

# A Comparative Study of the Effects of Using Higher Order Mechanical Priors in SPECT Reconstruction

Soo-Jin Lee<sup>#</sup>, Anand Rangarajan<sup>§</sup>, and Gene Gindi<sup>#¶</sup>

<sup>#</sup>Department of Electrical Engineering, SUNY at Stony Brook

<sup>§</sup>Department of Computer Science, Yale University

<sup>¶</sup>Department of Radiology, SUNY at Stony Brook

## Abstract

While the ML-EM reconstruction method in SPECT is unstable due to its ill-posed nature, the Bayesian reconstruction methods overcome this instability by introducing the prior information. The new prior, the *weak plate*, is attractive in that it models spatial structures more accurately than various forms of smoothness priors used in the past. Here we argue that more expressive priors able to model more complicated forms than a piecewise constant source may be better used in SPECT reconstructions. To evaluate quantitative performance of the reconstruction algorithms, in an ensemble sense, bias and variance were used as a metric. We show that the extension to a higher-order model significantly improves variance at approximately no cost in bias error. We also observe and characterize the behavior of the associated hyperparameters of the prior distributions in a systematic way.

## I. INTRODUCTION

The maximum likelihood (ML) approach using the expectation maximization (EM) algorithm has been useful in reconstruction for emission tomography. However, due to the ill-posed nature of the problem, the ML-EM suffers from instability. In contrast, the Bayesian reconstruction methods overcome this instability by incorporating prior information into the reconstruction problem. Early formulations [1, 2, 3] incorporated priors that imposed piecewise constant spatial constraints on the underlying object. Reconstruction algorithms that use this kind of assumption perform better, according to most reasonable metrics, than those that make no such assumption, when tested on such phantoms. However, the self-consistent loop of piecewise constant prior and piecewise constant phantom leads one to question whether these good results generalize to a realistic clinical setting where the underlying (patient) source distribution may not be piecewise constant.

In our preliminary work [4], we showed possibilities of improving reconstructions by using a more accurate model than a piecewise constant prior, the *weak plate* (WP) prior. Results in [4] show that the WP prior leads to improved estimate in terms of simple root mean-squared error (RMSE)

criteria compared to the weak membrane (WM) prior. In this paper, we evaluate and compare quantitative performance of these reconstruction algorithms more systematically in order to validate possible advantages of a higher order model. This work follows a methodology in [5] for bias/variance characterization. Since these reconstruction algorithms are dependent on two hyperparameters embedded in the prior distributions, we observe and characterize, in an ensemble sense, the effects of these parameters.

## II. BAYESIAN RECONSTRUCTION WITH MECHANICAL PRIORS

Since our priors encourage the formation of smooth regions punctuated by a locus of discontinuities, we use an unobservable line processes  $\mathbf{I}$  proposed by Geman and Geman [1] to preserve the discontinuities in the reconstructions. The MAP approach in the context of Bayesian framework is then to estimate both the source field  $\mathbf{f}$  and the discontinuity field  $\mathbf{I}$  simultaneously by maximizing the posterior probability, given as

$$\begin{aligned} \Pr(\mathbf{F} = \mathbf{f}, \mathbf{L} = \mathbf{l} | \mathbf{G} = \mathbf{g}) \\ = \frac{\Pr(\mathbf{G} = \mathbf{g} | \mathbf{F} = \mathbf{f}, \mathbf{L} = \mathbf{l}) \Pr(\mathbf{F} = \mathbf{f}, \mathbf{L} = \mathbf{l})}{\Pr(\mathbf{G} = \mathbf{g})}, \end{aligned} \quad (1)$$

where  $\mathbf{f}$ ,  $\mathbf{I}$ , and  $\mathbf{g}$  are 2-D vector fields for the source intensities, line processes, and projection data respectively, and  $\mathbf{F}$ ,  $\mathbf{L}$ ,  $\mathbf{G}$  are the associated random fields. Given the posterior distribution in (1), maximizing the posterior distribution is equivalent to minimizing  $-\log$  of the posterior probability and the MAP estimation reduces to  $(\hat{\mathbf{f}}, \hat{\mathbf{l}}) = \arg \min_{(\mathbf{f}, \mathbf{l})} [-\log \Pr(\mathbf{G} = \mathbf{g} | \mathbf{F} = \mathbf{f}) - \log \Pr(\mathbf{F} = \mathbf{f}, \mathbf{L} = \mathbf{l})]$ ,

where the two terms on the right side are the likelihood and the prior, respectively. For the likelihood, Poisson statistics are applied in a conventional way.

We use the familiar Gibbs distribution to model prior information concerning the piecewise smoothness of the image:

$$\Pr(\mathbf{F} = \mathbf{f}, \mathbf{L} = \mathbf{l}) = \frac{1}{Z_P} \exp(-E_P(\mathbf{f}, \mathbf{l})),$$

where  $Z_P$  is a normalization factor and  $E_P(\mathbf{f}, \mathbf{l})$  is the prior energy derived from the mechanical models ( $E_P^M(\mathbf{f}, \mathbf{l})$  for

WM and  $E_P^P(\mathbf{f}, \mathbf{l})$  for WP), given as [6]

$$E_P^M(\mathbf{f}, \mathbf{l}) = \lambda \sum_{ij} \{V_{i,j}^v(1 - l_{i,j}^h) + V_{i,j}^h(1 - l_{i,j}^v)\} + \alpha \sum_{ij} (l_{i,j}^h + l_{i,j}^v) \quad (2)$$

$$E_P^P(\mathbf{f}, \mathbf{l}) = \lambda \sum_{i,j} \{V_{i,j}(1 - l_{i,j})\} + \alpha \sum_{ij} l_{i,j}, \quad (3)$$

where  $V_{i,j}^v \stackrel{\text{def}}{=} f_v^2(i, j)$  and  $V_{i,j}^h \stackrel{\text{def}}{=} f_h^2(i, j)$  for WM, and  $V_{i,j} \stackrel{\text{def}}{=} f_{vv}^2(i, j) + 2f_{hv}^2(i, j) + f_{hh}^2(i, j)$  for WP. The binary variable  $l$  is line process (for WM,  $l^v$  and  $l^h$  are vertical and horizontal line processes, respectively), and  $\lambda$  and  $\alpha$  are positive parameters. In (2),  $f_v$  and  $f_h$  are the first derivatives along the vertical and the horizontal directions, respectively. The term  $V_{i,j}$  in (3) is quadratic variation for thin plate [6]. The terms involving  $V_{i,j}^v$  and  $V_{i,j}^h$  in (2) and  $V_{i,j}$  in (3) encourage smoothness except where discontinuities occur (i.e.  $l = 1$ ) and include a penalty  $\alpha$  for the creation of discontinuities. Note that while the WM encourages smoothness only in constant regions, the WP encourages smoothness in ramplike regions as well. A fundamental limitation of the WM is that it has the unfortunate effect of turning a ramp into stepped terraces [6].

As it stands, the energy functions above are difficult to minimize, since they contain mixed binary and continuous variables. Although the use of binary variables can be circumvented by integrating out the line processes, the new energy functions are non-convex and non-differentiable, thereby ruling out gradient-based descent methods. In [3], it is shown that the desired energy function can be approached by a sequence of energy functions indexed by a control parameter  $\beta$  – a continuation method. Readers are referred to [3] for more details on the derivation of a deterministic annealing algorithm.

To make the optimization tractable, we integrate out the line processes and embed the optimization within the familiar incomplete/complete data formulation of the Generalized EM (GEM) MAP algorithm [2]. The corresponding M-step objective function for WP is

$$M(\mathbf{f}|\hat{\mathbf{f}}^n; \beta) \stackrel{\text{def}}{=} \sum_{ij} \sum_{t\theta} \left[ H_{t\theta ij} f_{ij} - g_{t\theta} \frac{H_{t\theta ij} \hat{f}_{ij}^n}{\sum_{kl} H_{t\theta kl} \hat{f}_{kl}^n} \log f_{ij} \right] + \sum_{ij} \left[ -\frac{1}{\beta} \log(\exp(-\beta \lambda V_{i,j}) + \exp(-\beta \alpha)) \right],$$

where  $g_{t\theta}$  is the number of detected counts in the detector bin indexed by  $t$  at angle  $\theta$ , and  $\mathcal{H}_{t\theta, ij}$  is the contribution of a photon emitted from  $f_{i,j}$  to detector bin  $t$  at angle  $\theta$ . The prior energy for WM is represented by the two terms involving  $V_{i,j}^v$  and  $V_{i,j}^h$  [3]. Unfortunately, closed-form solution for  $f_{i,j}$  cannot be directly obtained by setting  $\frac{\partial M(\mathbf{f}|\hat{\mathbf{f}}^n; \beta)}{\partial f_{i,j}} = 0$ , as the differentiation results in an expression transcendental in  $f_{i,j}$ . One possible solution to

this problem is to consider the transcendental term to be a new, albeit dependent (on  $f_{i,j}$ ), variable and to descend on each such variable separately. In [3], we defined the new variable as

$$z_{i,j} \stackrel{\text{def}}{=} \frac{1}{1 + \exp(-\beta(\lambda V_{i,j} - \alpha))}. \quad (4)$$

With the new variables,  $M(\mathbf{f}|\hat{\mathbf{f}}^n; \beta)$  becomes  $M(\mathbf{f}, \mathbf{z}|\hat{\mathbf{f}}^n; \beta)$ . For the maximization step, we apply a coordinate-wise descent method by keeping  $f_{i,j}$  frozen while descending on  $z_{i,j}$  and keeping  $z_{i,j}$  frozen while descending on  $f_{i,j}$ . The update equations for WP are:

$$f_{ij} = \frac{-(\sum_{t\theta} \mathcal{H}_{t\theta, ij} - 2\lambda X_3) + \sqrt{(\sum_{t\theta} \mathcal{H}_{t\theta, ij} - 2\lambda X_3)^2 + 8\lambda X_2 X_1}}{4\lambda X_2} \quad (5)$$

where

$$\begin{aligned} X_1 &\stackrel{\text{def}}{=} \sum_{t\theta} g_{t\theta} \frac{\mathcal{H}_{t\theta, ij} \hat{f}_{ij}^n}{\sum_{kl} \mathcal{H}_{t\theta, kl} \hat{f}_{kl}^n} \\ X_2 &\stackrel{\text{def}}{=} 10(1 - z_{i,j}) + (1 - z_{i,j+1}) + 3(1 - z_{i,j-1}) \\ &\quad + (1 - z_{i+1,j}) + 3(1 - z_{i-1,j}) + 2(1 - z_{i-1,j-1}) \\ X_3 &\stackrel{\text{def}}{=} \{2(f_{i,j+1} + f_{i,j-1}) + 2(f_{i+1,j} + f_{i-1,j}) \\ &\quad + 2(f_{i,j+1} + f_{i+1,j} - f_{i+1,j+1})\}(1 - z_{i,j}) \\ &\quad + (2f_{i,j+1} - f_{i,j+2})(1 - z_{i,j+1}) \\ &\quad + \{(2f_{i,j-1} - f_{i,j-2}) \\ &\quad + 2(f_{i,j-1} - f_{i+1,j-1} + f_{i+1,j})\}(1 - z_{i,j-1}) \\ &\quad + (2f_{i+1,j} - f_{i+2,j})(1 - z_{i+1,j}) \\ &\quad + \{(2f_{i-1,j} - f_{i-2,j}) \\ &\quad + 2(f_{i-1,j} - f_{i-1,j+1} + f_{i,j+1})\}(1 - z_{i-1,j}) \\ &\quad + (2f_{i-1,j} + f_{i,j-1} - f_{i-1,j-1})(1 - z_{i-1,j-1}). \end{aligned}$$

Update equations for WM appear in [3].

### III. EXPERIMENTS AND RESULTS

We first tested the four reconstruction algorithms – GEM MAP with the WM prior, GEM MAP with the WP prior, and ML-EM with two different stopping rules. The first ML-EM reconstructions (EM-1) stopped at minimum RMSE and the second ML-EM (EM-2) stopped after minimum RMSE based on the starting point of deterioration of the smoothness in the reconstructed images. For the WM and WP reconstructions, the annealing schedule ran through 12 to 16 values with a doubling at each new value of  $\beta$ . The entire simulation was terminated when  $z_{i,j} \leq 0.9$  or  $z_{i,j} \geq 0.1$ .

The algorithms were tested on two phantoms (A and B for convenience). Phantom A (Fig. 1(a)) was derived from digitized rhesus monkey autoradiograph of the rCBF SPECT agent  $^{99m}\text{Tc-ECD}$ . The intent here is that such an autoradiograph contains edge structure for more realistic than a simple piecewise constant phantom. Phantom B

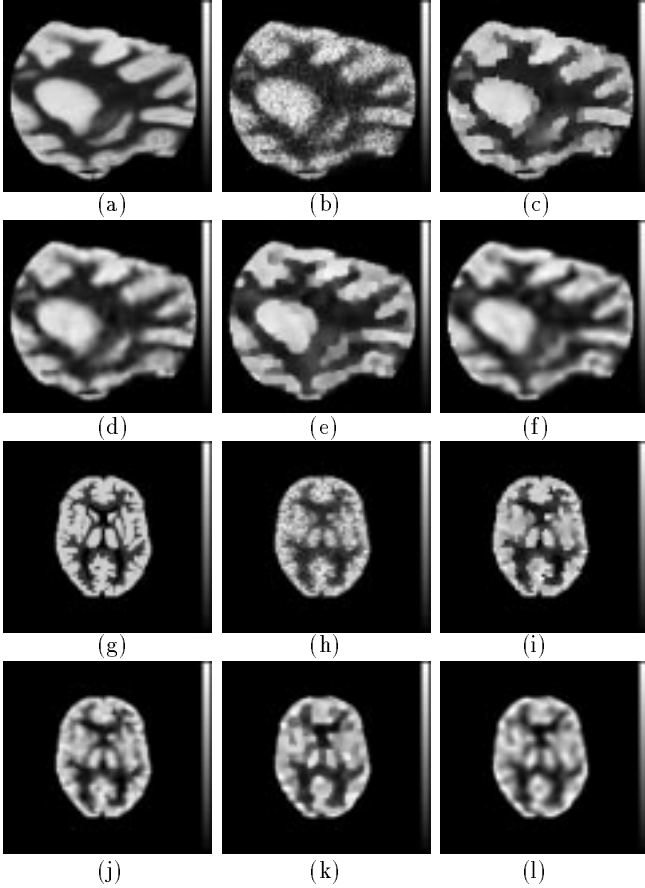


Figure 1: Anecdotal reconstructions for phantoms A and B. (a),(g): Phantoms. (b)(h),(c)(i),(d)(j): EM-2, WM, and WP reconstructions, respectively, for case of attenuation simulation and compensation only. (e)(k),(f)(l): WM and WP reconstructions respectively, for case of attenuation, scatter, and detector response simulation and compensation.

(Fig. 1(g)) is a 2-D Hoffman brain phantom which is piecewise constant.

For projection data from  $128 \times 128$  phantoms with 0.2cm pixels, we used 128 projection angles over  $360^\circ$  with 192 detector bins with 0.2cm intervals. Each projection ray was attenuated by the constant attenuation factor of  $0.12\text{cm}^{-1}$ . For measure of quantitative performance of the reconstruction algorithms in an ensemble sense, we generated two sets of 40 independent Poisson noise realizations of projection data for each phantom. One set simulated the effects of attenuation only and the second set included additional simulations for scatter and detector response with 11mm of full width at half maximum (FWHM). The total number of detector counts for phantoms A and B were approximately 732K and 300K, respectively.

Fig. 1 shows anecdotal reconstructed images from the four reconstruction algorithms. The ML-EM reconstructions look relatively noisy and the WM reconstructions look artificially patchy, a result not unexpected since WM tends to favor piecewise constant reconstructions. By extension to a higher order model, the WP reconstructions avoid the artifact of WM reconstructions. Fig. 2 shows profiles along two lines on phantom A. Both EM plots are noisy, the result for EM-1 (fewer iterations) is smoother

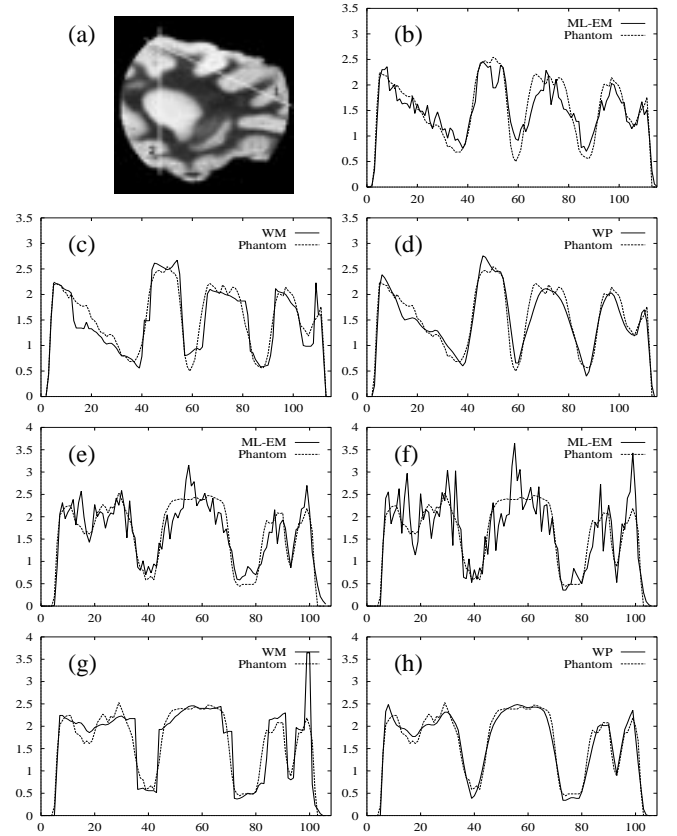


Figure 2: Profile plots for an anecdotal reconstruction of Phantom A. The dotted line indicates the phantom, and the solid line the reconstruction. (a) Phantom with profile lines shown. (b),(c),(d): Profiles along line 1 for EM-1, WM, WP, respectively. (e),(f),(g),(h): Profiles along line 2 for EM-1, EM-2, WM, WP, respectively.

than that for EM-2 but incurs a bias error. The WM reconstruction illustrates the tendency towards piecewise constant regions. The WM has reconstructed a ramp as a step edge typifying the kinds of ensemble accuracy and precision errors. Comparison of the WP result in Fig. 2(h) shows that the noise errors typical of EM (Figs. 2(e) and (f)) and edge approximation and hot spot errors typical of WM Fig. 2(g) have both been corrected.

To evaluate the reconstructions more quantitatively, we computed bias and standard deviation (STD) images (Figs. 3). A bias image,  $b_{i,j}$ , and a standard deviation image,  $s_{i,j}$ , are defined as  $b_{i,j} = \frac{1}{K} \sum_{k=1}^K (\hat{f}_{i,j}^k - f_{i,j})$  and  $s_{i,j} = \sqrt{\frac{1}{K-1} \sum_{k=1}^K (\hat{f}_{i,j}^k - \bar{f}_{i,j})^2}$ , respectively, where  $\hat{f}_{i,j}^k$  is the  $k^{\text{th}}$  reconstruction of phantom  $f$  at location  $(i,j)$  and  $\bar{f}_{i,j}$  is the mean of  $\hat{f}_{i,j}^k$  over  $K = 40$  independent noise trials. Despite the unsatisfactory WM anecdotal images, the bias for WM compares favorably with that of EM and WP. The STD images for EM show noise growing with signal intensity as expected. For the WM prior, the variance is high in both ramplike and steplike edge regions. This effect appears to be due to the fact that the WM prior results in unstable estimates of edge location. For comparable performance in bias, the WP prior produces excellent variance results as seen in Figs 3.

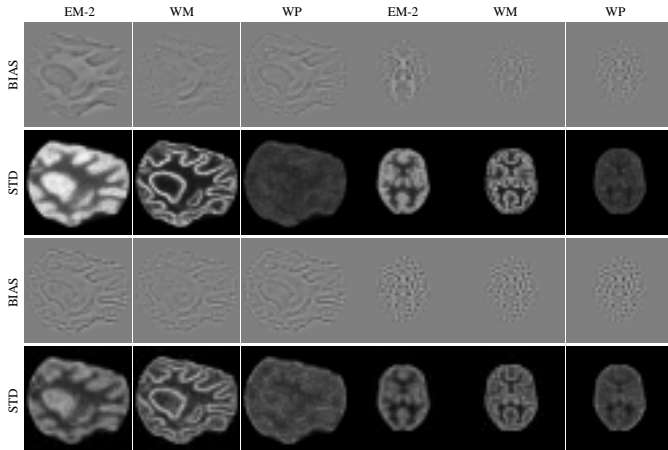


Figure 3: Bias-standard deviation for phantoms A and B obtained from 40 independent noise realizations. The first two rows include the effects of attenuation only and the second two rows include additional effects of scatter and detector response.

We measured regional bias and STD for each reconstruction algorithm in the 15 preselected ROI's shown in Fig. 4(a). Note that the boxed regions (9, 10, 11, 12, and 15) have higher contrast than the circled regions. We defined regional bias,  $b^R$ , and STD,  $s_1^R$ , as  $b^R = \frac{1}{m} \sum_{i,j \in R} b_{i,j}$  and  $s_1^R = \sqrt{\frac{1}{m} \sum_{i,j \in R} s_{i,j}^2}$ , respectively, where  $R$  denotes a region and  $m$  is the number of pixels in the region. Figs. 4(b),(c) illustrate clearly the bias and variance trade-off for EM-1 and EM-2. Comparison of Figs. 4(d),(e) shows that for (boxed) high contrast region, the WM variance is much higher than that for WP, but for (circled) smooth regions, WM compares favorably. Comparison of Figs. 4(c),(d),(e) shows that for comparable performance in bias, the WP exhibits superior performance in variance in regions including edge structure.

To observe the sensitivity of the parameters  $\lambda$  and  $\alpha$  for the WM and WP priors, we used a coarser version ( $64 \times 64$ ) of phantom A with the same number of detector counts. This phantom was designated C. We tested WM and WP priors with two different stopping rules. The first WM and WP reconstructions (WM-1 and WP-1, respectively) stopped at minimum RMSE and the second WM and WP (WM-2 and WP-2, respectively) stopped when  $z_{i,j} \leq 0.9$  or  $z_{i,j} \geq 0.1$ . We first searched recursively for the parameter pairs  $(\hat{\lambda}, \hat{\alpha})$  for WM-1 and WP-1 which yielded the reconstruction with the least RMSE. Having found  $(\hat{\lambda}, \hat{\alpha})$  for each algorithm, we generated  $20 \times 20$  different parameter pairs around  $(\hat{\lambda}, \hat{\alpha})$ . Fig. 5 shows contour plots drawn along the equi-RMSE's. Notice that the plots in Fig. 5 are unimodal. It is interesting to note that the region for the least RMSE for WM-1 is far from that for WM-2 (Figs. 5(a),(b)). In fact, the parameter  $\alpha$  for WM-2 approached infinity to produce the reconstruction with the least RMSE. In contrast, the WP exhibits the consistent parameter values for the least RMSE (Figs. 5(c),(d)). Comparison of Figs. 5(b),(d) shows that small changes in the parameters for the WM incur larger variations in

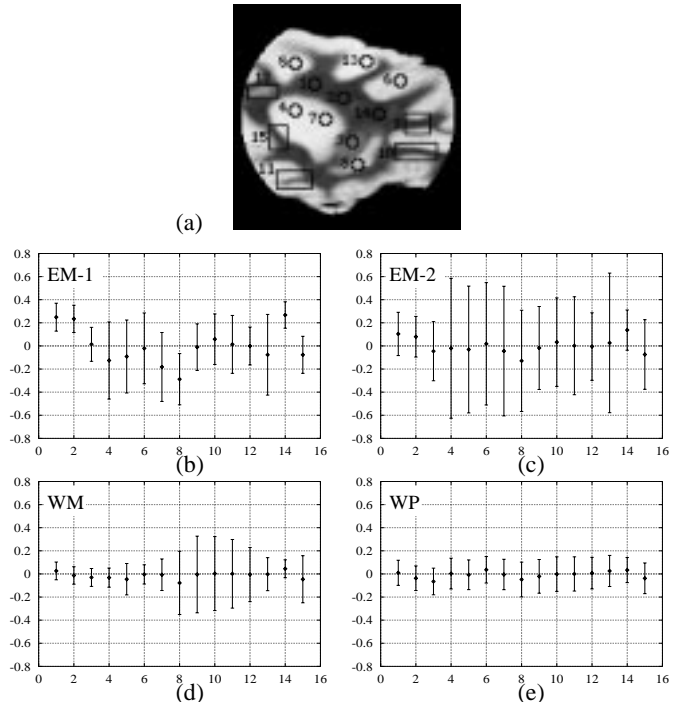


Figure 4: Regional bias-standard deviation plots for phantom A. (a) The phantom with region borders superposed. In (b)-(e), the abscissa indexes the region number, and plotted points with error bars indicate bias  $\pm$  one standard deviation.

RMSE than the WP case.

To characterize the effects of the parameter  $\lambda$  for the WM and WP priors, we used phantom C and generated 40 independent noise realizations of projection data with attenuation. We generated a set of  $\lambda$ 's for each reconstruction algorithm:  $\{\lambda_{-4}, \lambda_{-3}, \dots, \lambda_0, \dots, \lambda_3, \lambda_4\}$ , where  $\lambda_{i+1} = 2\lambda_i$  for  $i \in [-4, 3]$ . The parameter pair  $\hat{\lambda}$  and  $\hat{\alpha}$  obtained above were used for  $\lambda_0$  and  $\alpha$ , respectively. Since standard definitions for regional bias ( $b^R$ ) and standard deviation ( $s_1^R$ ) do not adequately account for the spatial structure on the reconstructed images, we used the vector fields (array of points in ROI)  $\mathbf{f}_R$  and  $\hat{\mathbf{f}}_R$  for the regions on the phantom and on the estimate, respectively. The new definitions for regional bias and standard deviation using these vector fields are:  $\mathbf{b}^R = \hat{\mathbf{f}}_R - \mathbf{f}_R$  for regional bias, and  $s_2^R = \sqrt{\frac{1}{K} \sum_{k=1}^K \|\hat{\mathbf{f}}_R^k - \mathbf{f}_R^k\|^2}$  for regional standard deviation, where  $\hat{\mathbf{f}}_R = \frac{1}{K} \sum_{k=1}^K \hat{\mathbf{f}}_R^k$ . Fig. 6 shows plots for standard deviation ( $s_2^R$ ) vs. norm of bias ( $\|\mathbf{b}^R\|$ ) of the representative four regions. For relatively constant regions, both bias and variance for the WM and WP reconstructions decreased until  $\lambda$  passed some critical value, at which point bias for both WM and WP started to increase (Figs. 6(a),(b)). For regions with edge structure, while the WP exhibited relatively small variance at most of the bias levels, variance for the WM increased significantly after some critical values (Figs. 6(c),(d)). This effect demonstrates that the WM mislocalizes edges in ramplike regions and large  $\lambda$  incurs large bias and variance errors in those regions.

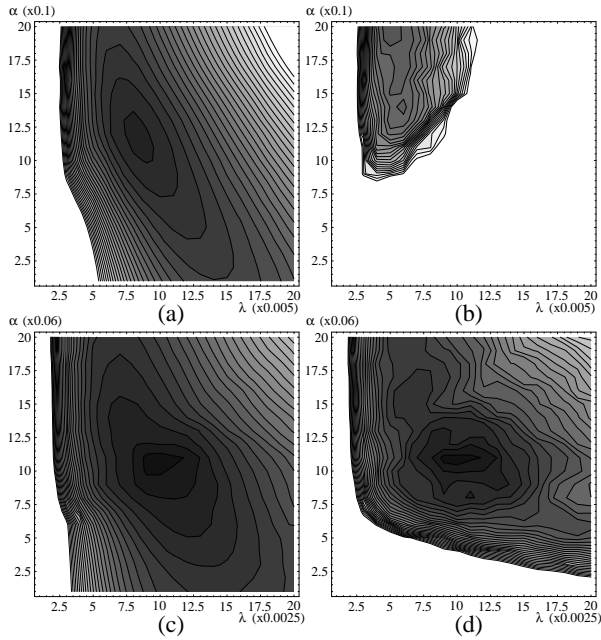


Figure 5: Contour plots for sensitivity of the parameters  $\lambda$  (abscissa) and  $\alpha$  (ordinate). Each contour level indicates the equi-error within a range. The darkest areas correspond to the parameters for the least RMSE's. (a),(c): Plots for WM and WP, respectively, stopped at the minimum RMSE. (b),(d): Plots for WM and WP, respectively, stopped when  $z_{i,j} \leq 0.9$  or  $z_{i,j} \geq 0.1$ .

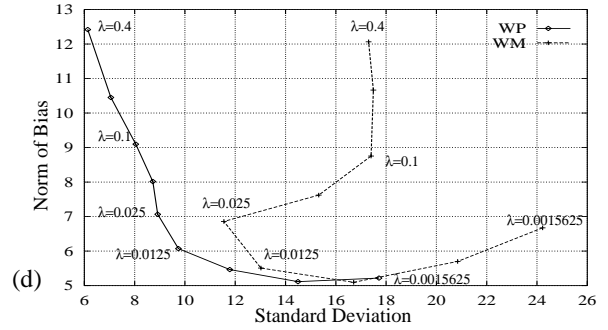
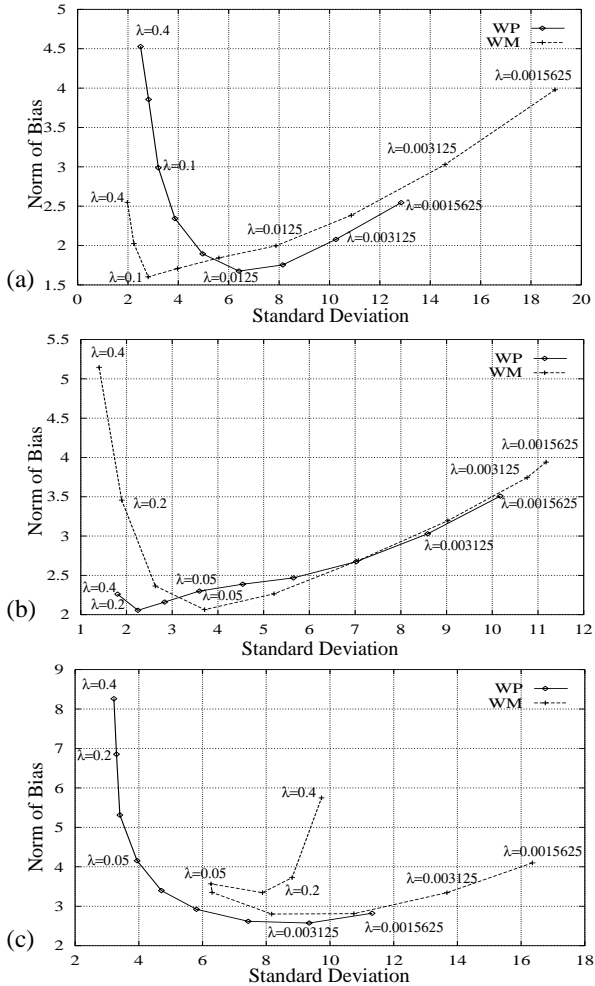


Figure 6: The effects of several different values for  $\lambda$  in WM and WP reconstructions. The four representative regions in phantom C – (a) high activity region (grey matter), (b) low activity region (white matter), (c) low contrast (ramplike) region, and (d) high contrast region (anatomical boundary) – were measured in terms of regional bias and standard deviation. The abscissa indicates regional standard deviation  $s_2^R$ , and the ordinate norm of regional bias,  $\|b^R\|$ .

## IV. CONCLUSIONS

We have considered quantitative performance measurements for Bayesian reconstruction algorithms with mechanical models as priors. Regional bias and variance analyses show that for comparable bias, the WP exhibits superior performance in variance in most of the regions, including edge structure. Observations for sensitivity of the parameters  $\lambda$  and  $\alpha$  show that small changes in the parameters for the WM incur larger variations in RMSE than those for the WP. Comparisons of the effects of the smoothing parameter  $\lambda$  demonstrate that the WP exhibits small variance at most of the bias levels. An overall conclusion is that the extension to a higher-order model significantly improves variance at approximately no cost in bias error for the types of “functionally correct” phantoms.

## REFERENCES

- [1] S. Geman and D. Geman, “Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images”, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, PAMI-6(6), pp. 721–741, November 1984.
- [2] T. Hebert and R. Leahy, “A Generalized EM Algorithm for 3-D Bayesian Reconstruction for Poisson Data using Gibbs Priors”, *IEEE Trans. on Medical Imaging*, MI-8(2), pp. 194–202, June 1989.
- [3] G. Gindi, A. Rangarajan, M. Lee, P. J. Hong, and G. Zubal, “Bayesian Reconstruction for Emission Tomography via Deterministic Annealing”, In H. Barrett and A. Gmitro, editors, *Information Processing in Medical Imaging*, pp. 322–338, Springer-Verlag, 1993.
- [4] S. J. Lee, A. Rangarajan, and G. Gindi, “Weak Plate Mechanical Models in Bayesian Reconstruction for Emission Tomography”, In *Proc. IEEE Nuclear Science Symposium and Medical Imaging Conference*, volume 3, pp. 1533–1537, November 1993.
- [5] Z. Zhou, R. M. Leahy, and E. U. Mumcuoglu, “A Comparative Study of the Effects of Using Anatomical Priors in PET Reconstruction”, In *Proc. IEEE Nuclear Science Symposium and Medical Imaging Conference*, volume 3, pp. 1749–1753, November 1993.
- [6] A. Blake and A. Zisserman, “*Visual Reconstruction*”, Artificial Intelligence, MIT Press, Cambridge, MA, 1987.