

Uncertainty Estimation for Bayesian Reconstructions from Low-Count SPECT Data¹

G.S. Cunningham and K.M. Hanson

MS P940, Los Alamos National Laboratory, Los Alamos, NM 87545

Abstract

Bayesian analysis is especially useful to apply to low-count medical imaging data, such as gated cardiac SPECT, because it allows one to solve the nonlinear, ill-posed, inverse problems associated with such data. One advantage of the Bayesian approach is that it quantifies the uncertainty in estimated parameters through the posterior probability. We compare various approaches to exploring the uncertainty in Bayesian reconstructions from SPECT data including: 1) the standard estimation of the covariance of an estimator using a frequentist approach, 2) a new technique called the "hard truth" in which one applies "forces" to the parameters and observes their displacements, and 3) Markov-chain Monte Carlo sampling of the posterior probability distribution, which in principle provides a complete uncertainty characterization.

I. INTRODUCTION

In a Bayesian formulation of the data analysis problem, one obtains the MAP estimate of the parameters by maximizing the posterior probability of model parameters given the measured data. The posterior probability can include prior information about model parameters, either physically-based or subjective, which can be critical for solving ill-posed problems such as limited-view tomography. The negative logarithm of the posterior probability, which is minimized, can be highly nonquadratic (due to non-Gaussianity of the noise, e.g.) or even possess more than one minimum (due to nonlinearities in the model, e.g.), and so iterative methods are employed for optimization.

We have implemented the Bayesian approach in a tool that we call the Bayes Inference Engine (BIE) to analyze image data acquired from 2D (and soon 3D) objects. This versatile and intuitive computer application allows one to develop complex geometric models for the objects under study, as well as complex models of the measurement process. The BIE permits one to compose a data-flow diagram (see Fig. 1) that produces a predicted image given a configuration of the object model, which could include geometric model(s) combined with models for variable intensities. Various aspects of the BIE are described elsewhere [1], [2], [3]. Geometric models have received increasing attention in medical imaging for tasks such as segmentation, reconstruction, restoration, and registration.

¹This work was supported by the United States Department of Energy under contract number W-7405-ENG-36.

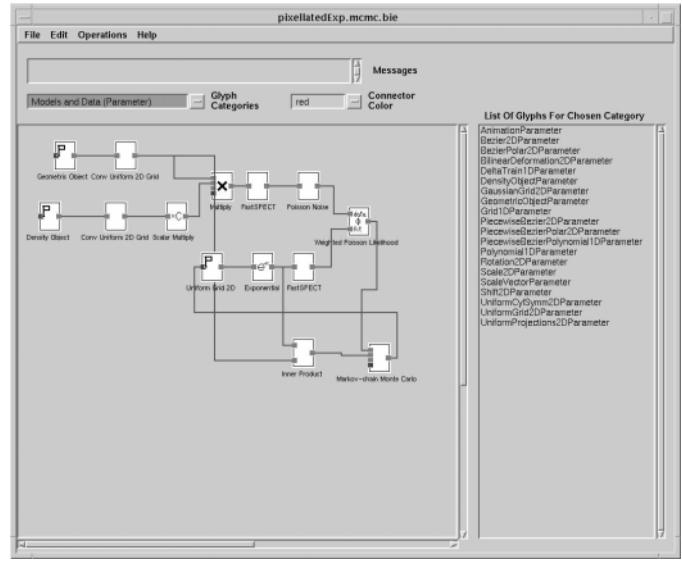


Figure 1: BIE canvas showing simulation of a 2D SPECT system.

In gated cardiac SPECT, a radiotracer is injected into the patient's circulatory system and travels to the heart, wherein disintegrations occur that produce photons which are individually detected. At certain instants in time, the radiotracer will perfuse one or the other of the chambers, and in this instant one can assume that the distribution of radiotracer will be relatively constant throughout a region defined by the walls of that chamber. Thus, a tomographic reconstruction of the radiotracer density should provide quantitative information about the location of the chamber walls. Normally, a reconstruction yields a voxellated image that must be segmented (often manually) to produce an estimate of the chamber volume. The ratio of the volumes at systoli and diastoli, called the ejection fraction, is an indicator of heart function.

In this article, we use 2D geometric models to reconstruct an area of constant, but unknown, radiotracer density from simulations of a low-count, limited-view, gated cardiac SPECT system. We estimate the contour of the area directly from the data using the BIE's modelling tools and gradient-based optimizer. Our near-term goal is to use 3D geometric models [4] to directly estimate volumes of (near) constant radiotracer density from real SPECT data. The small number of counts and limited number of views make gated cardiac SPECT ideally-suited for a Bayesian approach using geometric models. The computational cost associated with a Bayesian approach is similar to maximum likelihood approaches using simple voxellated models since the non-Gaussianity of the noise prohibits a closed-from solution to

the problem and iterative methods like maximum-likelihood expectation maximization (ML-EM) are used.

Ultimately, the quality of any reconstruction approach should be judged in terms of the bias and variance of the resultant estimator. In this paper, we use the BIE to apply three different approaches for quantifying the variance of estimators of radiotracer area using voxellated and geometric models: a frequentist approach, the “hard truth”, and Markov-Chain Monte Carlo (MCMC). We show that all three approaches agree with one another for the pixellated model, and that MCMC indicates that geometric models can reduce the variance of the area estimator if sufficient prior information can be applied to the problem regarding the smoothness of the 2D contour. We expect that the improvement in variance reduction will be even greater for 3D surface models.

II. BACKGROUND

A. The simulation of 2D gated cardiac SPECT

The data for this study were simulated using the real 2D system response of the University of Arizona’s FastSPECT machine [5]. The “2D” system response was measured over a 58x70 planar grid in object space with 2-mm voxel size by placing a 2-mm cube of radioactive tracer at the centers of each of the grid elements and measuring the response at each of the 24 detector grids (64x64 pixels). We selected the 27th row out of each of the first 13 detectors in an attempt to create a 2D problem for study. Thus, the system is modelled by a matrix \mathbf{H} that maps the mean intensities in a 58x70 object image \mathbf{f} into the mean intensities $\mathbf{g}=\mathbf{H}\mathbf{f}$ for each of 64 pixels in the 13 detectors (each detector pixel row was selected so that is approximately co-planar with the object slice). Systematic errors like scatter and attenuation were neglected for this study, but can be incorporated into a more complex model using the BIE.

The object slice was created by hand-tracing the contour for the right ventricle in an MRI reconstruction [6] and setting the pixel size so that the maximum width of the ventricle is about 6 cm. For this simulation, the radiotracer is assumed to be completely homogenous in density throughout the area defined by the contour, and the intensity per unit area is set so that the mean intensities at the detector yield an average of about 1/2 count per detector pixel, the average count level expected in a 20 msec frame for a real dynamic study. The total number of counts in the simulation used below is 488 (summed over a total of $64*13=832$ detector pixels).

B. Methods for quantifying uncertainty

1) Frequentist vs. Bayesian

For a linear, Gaussian data analysis problem, the uncertainty in the estimator or any linear functional of the

estimator is straightforward. Let $\mathbf{g}=\mathbf{H}\mathbf{f}+\mathbf{n}$, where \mathbf{f} is the object of interest, \mathbf{n} is additive, white Gaussian noise, and \mathbf{g} is the data. Further assume that \mathbf{H} is full-rank. The maximum likelihood estimator (MLE) for \mathbf{f} given \mathbf{g} is just $\hat{\mathbf{f}}=(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{g}$. If one is interested in a linear functional of \mathbf{f} , say $\mathbf{a}=\mathbf{h}^T\mathbf{f}$, then the MLE for \mathbf{a} is $\hat{\mathbf{a}}=\mathbf{h}^T\hat{\mathbf{f}}$. The covariance of $\hat{\mathbf{f}}$ is just $(\mathbf{H}^T\mathbf{H})^{-1}$. One can construct a confidence interval for $\hat{\mathbf{a}}$ that is independent of \mathbf{f} . That is, for every p one can find a number d such that the probability that $|\hat{\mathbf{a}}-\mathbf{a}|$ is greater than d is p , independent of \mathbf{f} . Note that the probabilistic variable here is $\hat{\mathbf{a}}$, not \mathbf{a} .

The Bayesian approach to this problem produces exactly the same result, but with a completely different interpretation. If one treats the likelihood as a posterior (with a uniform prior), then the posterior is Gaussian with covariance $(\mathbf{H}^T\mathbf{H})^{-1}$. Thus, one could construct a “credible interval” on $\hat{\mathbf{a}}$ and say that the probability that $|\hat{\mathbf{a}}-\mathbf{a}|$ is greater than d is p , for some value of p . The difference is that now \mathbf{a} is viewed as the probabilistic variable, not $\hat{\mathbf{a}}$. The estimator $\hat{\mathbf{a}}$ is fixed by the current realization of the data!

For a nonlinear, non-Gaussian problem though, confidence intervals do not, in general, exist. That is, the variance of the estimator does depend on the value of the real parameters \mathbf{f} , which are unknown. In this case, a Bayesian perspective is more satisfying, since one can still construct credible intervals on the parameters. One should take care, however, not to interpret credible intervals in the same way as confidence intervals.

2) The “hard truth”

We have introduced the concept that we call “the hard truth” and implemented it in the BIE [7]. The hard truth allows one to explore the uncertainty in complex models. This technique is understood in terms of an analogy between the negative log posterior and a physical potential. Near the MAP solution \mathbf{f}' , the minus log posterior can be approximated as a quadratic function of perturbations to the parameter set: $\phi=\phi'+.5(\mathbf{f}-\mathbf{f}')^T\mathbf{K}(\mathbf{f}-\mathbf{f}')$. Starting from \mathbf{f}' , the user can apply a “force” $c\mathbf{F}$ to the parameters by defining a new posterior $\phi=\phi'+.5(\mathbf{f}-\mathbf{f}')^T\mathbf{K}(\mathbf{f}-\mathbf{f}')-c\mathbf{F}^T(\mathbf{f}-\mathbf{f}')$, re-optimizing to obtain a new “MAP” solution \mathbf{f}'' and observe the displacements in \mathbf{f} as the solution converges to \mathbf{f}'' . The displacements $(\mathbf{f}''-\mathbf{f}')$ are related to the curvature of the minus-log-posterior in the limit as the force becomes infinitesimal (c goes to zero). That is, the response can be shown to be equal to the covariance matrix of a Gaussian approximation to the posterior times the force: $(\mathbf{f}''-\mathbf{f}')=c\mathbf{K}^{-1}\mathbf{F}$. If the value of the posterior at \mathbf{f}'' is ϕ'' , then we expect $\phi''-\phi'=5c^2\sigma^2$, where σ^2 is the variance of $\mathbf{F}^T\mathbf{f}$. Confidence intervals on linear measurements of the parameters ($\mathbf{F}^T\mathbf{f}$) can be made for the case in which the posterior is Gaussian. Specifically, $\sigma^2 = \mathbf{F}^T(\mathbf{f}''-\mathbf{f}')/c$. For the non-Gaussian case, approximate credible intervals can be obtained.

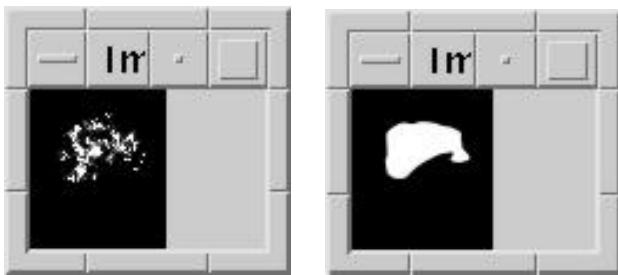
3) Markov-chain Monte Carlo

An alternative way to explore the uncertainties in a reconstruction and in quantities derived from it is to generate a sequence of random realizations drawn from the posterior. Such a sequence can be generated using the Markov-chain Monte Carlo (MCMC) technique in which one moves through a probability distribution by accepting or rejecting a proposed random step in the parameter space based on a Metropolis-Hastings algorithm[8]. This powerful approach allows one to study the full, marginalized probability distribution for any function of any combination of parameters.

III. ML-EM RECONSTRUCTION AND UNCERTAINTY

The traditional approach for analyzing gated cardiac SPECT data is to perform a voxel-based 3D reconstruction using maximum likelihood expectation-maximization (ML-EM) or a faster reconstruction technique like pre-processed filtered backprojection, and then to perform segmentation on the 3D volume to determine the 3D surfaces of the ventricles. The 3D surfaces define a volume which is measured at systoli and diastoli, and the ratio of the two volumes is computed to produce an estimate of the ejection fraction.

The ML-EM algorithm was employed for our 2D simulated data set and the result after 60 iterations is plotted in Fig. 2 alongside the original object. Increasing the number of iterations to 120 only decreased the log likelihood by 2.60 and did not noticeably change the appearance. Note that the ML-EM reconstruction has a very spiky appearance due to the fact that a zero detector pixel value strongly favors an estimate for the object intensity in which the path integral of the object intensity that produces that detector output is zero. Since the object intensity cannot be less than zero, a path integral of zero means that every pixel along the path must also be zero.



A) B)
Figure 2. A) ML-EM reconstruction after 60 iterations, B) original object, both plotted with a contrast of $[0,8e-4]$.

Ideally, what we'd like to do is incorporate the segmentation technique as an automatic subroutine so that we could apply the three techniques for quantifying uncertainty discussed above to the entire traditional process for estimating areas in the case of our 2D simulation. Unfortunately, much of the gated cardiac SPECT data is still segmented by hand, or is inherently 3D in nature, and so we

were unable to incorporate the segmentation into our results. The author's manual segmentation of the ML-EM estimate of intensity produced an estimate of the area that was in error by about 25% (too large).

An alternative to incorporating the segmentation into the uncertainty analysis is to compute the intensity estimated by ML-EM in the known area used for the simulation. This is the approach that we adopted. This approach will only tell us how wrong the reconstruction was in the sense of being too large, since, if all of the intensity in the reconstruction is placed inside the known area (but potentially in a much smaller area) then the “error” we calculate will be zero.

A. Uncertainty analysis using frequentist approach

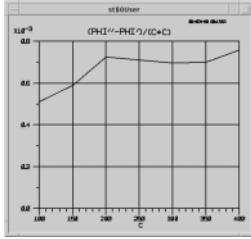
We generated 200 datasets with the same mean intensity as the simulation that produced 488 total counts (for 832 detector pixels) but with different realizations of the Poisson process. Each dataset was fed to the ML-EM algorithm with the initial guess for the object set at a constant value of 1.0 (the correct value is $8e-4$). 60 iterations of ML-EM were performed for each dataset, and the total intensity in the known area was computed. The mean of the 200 estimated total intensities was 0.411 (or 89.3% of the correct total intensity - 0.460) and the variance was $6.1e-4$, yielding a standard of deviation of 0.0247 (or 5.4% of the correct total intensity).

B. Uncertainty analysis using the “hard truth”

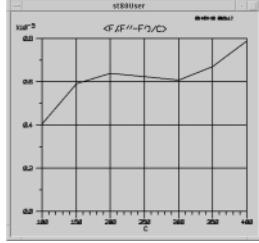
The “hard truth” was also applied to this dataset. There are two competing numerical issues in trying to apply the hard truth to a minus-log-posterior in the vicinity of the MAP solution: a) the constant c has to be chosen large enough to produce perturbations ($f' - f$) that yield substantial changes in the posterior, and b) we don't want to make c so large that we start exploring the non-Gaussian behavior of the posterior, at least for this example.

Since we are using a 60-iteration ML-EM as the MAP solution, and a 120-iteration ML-EM yields a change in minus log posterior $\phi' - \phi$ of about 2 units, we'd like to make the constant c large enough to produce $\phi' - \phi >> 2$ units. The relationship $\phi' - \phi = .5c^2\sigma^2$, where σ^2 is the variance of the intensity in the known area, can then give us an idea of what range to look at for c . In this case σ^2 was estimated using the frequentist approach as $6e-4$. This would indicate a value of $c=182.5$ in order to get a change in minus log posterior of about 10 units. Figure 3A shows the quantity $(\phi' - \phi)/c^2$, plotted for c in the range 100-400. Notice that the values at $c=100$ and 150 do not lie on the same flat line that the higher values do, most likely due to the numerical issue a) above. At a value $c=1000$, the optimizer essentially blew up. Notice also that the non-negativity constraint was still in place during these re-optimizations (which were done using 20-30 global steps of conjugate gradient), and that the value for σ^2 determined from Fig. 3A is about $7.0e-4$, within 15% of the value determined using the frequentist approach. However,

another measure of σ^2 can be determined from the relationship $\sigma^2 = F^T (f' - f)/c$. This quantity is plotted as a function of c in Fig. 3B. These values for σ^2 also vary by about 15% (6.1e-4 to 7.9e-4).



A)



B)

Figure 3. A) An indication of the quadratic behavior of the minus-log-posterior for the pixellated model: $(\phi'' - \phi)/c^2$, where c is the strength of the force and $\phi'' - \phi$ is the actual change in minus log posterior, and B) the variance of the estimator as defined by $\sigma^2 = F^T (f' - f)/c$.

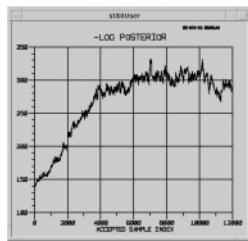
C. Uncertainty analysis using MCMC

Our first attempt at using MCMC on this problem failed miserably due to the non-negativity constraint. One cannot simply use a random walk to propose the next step and then project that on the constraint space (if the constraint space is nonlinear) before performing the accept/reject step.

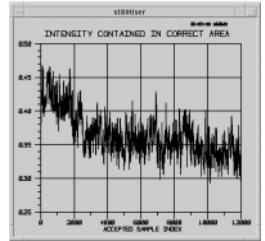
One quick fix for this problem is to re-parameterize the pixel values as the exponential of a set of unconstrained pixel values. Since the pixel values we are interested in are now just a simple function of another set of pixels values on which we can easily perform MCMC, our problem is solved.

We ran an MCMC chain of 25,000 trials on the re-parameterized model, of which 12,167 were accepted. The “burn-in” period was approximately the first 4000 accepted samples, as seen in Figs 4A and 4B. The sample mean of the intensity in the known area is 0.352 (23.4% error) and the sample variance is 4.4e-4 (so that the standard of deviation is 4.6% of the true area). The diagnostic program `gibbsit` [9] indicates that this run is adequate for obtaining the 2.5% quantile to reasonable accuracy.

The fact that the mean of the posterior is lower than the MAP estimate makes sense, since the MAP estimate will pin many pixel values to 0.0 intensity, leaving them with nowhere to go but up when the posterior is sampled. If intensity is put into these background pixels, it must be taken away from the area where the intensity is known to be (since the total number of counts in all of the detector pixels presumably determines the total intensity in the image to 3% or so). The lower sample standard of deviation may be due to an inadequate number of samples in MCMC, as the `gibbsit` diagnostic for the 80% quantile showed that only every 463rd accepted sample was independently drawn from the posterior!



A)



B)

Figure 4. A) Minus log posterior and B) intensity in known area for accepted samples drawn using MCMC for the exponentially reparameterized pixellated model.

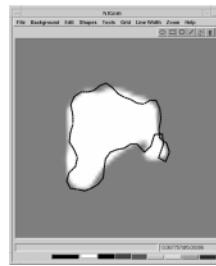
IV. GEOMETRIC MODELLING AND UNCERTAINTY

The BIE was used to construct a 50-vertex simple polygon model of the contour describing the area with constant intensity. A curvature prior [10] with user-manipulable strength was used to penalize non-smooth realizations of the contour.

The MAP solution for the curve using a strong prior is plotted in Fig. 5A, and the MAP solution for a weak prior (the strength is only 3% of that used for the strong prior) is plotted in Fig. 5B. Obviously, the smooth parts of the contour are reconstructed very well using the strong prior, while the high-curvature deviations from smoothness are missed completely. On the other hand, the reconstruction using the weak prior does marginally better at capturing the high-curvature parts of the true contour while doing substantially worse on the smooth parts.



A)



B)

Figure 5. The MAP estimate for the geometric model with prior strength equal to A) 1.0 and B) 0.03.

The MAP estimate for the area using the strong prior was in error by about 1% of the true area, which is better than one might guess is possible by simply looking at the signal-to-noise ratio in the total number of counts (the square root of 488 counts is 22.1, meaning there is a standard of deviation of roughly 5% in the total number of counts!). The MAP estimate for area using the weak prior was in error by about 10% of the true area (comparable to the pixellated model using ML-EM).

A) Uncertainty analysis

The frequentist approach to measuring the uncertainty in this case is difficult because the BIE’s optimizer requires

some user interaction when optimizing highly non-quadratic functions like the one obtained when using a geometric model [11]. Thus, as with segmentation, we can't just put the optimizer in a black box and run multiple realizations of the data through it in order to obtain sample statistics.

The hard truth was not attempted for this example, either, as we were not interested in exploring the higher-order correlations in the geometry that can be discovered using the "hard truth" [7], but rather in connecting the "hard truth" result to other methods of quantifying uncertainty, which we have already done for the pixellated model.

We did run MCMC on the geometric model using the very strong prior in order to determine how much better than a pixellated model it was for determining area. We ran 20,800 trials, of which approximately 6000 were accepted. The error in the sample mean of the area was 2.74% of the true area, and the sample standard of deviation was 3.26% of the true area. The `gibbsit` diagnostic indicated good convergence of the 2.5% quantile statistic with very little burn-in needed. Once again, though, the higher quantiles indicated a need to go to much larger samples to get the desired precision. For example, the 75% quantile indicated that only every 268th sample was drawn independently.

V. CONCLUSIONS

Geometric modelling has been shown to have excellent potential for improving the quality of ejection fraction estimates from low-count, gated cardiac SPECT data when strong prior information about the curvature of the chamber walls can be employed. The MCMC runs show a reduction in variance of about a factor of 2 over traditional pixellated models, while the MAP solutions showed a reduction in error of a factor of 10!

On the other hand, the weak prior geometric model performed quite poorly, and the MAP solution here still had a curvature value that was only about 50% of what the true contour had. If aspects of the heart wall that have high curvature are expected to be estimated well, then more counts or a better model of the wall that incorporates prior information about where the high curvature is spatially located, and where it is not, may be needed. We plan to use 3D surface models in the near future to estimate ejection fractions from real data using the FastSPECT system at the University of Arizona. The 3Dness of the surface models may help substantially in reducing the variance of the volume estimates over what might be expected by extrapolating this 2D study.

Three alternative methods for exploring the uncertainty in reconstructions from simulated 2D SPECT data have been demonstrated: frequentist, "hard truth", and MCMC. All three are in fair agreement for the pixellated model, although the MCMC showed a reduction in variance of about 30% relative to the other two. The agreement between the frequentist approach and the "hard truth" indicates a high

degree of Gaussianity for the problem of estimating the area of constant intensity from limited-view, low-count tomographic data. This is most likely due to the fact that the area is quite large, and so the Poisson counting statistics accumulate to an approximately Gaussian distribution. If one is interested in small areas, or specific contour location, it may be that the Gaussianity disappears.

VI. ACKNOWLEDGMENTS

We would like to thank Bob McKee and Marion Yapuncich for their contributions to the BIE; and Don Wilson, Michel Rogulski, and Harrison Barrett for the data on FastSPECT.

VII. REFERENCES

- [1] K. M. Hanson and G. S. Cunningham, "The Bayes inference engine," Maximum Entropy and Bayesian Methods, Dordrecht, Kluwer Academic, pp. 1-20, 1995 (to be published).
- [2] G. S. Cunningham, K. M. Hanson, G. R. Jennings, Jr., and D. R. Wolf, "An object-oriented optimization system," *Proc. IEEE Int. Conf. Image Processing*, vol.~III, pp. 826-830, 1994.
- [3] G. S. Cunningham, K. M. Hanson, G. R. Jennings, Jr., and D. R. Wolf, "An object-oriented implementation of a graphical-programming system," *Proc. SPIE*, vol. 2167, pp. 914-923, 1994.
- [4] X.L. Battle, K.M. Hanson, and G.S. Cunningham, "3D tomographic reconstructions using geometrical models," to be presented at SPIE Medical Imaging, February 1997.
- [5] M. Rogulski, An electronic poster is available at <http://www.radiology.arizona.edu/~fastspec>.
- [6] D.R. Thedens, D.J. Skorton, and S.R. Fleagle, "Methods of graph searching for border detection in image sequences with applications to cardiac magnetic resonance imaging," *IEEE Tran. Med. Imaging*, vol. 14, pp. 42-55, 1995.
- [7] K. M. Hanson and G. S. Cunningham, "Exploring the reliability of Bayesian reconstructions," *Proc. SPIE*, vol. 2434, pp. 416-423, 1995.
- [8] J. Besag, P. Green, D. Higdon and K. Mengersen, "Bayesian computation and stochastic systems", *Stat. Sci.*, vol. 10, pp. 3-66, 1995.
- [9] W.R. Gilks, S. Richardson, and D.J. Spiegelhalter, *Markov Chain Monte Carlo in Practice*, London, Chapman and Hall, 1996, chapter 7.
- [10] K.M. Hanson, R.L. Bilisoly, and G.S. Cunningham, "Kinky tomographic reconstruction," to appear in *proc. SPIE Medical Imaging*, 1996.
- [11] G.S. Cunningham, K.M. Hanson, and I.Koyfman, "Improved convergence of gradient-based reconstruction using multi-scale models," to appear in *proc. SPIE Medical Imaging*, 1996.