pixel window around the expected position. As shown in Fig. 23 (top row), $\tilde{s}_1^{(f)}$ and $\tilde{s}_2^{(f)}$ are sinusoidal functions of gantry angle, and a $\tilde{s}_1^{(f)} - \tilde{s}_2^{(f)}$ parametric plot describes a 180 degree circle segment with radius and phase corresponding to the error $\epsilon_s = \hat{s}_0 - s_0$. To estimate the radius and phase of the circle segment, we fit two sine waves $\bar{s}_1^{(f)}$ and $\bar{s}_2^{(f)}$ to the sequences $\tilde{s}_1^{(f)}$ and $\tilde{s}_2^{(f)}$

$$\bar{s}_1^{(f)} = \alpha \sin(f\pi/F + \phi_1) + \beta_1 \tag{19}$$

$$\bar{s}_2^{(f)} = \alpha \sin(f\pi/F + \phi_2) + \beta_2 \tag{20}$$

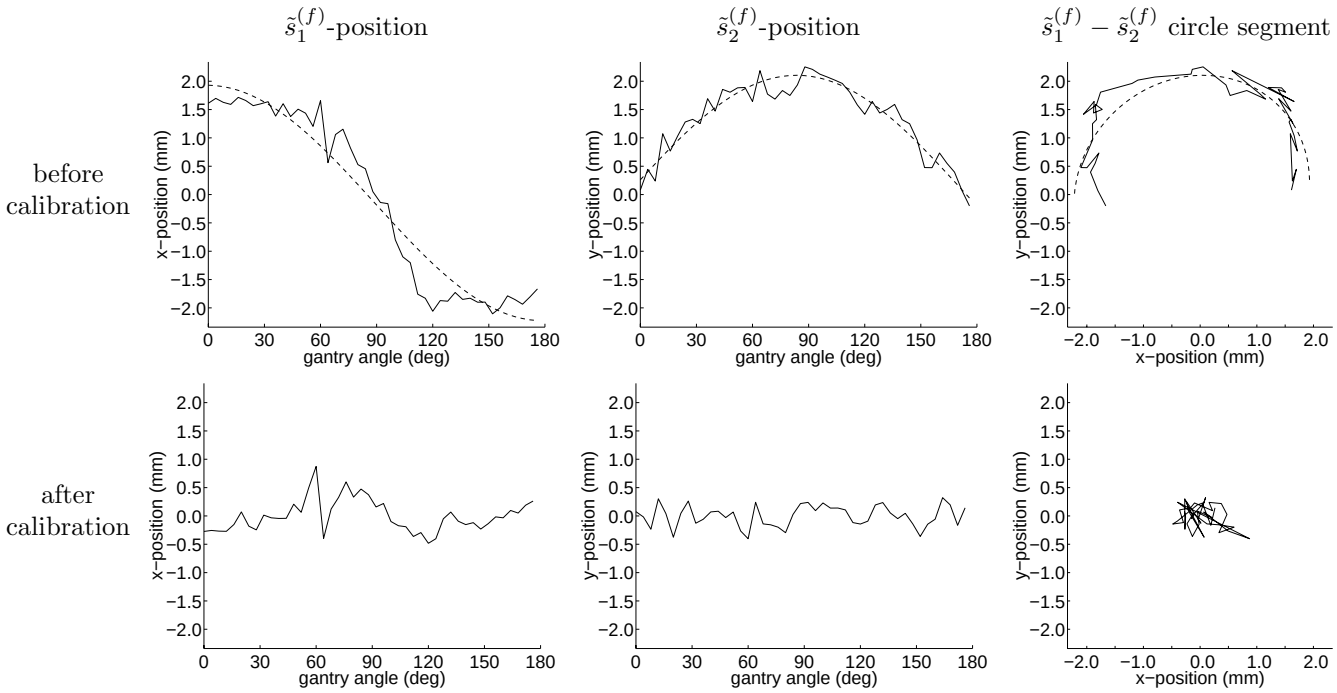


Figure 23: Plots of point source position estimated from single frame reconstruction images. The left two columns show the horizontal $\tilde{s}_1^{(f)}$-position and the vertical $\tilde{s}_2^{(f)}$-position of the point source as a function of gantry angle. The right column shows the $\tilde{s}_1^{(f)} - \tilde{s}_2^{(f)}$ parametric plot for the 180 degree gantry sweep. The top and bottom rows show the result before and after calibration respectively. The result before calibration was obtained using an initial estimate of the calibration line source position s_0 that was additionally perturbed by 2.4 mm to demonstrate the circular blurring effect. The solid lines show the estimated positions for each frame whereas the dashed lines indicate the fitted sine waves. The radius and phase of the fitted circle were used to correct the estimate of s_0, resulting in the calibrated result shown on the bottom.

where α, ϕ_1, ϕ_2, β_1 and β_2 are determined to obtain a least squares fit to $\tilde{s}_1^{(f)}$ and $\tilde{s}_2^{(f)}$. Figure 23 shows an example of these fits (dashed lines, top row). Based on the fit, the components of $\epsilon_s = \hat{s}_0 - s_0$ are estimated as $\epsilon_{s,1} = \bar{s}_1^{(0)} = \alpha \sin(\phi_1)$ and $\epsilon_{s,2} = \bar{s}_2^{(0)} = \alpha \sin(\phi_2)$. A corrected estimate $\hat{s}_0'$ is then computed as $\hat{s}_0' = \hat{s}_0 - \epsilon_s$. This more accurate estimate is then used to obtain a more precise system kernel $h(\theta, \Delta t)$. The new $h(\theta, \Delta t)$ can be tested for accuracy by repeating the frame reconstructions of the second data set. The bottom row of Fig. 23 shows the position estimates $\tilde{s}^{(f)}$ after calibration.

References

[1] R. E. Carson, Y. Yan, B. Chodkowski, T. K. Yap, and M. E. Daube-Witherspoon. Precision and accuracy of regional radioactivity quantitation using the maximum likelihood EM reconstruction algorithm. *IEEE Trans. on Medical Imaging*, 13(3):526–537, September 1994.

[2] C. A. Johnson, Y. Yan, R. E. Carson, R. L. Martino, and M. E. Daube-Witherspoon. A system for the 3D reconstruction of retracted-septa PET data using the EM algorithm. *IEEE Trans. on Nuclear Science*, 42(4):1223–1227, August 1995.

[3] E. Ü. Mumcuoğlu, R. M. Leahy, S. R. Cherry, and E. Hoffman. Accurate geometric and physical response modelling for statistical image reconstruction in high resolution PET. In *Proc. of IEEE Nucl. Sci. Symp. and*

Med. Imaging Conf., volume 3, pages 1569–1573, Anaheim, CA, November 2-9 1996.

[4] L. Shepp and Y. Vardi. Maximum likelihood reconstruction for emission tomography. *IEEE Trans. on Medical Imaging*, MI-1(2):113–122, October 1982.

[5] J-S. Liow and S. C. Strother. Practical tradeoffs between noise, quantitation, and number of iterations for maximum likelihood-based reconstructions. *IEEE Trans. on Medical Imaging*, 10(4):563–571, December 1991.

[6] J. Qi, R. M. Leahy, S. R. Cherry, A. Chatziioannou, and T. H. Farquhar. High-resolution 3D Bayesian image reconstruction using the microPET small-animal scanner. *Phys. Med. Biol.*, 43(4):1001–1013, April 1998.

[7] T. Hebert and R. Leahy. A generalized EM algorithm for 3-D Bayesian reconstruction from Poisson data using Gibbs priors. *IEEE Trans. on Medical Imaging*, 8(2):194–202, June 1989.

[8] Peter J. Green. Bayesian reconstruction from emission tomography data using a modified EM algorithm. *IEEE Trans. on Medical Imaging*, 9(1):84–93, March 1990.

[9] K. Sauer and C. A. Bouman. A local update strategy for iterative reconstruction from projections. *IEEE Trans. on Signal Processing*, 41(2):534–548, February 1993.

[10] L. Kaufmann. Maximum likelihood, least squares, and penelized least squares for PET. *IEEE Trans. on Medical Imaging*, 12(2):200–214, June 1993.

[11] J.A. Fessler and A.O. Hero. Space-alternating generalized expectation-maximization algorithms. *IEEE Trans. on Acoustic Speech and Signal Processing*, 42(10):2664–2677, October 1994.

[12] E. $\ddot{U}$. Mumcuoğlu, R. Leahy, S. R. Cherry, and Z. Zhou. Fast gradient-based methods for Bayesian reconstruction of transmission and emission PET images. *IEEE Trans. on Medical Imaging*, 13(4):687–701, December 1994.

[13] A.R. De Pierro. A modified expectation maximization algorithm for penalized likelihood estimation in emission tomography. *IEEE Trans. on Medical Imaging*, 14(1):132–137, March 1995.

[14] C. A. Bouman and K. Sauer. A unified approach to statistical tomography using coordinate descent optimization. *IEEE Trans. on Image Processing*, 5(3):480–492, March 1996.

[15] A. Blake and A. Zisserman. *Visual Reconstruction*. MIT Press, Cambridge, Massachusetts, 1987.

[16] R. Stevenson and E. Delp. Fitting curves with discontinuities. *Proc. of the first international workshop on robust computer vision*, pages 127–136, October 1-3 1990.

[17] D. Geman and G. Reynolds. Constrained restoration and the recovery of discontinuities. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 14(3):367–383, March 1992.

[18] C. A. Bouman and K. Sauer. A generalized Gaussian image model for edge-preserving MAP estimation. *IEEE Trans. on Image Processing*, 2(3):296–310, July 1993.

[19] S. Brette, J. Idier, and Ali Mohammad-Djafari. Scale invariant Markov models for Bayesian inversion of linear inverse problems. In J. Skilling and S. Sibisi, editors, *Maximum Entropy and Bayesian Methods*, pages 199–212. Kluwer Academic Publishers, Netherlands, 1996.

[20] J. Qi, R. M. Leahy, C. Hsu, T. H. Farquhar, and S. R. Cherry. Fully 3D Bayesian image reconstruction for the ECAT EXACT HR+. *IEEE Trans. on Nuclear Science*, 45(3):1096–1103, June 1998.

[21] S. R. Cherry, Y. Shao, S. Siegel, R. W. Silverman, K. Meadors, J. Young, W. F. Jones, D. Newport, C. Moyers, E. $\ddot{U}$. Mumcuoğlu, M. Andreaco, M. Paulus, D. Binkley, R. Nutt, and M. E. Phelps. MicroPET: a high resolution PET scanner for imaging small animals. *IEEE Trans. on Nuclear Science*, 44(3):1161–1166, June 1997.

[22] G. Brix, J. Zaers, L-E Adam, M. E. Bellemann, H. Ostertag, H. Trojan, U. Haberkorn, J. Doll, F. Oberdorfer, and W. J. Lorenz. Performance evaluation of a whole-body PET scanner using the NEMA protocol. *Journal of Nuclear Medicine*, 38(10):1614–1623, October 1997.

[23] A. Chatziioannou, J. Qi, A. Moore, A. Annala, K. Nguyen, R. M. Leahy, and S. R. Cherry. Comparison of 3D maximum a posteriori and filtered backprojection algorithms for high resolution animal imaging with microPET. *IEEE Trans. on Medical Imaging*, 19(5):507–512, May 2000.

[24] N. C. Rouze, W. Winkle, and G. D. Hutchins. IndyPET - a high resolution, high sensitivity dedicated research scanner. In *Proc. of IEEE Nucl. Sci. Symp. and Med. Imaging Conf.*, volume 3, pages 1460–1464, Seattle, WA, October 24-30 1999.

[25] T. J. Holmes and D. C. Ficke. Analysis of positron-emission tomography scintiallation-detectors with wedge faces and inter-crystal septa. *IEEE Trans. on Nuclear Science*, 32(1):826–830, February 1985.

[26] Zhengrong Liang. Detector response restoration in image reconstruction of high resolution positron emission tomography. *IEEE Trans. on Medical Imaging*, 13(2):314–321, June 1994.

[27] Neil A. Keller and L. R. Lupton. PET detector ring aperture function calculations using monte carlo techniques. *IEEE Trans. on Nuclear Science*, 30(1):676–680, February 1983.

[28] R. Lecomte, D. Schmitt, and G. Lamoureux. Geometric study of a high resolution PET detection system using small detectors. *IEEE Trans. on Nuclear Science*, 31(1):556–561, February 1984.

[29] R. H. Huesman, E. M. Salmeron, and J. R. Baker. Compensation for crystal penetration in high resolution positron tomography. *IEEE Trans. on Nuclear Science*, 36(1):1100–1107, February 1989.

[30] T. Frese, C. A. Bouman, N. C. Rouze, G. D. Hutchins, and K. Sauer. Bayesian multiresolution algorithm for PET reconstruction. In *Proc. of IEEE Int'l Conf. on Image Proc.*, pages 613–616, Vancouver, BC, Canada, September 10-13 2000.

[31] T. Frese, C. A. Bouman, and K. Sauer. Adaptive wavelet graph model for Bayesian tomographic reconstruction. *To appear in the IEEE Trans. on Image Processing.*

[32] J. A. Fessler and W. L. Rogers. Spatial resolution properties of penalized-likelihood image reconstruction: Space-invariant tomographs. *IEEE Trans. on Image Processing*, 5(9):1346–1358, September 1996.

[33] K. Wienhard, M. Dahlbom, L. Eriksson, C. Michel, T. Bruckbauer, U. Pietrzyk, and W-D. Heiss. The ECAT EXACT HR: performance of a new high resolution positron scanner. *J. Comput. Assist. Tomogr.*, 18(1):110–118, Jan.-Feb. 1994.

[34] S. Batchelor, G. M. Blake, and J. E. Saunders. A comparison of three commercially available PET imaging systems. *Nuclear Medicine Communications*, 13(1):20–27, January 1992.

[35] David Barnes, Gary Egan, Graeme O'Keefe, and David Abbott. Characterization of dynamic 3-D PET imaging for functional brain mapping. *IEEE Trans. on Medical Imaging*, 16(3):261–269, June 1997.

[36] S. G. Azevedo, D. J. Schneberk, J. P. Fitch, and H. E. Martz. Calculation of the rotational centers in computed tomography sinograms. *IEEE Trans. on Nuclear Science*, 37(4):1525–1540, August 1990.

[37] M. Yavuz and J. A. Fessler. Statistical image reconstruction methods for randoms-precorrected PET scans. *Medical Image Analysis*, 2(4):369–378, December 1998.

[38] L. Kaufmann. Implementing and accelerating the EM algorithm for positron emission tomography. *IEEE Trans. on Medical Imaging*, MI-6(1):37–51, 1987.

[39] E. *Ü*. Mumcuoğlu, R. Leahy, and S. R. Cherry. Bayesian reconstruction of PET images: Methodology and performance analysis. *Phys. Med. Biol.*, 41(9):1777–1807, September 1996.

[40] M. T. Chan, R. M. Leahy, E. Ü. Mumcuoğlu, S. R. Cherry, J. Czernin, and A. Chatziioannou. Comparing lesion detection performance for PET image reconstruction algorithms: a case study. *IEEE Trans. on Nuclear Science*, 44(4):1558–1563, August 1997.

[41] J. Llacer, E. Veklerov, L. R. Baxter, S. T. Grafton, L. K. Griffeth, R. A. Hawkins, C. K. Hoh, J. C. Mazziotta, E. J. Hoffman, and C. E. Metz. Results of a clinical receiver operating characteristic study comparing filtered backprojection and maximum likelihood estimator images in FDG PET studies. *Journal of Nuclear Medicine*, 34(7):1198–1203, July 1993.