Noise Characterization of Block-Iterative Reconstruction Algorithms: II. Monte Carlo Simulations

Edward J. Soares*, *Member, IEEE*, Stephen J. Glick, *Member, IEEE*, and John W. Hoppin

***Abstract*—In Soares *et al.* (2000), the ensemble statistical properties of the rescaled block-iterative expectation-maximization (RBI-EM) reconstruction algorithm and rescaled block-iterative simultaneous multiplicative algebraic reconstruction technique (RBI-SMART) were derived. Included in this analysis were the special cases of RBI-EM, maximum-likelihood EM (ML-EM) and ordered-subset EM (OS-EM), and the special case of RBI-SMART, SMART. Explicit expressions were found for the ensemble mean, covariance matrix, and probability density function of RBI reconstructed images, as a function of iteration number. The theoretical formulations relied on one approximation, namely that the noise in the reconstructed image was small compared to the mean image. In this paper, we evaluate the predictions of the theory by using Monte Carlo methods to calculate the sample statistical properties of each algorithm and then compare the results with the theoretical formulations. In addition, the validity of the approximation will be justified.**

***Index Terms*—Covariance matrix, maximum-likelihood expectation-maximization, noise characterization, ordered subsets, rescaled block-iterative image reconstruction, simultaneous multiplicative algebraic reconstruction technique.**

I. Introduction

Part I of this paper [1], characterized the ensemble statistical properties of the rescaled block-iterative expectation-maximization (RBI-EM) reconstruction algorithm and rescaled block-iterative simultaneous multiplicative algebraic reconstruction technique (RBI-SMART). Also included in this analysis were the special cases of RBI-EM, maximum-likelihood EM (ML-EM) and ordered-subset EM (OS-EM), and the special case of RBI-SMART, SMART. Obtained were explicit expressions that approximate the ensemble mean, covariance matrix, and probability density function (PDF) for images reconstructed using the aforementioned algorithms. The theoretical formulations relied on one approximation, namely that the noise in the reconstructed image was small compared to the mean image.

The theoretical approach taken in Part I [1], which facilitates our derivation of analytic expressions for the ensemble statistical parameters for the aforementioned algorithms, was first motivated by the work of Barrett *et al.* [2], [3] for the ML-EM algorithm. Their approach in [2] was to linearize the ML-EM algorithm in log-image space, which then allowed for the computation of the approximate ensemble statistical parameters. Subsequently, they confirmed the validity of their approach in [3] using Monte Carlo methods. Researchers have also investigated the noise properties for regularized versions of the EM algorithm, including Fessler [4] and Wang and Gindi [5], [6].

Knowledge of the ensemble statistical properties of block iterative reconstructed images is important for several reasons. For example, human observers tend to exhibit decreased performance in signal detection tasks with images containing certain types of correlated noise [7]. Information regarding the noise correlation structure, in the form of the ensemble covariance matrix, would then be helpful in predicting such decreased performance. In addition, the ensemble mean and covariance matrix are useful for image quality evaluation, particularly for computing model observer signal-to-noise ratios for detection or estimation tasks using mathematical model observers [8], [9]. In some cases, image reconstruction algorithms such as OS-EM have been evaluated using image quality metrics derived from sample statistical quantities [10]. However, an adequate sample size must be used for estimating the sample statistical parameters, in order to accurately estimate their respective ensemble counterparts. This is especially impractical when the dimensionality of the feature vector is large or if the inverse of the entire image covariance matrix is needed. In such cases, it is necessary to know the ensemble statistical parameters of the reconstructed images.

In this paper, the theoretical formulations derived in Part I [1] will be evaluated by comparing the approximate ensemble statistical parameters with their respective sample counterparts. Using Monte Carlo methods, we will calculate sample estimates of the mean, covariance matrix, and point PDF of RBI reconstructed images for various noise levels, iteration points, and subset orderings, for purposes of comparison. In addition, the approximation that serves as the basis for the theoretical formulations in Part I [1] will be justified to determine when the theory break down. Thus, we will ascertain under what conditions the theoretical formulations accurately predict the true ensemble statistical properties.

Manuscript received July 14, 2004; revised August 18, 2004. This work was supported in part by the National Cancer Institute under Grant 78573 and in part by the National Heart, Lung, and Blood Institute under Grant HL 50349. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Cancer Institute nor the National Heart, Lung, and Blood Institute. The Associate Editor responsible for coordinating the review of this paper and recommending its publication was M. Viergever. *Asterisk indicates corresponding author.*

*E. J. Soares is with the Department of Mathematics and Computer Science, College of the Holy Cross, Worcester, MA 01610 USA (e-mail: esoares@holycross.edu).

S. J. Glick is with the Division of Nuclear Medicine, University of Massachusetts Medical School, Worcester, MA 01655 USA (e-mail: Stephen.Glick@umassmed.edu).

J. W. Hoppin was with the Department of Mathematics and Computer Science, College of the Holy Cross, Worcester, MA 01610 USA. He is now with the Central Institute for Electronics Research Center Jülich, 52425 Jülich, Germany (e-mail: j.hoppin@fz-juelich.de).

Digital Object Identifier 10.1109/TMI.2004.836876

Following the notation used in Part I [1], we shall denote vectors with lower-case bold letters, matrices with upper-case bold letters, and scalars with lower-case letters.

II. Methods

A. Sample Image Statistics

In order to test the predictions of the theory outlined in Part I [1], Monte Carlo techniques were employed. By generating a large sample of noisy image reconstructions, we were able to calculate sample estimates of the mean, covariance matrix, and point PDF for the RBI-EM, OS-EM, and RBI-SMART algorithms. As the sample size increases, the sample statistical measures better estimate their true ensemble counterparts. Therefore, by comparing the approximate ensemble and sample statistical properties, one can determine to what degree the theoretical formulations accurately predict the true ensemble statistical properties of the RBI reconstructed images.

Let us denote the nth noisy two-dimensional (2-D) image estimate at the kth iteration by $\hat{\mathbf{f}}^{(k),n}$, where the reconstructed image has been lexicographically ordered into a $J \times 1$ column vector and N is the sample size. Thus, $\{\hat{\mathbf{f}}^{(k),n}\}_{n=1}^{N}$ is the sample of N noisy reconstructions used to determine the sample statistical parameters at the kth iteration. By definition, the sample mean image $\bar{\hat{\mathbf{f}}}^{(k)}$ is given by

$$\bar{\hat{\mathbf{f}}}^{(k)} = \frac{1}{N}\sum_{n=1}^{N}\hat{\mathbf{f}}^{(k),n}. \tag{1}$$

Also, the unbiased form of the sample covariance matrix $\mathbf{S}_{\hat{\mathbf{f}}^{(k)}}$ is given by

$$\mathbf{S}_{\hat{\mathbf{f}}^{(k)}} = \frac{1}{N-1}\sum_{n=1}^{N}\left(\hat{\mathbf{f}}^{(k),n} - \bar{\hat{\mathbf{f}}}^{(k)}\right)\left(\hat{\mathbf{f}}^{(k),n} - \bar{\hat{\mathbf{f}}}^{(k)}\right)^{t}. \tag{2}$$

Furthermore, from this set of N noisy images, one can also compute a normalized histogram for any particular RBI reconstructed image pixel, which could be used to estimate the univariate point PDF for that given pixel.

When J is large, as is usually the case for clinical images, it is impossible to examine the entire covariance matrix for purposes of comparison. Thus, we will only consider selected parts of the covariance matrix when making our evaluations, namely the variance and local covariance images as described by Wilson *et al.* [3]. The variance image is constructed by ordering the diagonal elements of the covariance matrix into a 2-D image. Since the noise is nonstationary in single photon emission computed tomography [11], the correlation structure changes with pixel location. Therefore, the entire covariance matrix is needed to describe the noise structure at every pixel location. However, one can use the definition of the local covariance [3], [12] to examine the noise properties at a particular image pixel location. Based upon our lexicographic ordering of the original 2-D object into a $J \times 1$ vector, the local covariance at pixel j is determined by unlexicographically ordering row (or column) j of the covariance matrix into a 2-D image.

B. Ensemble Image Statistics

Part I [1] derived analytic expressions for the ensemble mean, covariance matrix, and PDF for the RBI methods. This derivation was accomplished by assuming a low-noise approximation, that states that the noise in the reconstructed image is small compared to the mean image. By linearizing the RBI algorithms in log-image space, they were able to show that the noise in the logarithm of the reconstructed image could be expressed as a linear transformation of the noise in the projection data. With knowledge of the noise-free projection data, one could then calculate the ensemble statistical properties for images reconstructed with any of the RBI algorithms.

For comparison with their respective sample counterparts, we computed the approximate ensemble mean, covariance matrix, and point PDF for the RBI-EM, OS-EM, and RBI-SMART algorithms, using [1, eqs. (87), (88), and (90)].

Fig. 1. Simulated images of the uniform disk object (left), MCAT activity map (center), and MCAT attenuation map (right) used in our study.

C. Simulations

In Fig. 1, we show the two objects used in our evaluations. The first is a uniformly emitting disk of radius 13 pixels, where the image size is 32×32 pixels. The second object consists of a single 128×128 slice through an anthropomorphic phantom called the MCAT [13], along with corresponding nonuniform attenuation map. The disk object will be used in our comprehensive evaluation, whereas the results obtained using the MCAT are meant to demonstrate the effectiveness of the theory for a more complex object and more realistic imaging scenario. For the disk object, noise-free projection data of dimension 32×32 were created using an analytic ray-driven projector [14] based on Siddon's ray-tracing algorithm [15]. Nonuniform photon attenuation, distance-dependent detector response, and scatter were not modeled in the simulation.

A total of 10 000 noisy data sets, for each of three noise levels, were generated by adding pseudorandom Poisson noise to the noise-free projection data. Considered were a high signal-to-noise case of 1 500 000 counts, a moderate signal-to-noise case of 500 000 counts, and a low signal-to-noise case of 50 000 counts. These levels were selected based upon the previous work of Wilson *et al.* [3]. The data sets were then reconstructed with each of the three aforementioned algorithms.

The algorithms all rely on the subset orderings of the data and the stopping point in the iteration process. We chose to examine the cases of 4 subsets (8 angles per subset), 8 subsets (4 angles per subset), and 16 subsets (2 angles per subset). These were selected based upon the fact that faster acceleration is achieved with increasing numbers of subsets. Not all possible combinations of noise level and subset ordering were considered. Instead, we chose to fix one parameter and vary the other. The noise level-subset ordering pairs considered are listed in Table I.

Additionally, the stopping points in the iterative process considered were 16, 32, 64, 128, 256, 512, and 1024 iterations. As discussed in Part I [1], we consider one iteration of a block-iterative method to be a single pass through one subset of the data. Most authors regard one iteration as a complete pass through

TABLE I
THE VARIOUS COMBINATIONS OF NOISE LEVEL AND SUBSET ORDERING USED IN OUR ANALYSIS OF THE RBI ALGORITHMS

	50K counts	500K counts	1,500K counts
4 subsets		√	
8 subsets	√	√	√
16 subsets		√	

all of the subsets, which we call a "full iteration." To maintain consistency with our previous convention, we shall present all results in terms of full iterations of each algorithm. Thus, the aforementioned stopping points correspond to different full iterations, depending on how many subsets are used.

For the MCAT simulation, noise-free projection data of dimension 128×128 were created using the same ray-driven projector which modeled the effects of nonuniform photon attenuation. From this, we generated 1000 noisy data sets by adding pseudorandom Poisson noise to the noise-free projection data. A single count level of 500 000 counts was considered. The data sets were reconstructed with RBI-EM, OS-EM, and RBI-SMART, using 8 subsets (16 angles per subset) and stopping at iteration 128 (16 full iterations).

D. Measures

In our evaluations of the theoretical formulations, we attempted to quantify the discrepancies between the theoretical and Monte Carlo estimates of the true ensemble statistical parameters, using measures described in Wilson *et al.* [3]. One measure used was the root mean-square (RMS) percent error, defined to be

$$\text{RMS \% error} = \frac{100}{\sqrt{L}} \sqrt{\sum_{l=1}^{L} \left(\frac{\hat{\text{stat}}_l^{\text{mc}} - \hat{\text{stat}}_l^{\text{th}}}{\hat{\text{stat}}_l^{\text{mc}}} \right)^2}. \quad (3)$$

Due to the fact that the RMS measure contains errors from the use of the Monte Carlo estimate to represent the true ensemble statistic, we also employed the average (AVG) percent bias, defined to be

$$\text{AVG \%bias} = \frac{100}{L} \sum_{l=1}^{L} \frac{\hat{\text{stat}}_l^{\text{mc}} - \hat{\text{stat}}_l^{\text{th}}}{\hat{\text{stat}}_l^{\text{mc}}}. \quad (4)$$

In either measure, the quantity *stat* can either represent the *mean* or *variance*, depending on the statistic being compared. With this in mind, $\hat{\text{stat}}_l^{\text{mc}}$ refers to the Monte Carlo estimate of the particular statistic at pixel l, while $\hat{\text{stat}}_l^{\text{th}}$ refers to the analogous theoretical estimate. Both the RMS % error and AVG % bias were quantified in a region-of-interest within the interior of the disk that contained L total pixels.

III. RESULTS

A. Probability Distributions

In Fig. 2, we compare the approximate ensemble and sample point PDF's for pixel (16, 16), which is near the center of the disk object, for RBI-EM, OS-EM and RBI-SMART reconstructed images. The projection data contained 500 000 total counts and were partitioned into 8 subsets. Shown are sample point PDF's generated from 10 000 noisy realizations for 2, 16, and 128 full iterations of each algorithm. The lines indicate the predicted values of the log-normal point PDF to the sample point PDF using [1, eq. (90)].

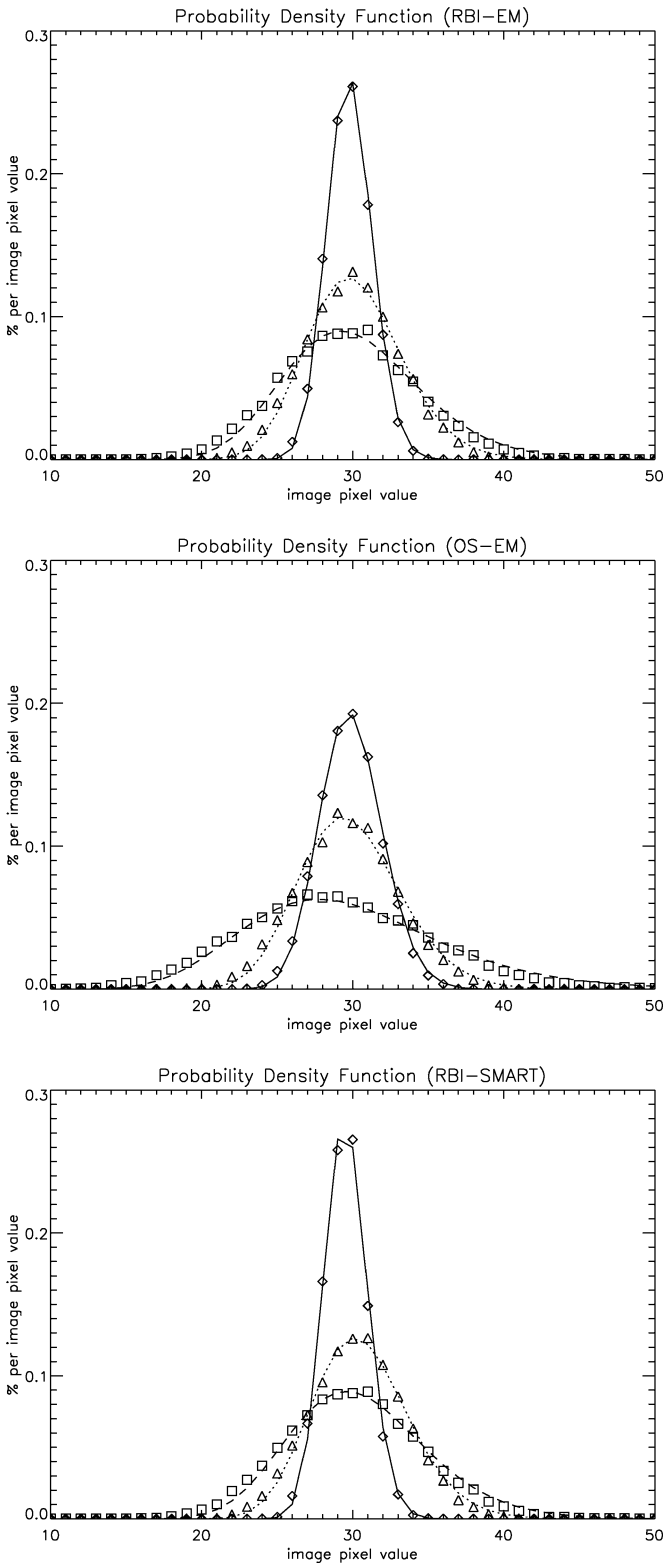

Fig. 2. Probability density functions (PDF's) for a pixel near the center of the disk for the RBI-EM (top), OS-EM (center), and RBI-SMART (bottom) algorithms. Shown are sample PDF's for 2 (◇), 16 (△), and 128 (□) full iterations of each algorithm. The lines indicate the predicted values of the log-normal point PDF to the sample PDF. The projection data contained 500 000 total counts and were partitioned into 8 subsets.

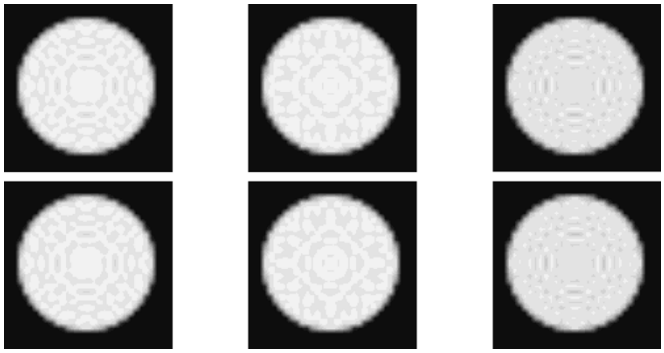

Fig. 3. Ensemble mean (top row) and sample mean (bottom row) images of the disk object for the RBI-EM (left), OS-EM (center), and RBI-SMART (right) algorithms. Both ensemble and sample mean images were generated for 16 full iterations of each algorithm. The projection data contained 500 000 total counts and were partitioned into 8 subsets.

We can see that the approximate ensemble point PDF's show excellent agreement with their respective sample counterparts, at each iteration point. However, slight deviations between the sample points and solid curves can be seen for higher iterations. This is the effect of the low-noise approximation becoming less valid, as the noise is no longer small compared to the mean image. Also, the respective point PDF's becomes wider, exhibiting an increase in pixel variance. This effect is due to the buildup of noise in the data as iteration increases.

B. Comparison Between Ensemble and Sample Parameters

1) Mean Image: The results comparing the approximate ensemble and sample means are shown in Figs. 3 and 4, and Table II. In Fig. 3, we show images of the approximate ensemble mean and sample mean for the three algorithms. The object considered was the uniform disk with 500 000 total counts in the projection data, which were partitioned into 8 subsets. The images correspond to full iteration 16. Fig. 4 illustrates central profiles through the respective images in Fig. 3. In Table II, we show estimation errors between the approximate ensemble and sample mean images as a function of full iteration number for the three algorithms.

Examining Figs. 3 and 4, we qualitatively note the excellent agreement between the respective approximate ensemble and sample means. Such agreement at the other iteration points can also be inferred by examining Table II. We note that at each iteration, the RMS % error is less than 2% and the AVG % bias is less than 1% in all cases. These measures were also computed (although not shown) for the case of the 500 000 count data partitioned into 4 subsets and 16 subsets. In the 4-subset case, the RMS % error and AVG % bias were all less than 1% at each iteration. In the 16-subset case, the RMS % error and AVG % bias were all less than 3% at each iteration. Considering the case of the 1 500 000 count data partitioned into 8 subsets, the RMS % error and AVG % bias were all less than 1% at each iteration.

However, when considering the 50 000 count data partitioned into 8 subsets, the RMS % error and AVG % bias were higher. The RBI-EM algorithm yielded errors less than 2% at each iteration, the errors for the OS-EM algorithm were less than 4% and for RBI-SMART, less than 8%. This increase in error is due to the fact that the low-noise approximation is less valid when the data are low-count.

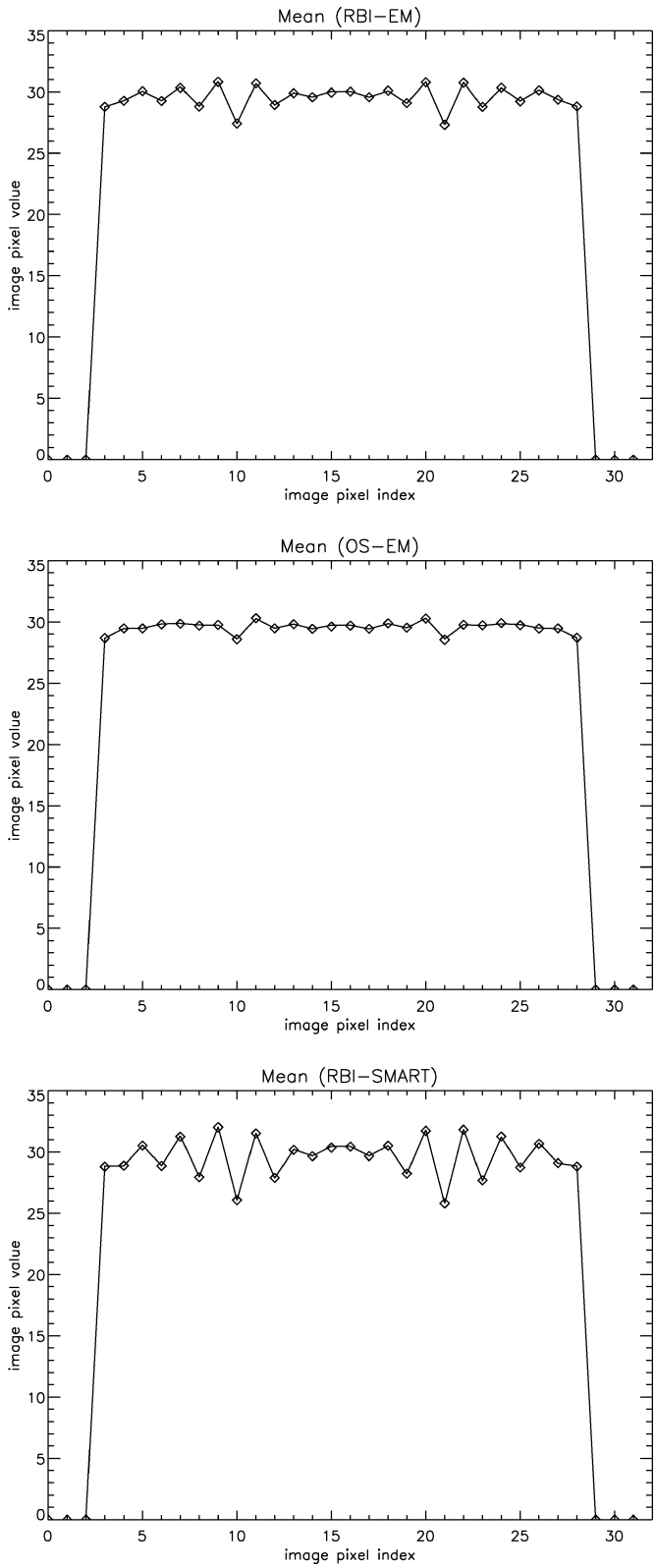

Fig. 4. Central profiles through the ensemble mean (solid lines) and sample mean (◇) images of the disk object, for the RBI-EM (top), OS-EM (center), and RBI-SMART (bottom) algorithms, as shown in Fig. 3.

2) Variance Image: The results comparing the approximate ensemble and sample variance images are shown in Figs. 5 and 6, and Tables III–VII. In Fig. 5, we show images of the approximate ensemble variance and sample variance for the three

TABLE II
ESTIMATION ERRORS BETWEEN THE APPROXIMATE ENSEMBLE AND SAMPLE MEANS AS A FUNCTION OF ITERATION NUMBER FOR THE RBI-EM, OS-EM AND RBI-SMART ALGORITHMS. THE PROJECTION DATA CONTAINED 500 000 TOTAL COUNTS, AND WERE PARTITIONED INTO 8 SUBSETS

	RBI-EM		OS-EM		RBI-SMART	
full iter. no.	RMS % error	AVG % bias	RMS % error	AVG % bias	RMS % error	AVG % bias
2	0.05	0.00	0.07	0.00	0.06	-0.04
4	0.07	0.00	0.09	0.00	0.09	-0.03
8	0.09	0.00	0.11	0.00	0.11	-0.03
16	0.11	0.00	0.15	0.00	0.13	-0.03
32	0.13	0.00	0.25	0.00	0.18	-0.03
64	0.21	0.01	0.50	-0.03	0.28	-0.03
128	0.39	0.00	1.04	-0.11	0.47	-0.03

TABLE III
ESTIMATION ERRORS BETWEEN THE APPROXIMATE ENSEMBLE AND SAMPLE VARIANCES AS A FUNCTION OF ITERATION NUMBER FOR THE RBI-EM, OS-EM AND RBI-SMART ALGORITHMS. THE PROJECTION DATA CONTAINED 1 500 000 TOTAL COUNTS, AND WERE PARTITIONED INTO 8 SUBSETS

	RBI-EM		OS-EM		RBI-SMART	
full iter. no.	RMS % error	AVG % bias	RMS % error	AVG % bias	RMS % error	AVG % bias
2	1.52	-0.01	1.51	-0.02	1.51	0.01
4	1.50	-0.03	1.50	-0.06	1.51	-0.00
8	1.49	-0.06	1.46	-0.08	1.49	-0.03
16	1.45	-0.04	1.41	-0.10	1.46	-0.01
32	1.42	-0.04	1.38	-0.22	1.42	-0.00
64	1.40	-0.15	1.79	-1.02	1.40	-0.08
128	1.48	-0.46	7.50	-6.75	1.46	-0.38

TABLE IV
ESTIMATION ERRORS BETWEEN THE APPROXIMATE ENSEMBLE AND SAMPLE VARIANCES AS A FUNCTION OF ITERATION NUMBER FOR THE RBI-EM, OS-EM AND RBI-SMART ALGORITHMS. THE PROJECTION DATA CONTAINED 500 000 TOTAL COUNTS AND WERE PARTITIONED INTO 8 SUBSETS

	RBI-EM		OS-EM		RBI-SMART	
full iter. no.	RMS % error	AVG % bias	RMS % error	AVG % bias	RMS % error	AVG % bias
2	1.44	-0.11	1.47	-0.17	1.44	-0.01
4	1.46	-0.20	1.48	-0.28	1.45	-0.09
8	1.49	-0.30	1.48	-0.25	1.47	-0.19
16	1.49	-0.25	1.50	-0.19	1.47	-0.13
32	1.49	-0.26	1.67	-0.61	1.48	-0.08
64	1.61	-0.66	3.88	-3.31	1.54	-0.42
128	2.32	-1.72	17.77	-16.52	2.14	-1.44

TABLE V
ESTIMATION ERRORS BETWEEN THE APPROXIMATE ENSEMBLE AND SAMPLE VARIANCES AS A FUNCTION OF ITERATION NUMBER FOR THE RBI-EM, OS-EM AND, RBI-SMART ALGORITHMS. THE PROJECTION DATA CONTAINED 50,000 TOTAL COUNTS AND, WERE PARTITIONED INTO 8 SUBSETS

	RBI-EM		OS-EM		RBI-SMART	
full iter. no.	RMS % error	AVG % bias	RMS % error	AVG % bias	RMS % error	AVG % bias
2	1.48	0.21	1.54	0.09	2.38	1.53
4	1.55	-0.25	1.72	-0.66	2.88	1.19
8	1.82	-0.93	1.69	-0.16	4.26	1.17
16	1.72	-0.45	1.90	-0.14	5.77	2.87
32	2.10	-0.69	5.63	-4.70	6.66	3.77
64	5.88	-4.95	23.70	-22.55	6.69	0.20
128	16.99	-16.10	80.99	-77.92	14.05	-10.80

algorithms. The object considered was the uniform disk with 500 000 total counts in the projection data, which were partitioned into 8 subsets. The images correspond to full iteration 16. Fig. 6 illustrates central profiles through the respective images in Fig. 5. In Tables III–VII, we show estimation errors between the approximate ensemble and sample variance images as a function of full iteration number for the three algorithms. Tables III, IV, and V reflect results when the number of subsets was fixed at 8 and the count level decreased, while Tables VI, IV and VII reflect results when the count level was fixed at 500 000 counts and the number of subsets increased.

Let us first consider the results when the data partitioning is fixed at 8 subsets and the noise level decreases from the high-count case of 1 500 000 to the low-count case of 50 000 counts. Subsequently, we will evaluate the results when the noise level is fixed at a moderate count level of 500 000 counts and the data partitioning increases from 4 subsets to 16 subsets.

Examining Figs. 5 and 6, we again qualitatively note the strong agreement between the respective approximate ensemble and sample variance images. Such agreement is typical for low iterations and moderate to high count levels. As with the mean images, deviations occur when iteration increases or when the

TABLE VI
ESTIMATION ERRORS BETWEEN THE APPROXIMATE ENSEMBLE AND SAMPLE VARIANCES AS A FUNCTION OF ITERATION NUMBER FOR THE RBI-EM, OS-EM AND RBI-SMART ALGORITHMS. THE PROJECTION DATA CONTAINED 500 000 TOTAL COUNTS AND, WERE PARTITIONED INTO 4 SUBSETS

	RBI-EM		OS-EM		RBI-SMART	
full iter. no.	RMS % error	AVG % bias	RMS % error	AVG % bias	RMS % error	AVG % bias
4	1.45	-0.14	1.46	-0.17	1.44	-0.02
8	1.47	-0.23	1.48	-0.28	1.46	-0.12
16	1.50	-0.31	1.49	-0.29	1.48	-0.20
32	1.50	-0.24	1.49	-0.21	1.48	-0.10
64	1.51	-0.30	1.55	-0.42	1.49	-0.11
128	1.70	-0.82	1.98	-1.22	1.61	-0.56
256	2.63	-2.06	3.85	-3.35	2.43	-1.77

TABLE VII
ESTIMATION ERRORS BETWEEN THE APPROXIMATE ENSEMBLE AND SAMPLE VARIANCES AS A FUNCTION OF ITERATION NUMBER FOR THE RBI-EM, OS-EM AND, RBI-SMART ALGORITHMS. THE PROJECTION DATA CONTAINED 500 000 TOTAL COUNTS AND WERE PARTITIONED INTO 16 SUBSETS

	RBI-EM		OS-EM		RBI-SMART	
full iter. no.	RMS % error	AVG % bias	RMS % error	AVG % bias	RMS % error	AVG % bias
1	1.47	-0.09	1.47	-0.05	1.45	0.01
2	1.47	-0.17	1.49	-0.24	1.45	-0.07
4	1.48	-0.28	1.46	-0.15	1.46	-0.18
8	1.48	-0.23	1.47	-0.13	1.47	-0.11
16	1.48	-0.22	1.60	-0.47	1.48	-0.05
32	1.59	-0.59	2.28	-1.50	1.53	-0.35
64	2.27	-1.63	4.33	-3.62	2.09	-1.35

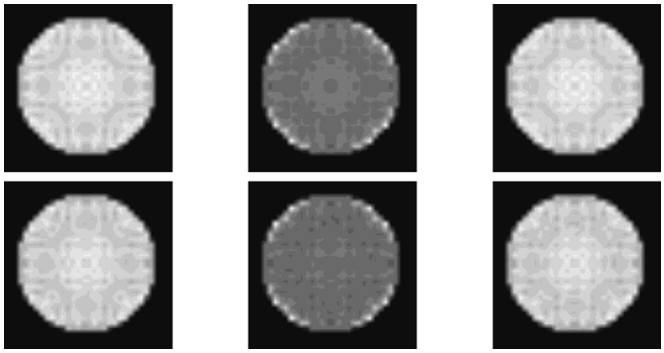

Fig. 5. Ensemble variance (top row) and sample variance (bottom row) images of the disk object for the RBI-EM (left), OS-EM (center), and RBI-SMART (right) algorithms. The ensemble and sample variance images were generated for 16 full iterations of each algorithm. The projection data contained 500 000 total counts and were partitioned into 8 subsets.

low-noise approximation is violated. This agreement at the other iteration points can be seen by examining Tables III–V.

When the data are high count (Table III), the RMS % errors are all less than 2% for each algorithm, with the exception of OS-EM, which shows an error exceeding 7% for iteration 128. Likewise, the AVG % bias is less than 2% for each algorithm, with the exception of OS-EM, which shows an AVG % bias exceeding 6% at iteration 128. One explanation for this increase in error for OS-EM relative to the other algorithms is that OS-EM exhibits a higher degree of acceleration than do RBI-EM or RBI-SMART for the same iteration point.

For moderate count data (Table IV), we note that the RMS % errors are all less than 4% for each algorithm, with the exception of OS-EM, which shows an error exceeding 17% for iteration 128. Likewise, the AVG% bias is less than 4% for each algorithm, with the exception of OS-EM, which shows an AVG % bias exceeding 16% at iteration 128. Again, this is due to the fact that OS-EM provides a greater degree of acceleration than the other two algorithms for the same iteration point.

When the data are low count (Table V), we can see that the RMS % error and AVG % bias all exhibit a marked increase, in relation to the previous results. Although these errors are generally on the order of 6% or less for most iteration numbers, we see that iterating too far yields significant errors and so the approximate ensemble variance is not a good approximation to the true ensemble variance, which we estimate using the sample variance. We again note that OS-EM yields a much higher RMS % error and AVG % bias for iteration 128 than do RBI-EM and RBI-SMART, although the results for all three algorithms are unacceptable.

Now, let us consider our findings when the count level is fixed at 500 000 counts and the number of subsets increases. As previously mentioned, all of the results were computed for stopping points 16, 32, 64, 128, 256, 512, and 1024 in the iterative process. These will correspond to different full iterations, depending on how many subsets are used. In the 4-subset case, these points correspond to full iterations 4, 8, 16, 32, 64, 128, and 256, respectively. For the 8-subset configuration, they correspond to full iterations 2, 4, 8, 16, 32, 64, and 128, while in the 16-subset configuration, they correspond to full iterations 1, 2, 4, 8, 16, 32, and 64. This fact will be important when we compare the estimation errors at the same stopping point.

In the 4-subset case (Table VI), the RMS % errors were all less than 2% and the AVG % biases were all less than 1%, up to and including iteration 64, for each algorithm. The same is true for the 8-subset case (Table IV), with the exception of the OS-EM algorithm at iteration 64, which yields an RMS % error and AVG % bias of approximately 4%. When the data are partitioned into 16 subsets (Table VII), the RMS % errors were all less than 5% and the AVG % biases were all less than 4%, up to

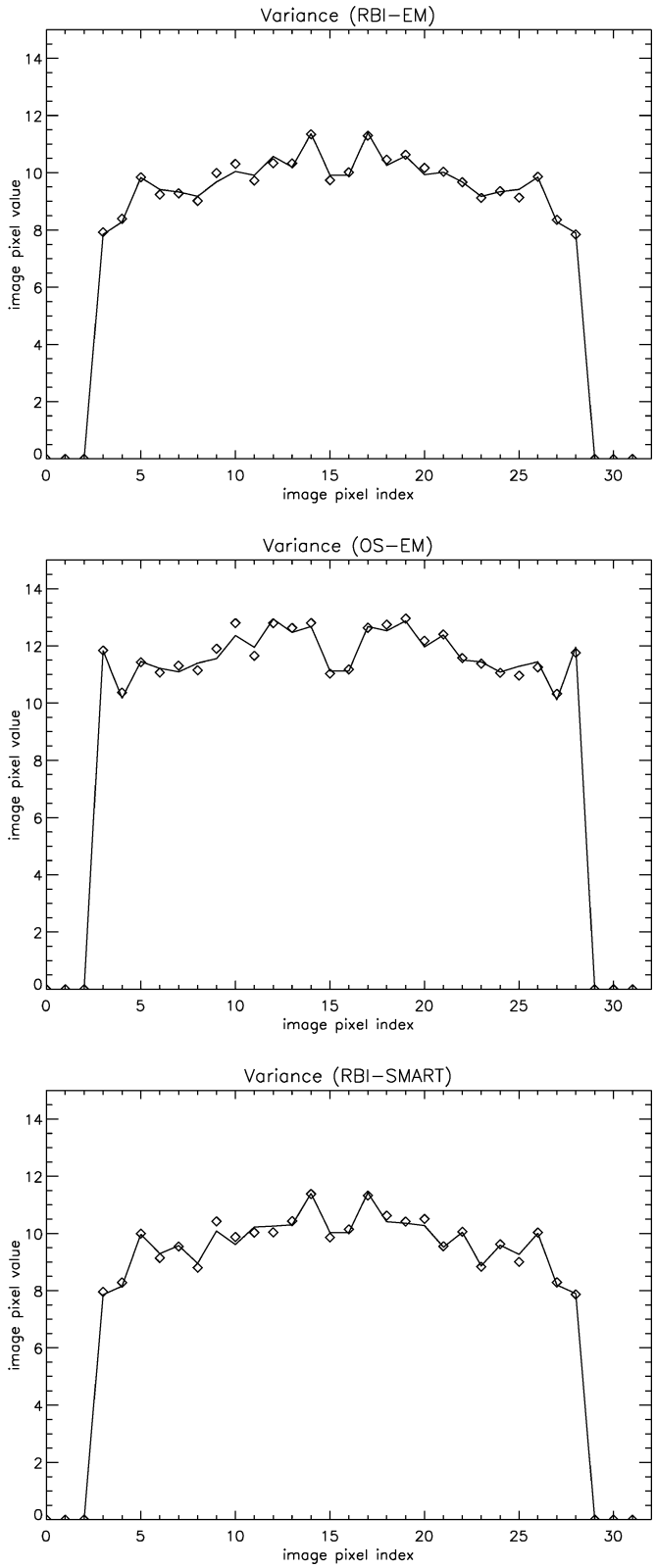

Fig. 6. Central profiles through the ensemble variance (solid lines) and sample variance ($\diamond$) images of the disk object, for the RBI-EM (top), OS-EM (center), and RBI-SMART (bottom) algorithms, as shown in Fig. 5.

and including iteration 64, for each algorithm. As the number of subsets increases, faster acceleration is seen for each algorithm.

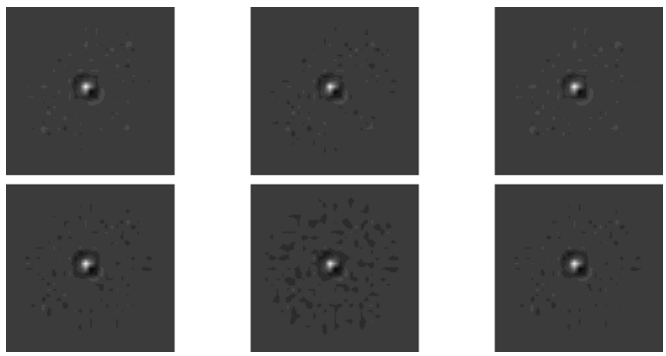

Fig. 7. Ensemble (top row) and sample (bottom row) local covariance images of the disk object for the RBI-EM (left), OS-EM (center), and RBI-SMART (right) algorithms. The ensemble and sample local covariance images were generated for 16 full iterations of each algorithm. The projection data contained 500 000 total counts and were partitioned into 8 subsets.

This increase impacts the reconstructed image by increasing noise variance, which subsequently violates the low-noise approximation. Thus, we would expect to see increased RMS % errors and AVG % bias when the noise level is fixed and the number of subsets increases.

3) Local Covariance Image: The results comparing the approximate ensemble and sample local covariance images are shown in Figs. 7 and 8. In Fig. 7, we show images of the approximate ensemble and sample local covariance for pixel (16,16), for the three algorithms. The object considered was the uniform disk with 500 000 total counts in the projection data, which were partitioned into 8 subsets. The images correspond to full iteration 16. Fig. 8 illustrates central profiles through the respective images in Fig. 7.

Since the covariance matrix is of dimension 1024×1024, it is impossible to display and analyze results for all other image locations. The results presented herein are meant to give the reader some perspective on the quality of the agreement that might be seen between the approximate ensemble and sample local covariance images, for a given pixel location.

Examining Figs. 7 and 8, we again qualitatively note the excellent agreement between approximate ensemble and sample local covariance images. As was the case for the mean and variance images, such agreement is typical for low iterations and moderate to high count levels. Again, deviations typically occur when iteration increases or when the data are low-count.

C. MCAT Simulation

The results to illustrate the validity of the theoretical formulation for a more complex object and realistic imaging scenario are presented in Figs. 9–12. We compare the approximate ensemble and sample means in Figs. 9 and 10, which show images and x-axis profiles, respectively, of the approximate ensemble mean and sample means for the three algorithms, while Figs. 11 and 12 show images and x-axis profiles, respectively, through the approximate ensemble and sample variance images. The projection data contained 500 000 total counts, which were partitioned into 8 subsets. The images correspond to full iteration 16.

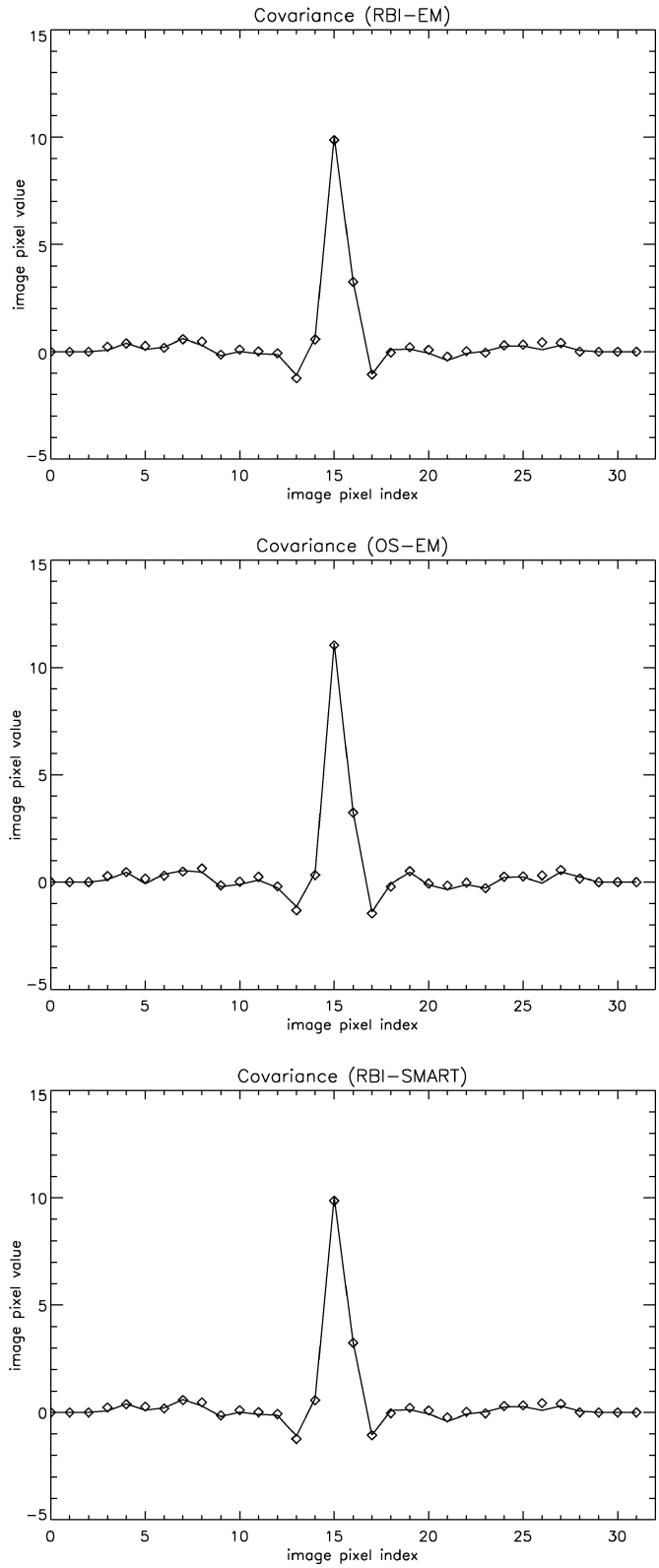

Fig. 8. Central profiles through the ensemble local covariance (solid lines) and sample local covariance ($\diamond$) images of the disk object, for the RBI-EM (top), OS-EM (center), and RBI-SMART (bottom) algorithms, as shown in Fig. 7.

Considering Figs. 9–12, we see that the theory well-predicts the true ensemble statistical parameters. Both high and low regions of activity are accounted for when comparing approximate ensemble and sample means. In addition, high and low regions of noise variance are accurately predicted, although the reader will note slight discrepancies between the approximate ensemble and sample variances in Fig. 12.

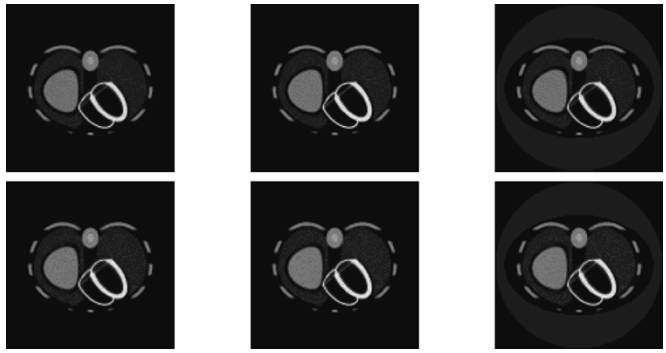

Fig. 9. Ensemble mean (top row) and sample mean (bottom row) images of the MCAT simulation for the RBI-EM (left), OS-EM (center), and RBI-SMART (right) algorithms. Both ensemble and sample mean images were generated for 16 full iterations of each algorithm. The projection data contained 500 000 total counts and were partitioned into 8 subsets.

IV. Discussion

We have attempted to fully validate the mathematical formulations in Part I [1], which allow us to compute the ensemble statistical properties of RBI reconstruction algorithms. In light of the large amount of data generated in the many tables and figures contained herein, we would like to highlight a few noteworthy aspects of the results for the reader.

We have compared the approximate ensemble statistical parameters to their respective sample counterparts, using a uniform disk and MCAT simulation. In doing this, Monte Carlo estimates were used to represent the true ensemble statistical parameters. Typically, error bars quantifying the corresponding confidence interval for a given parameter estimate would be included. In this work, because such large sample sizes were used, the corresponding confidence intervals for the Monte Carlo estimates were extremely small and, thus, if included in the appropriate figure, would be visually imperceptible to the reader. For example, in the mean image plot for RBI-EM (Fig. 4, top), using the expression for the large-sample 95% confidence interval [16] for pixel 15 with sample mean $\bar{x}$, sample standard deviation s and sample size N, we get

$$\bar{x} \pm z_{.025}\frac{s}{\sqrt{N}} \approx 30 \pm 1.96\frac{\sqrt{10}}{\sqrt{10\,000}} = 30 \pm 0.062 \tag{5}$$

which, if plotted, would be too small to be visually perceived by the reader, given the scale of the y-axis. For this reason, error bars were not included for Monte Carlo plots in Figs. 4, 6, 8, 10 and 12.

In emission tomographic imaging, there are two sources of randomness that we must attempt to quantify: 1) randomness due to quantum noise for a fixed object, and 2) randomness due to variability in the object itself. It is important to note that object variability was not included in the analysis in Part I [1], and

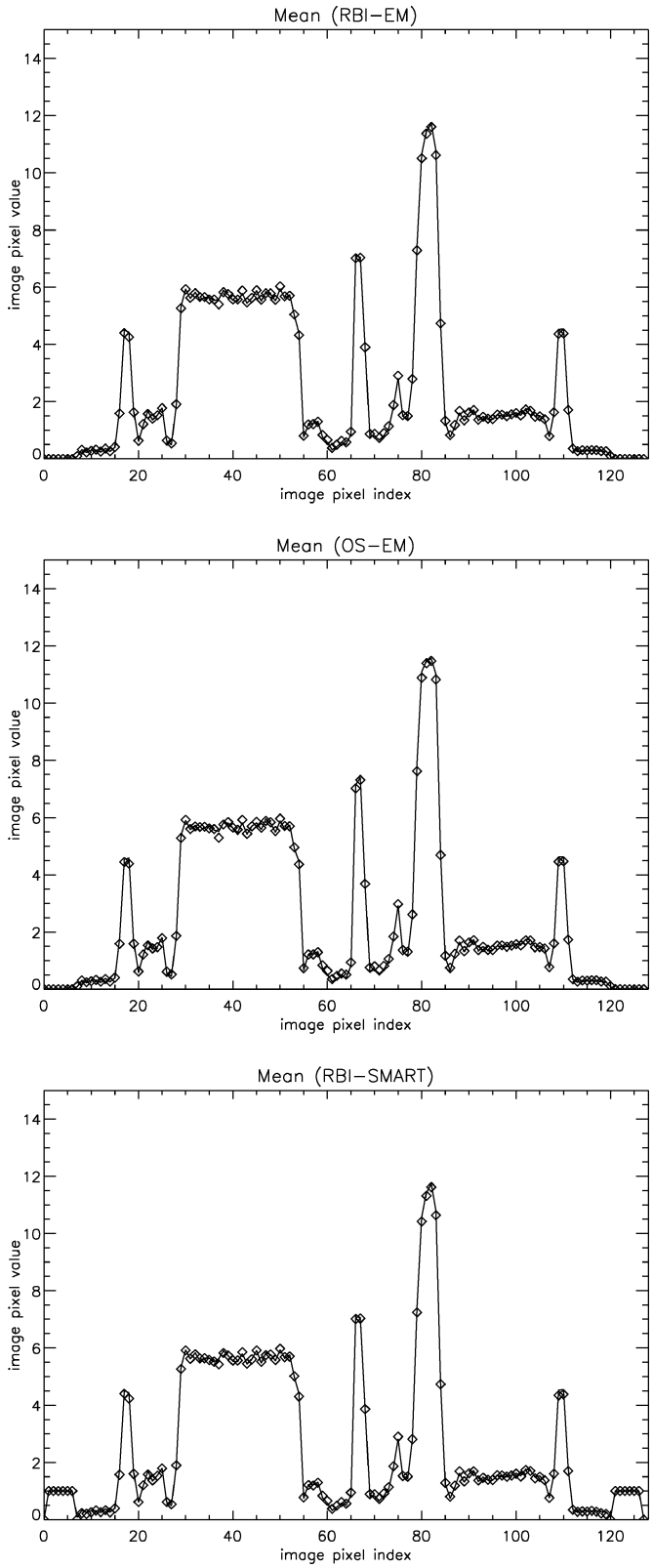

Fig. 10. Central profiles along the x-axis through the ensemble mean (solid lines) and sample mean ($\diamond$) images of the MCAT simulation, for the RBI-EM (top), OS-EM (center), and RBI-SMART (bottom) algorithms, as shown in Fig. 9.

so these validations also do not take object variation into account. However Abbey [17] provides an equivalent derivation

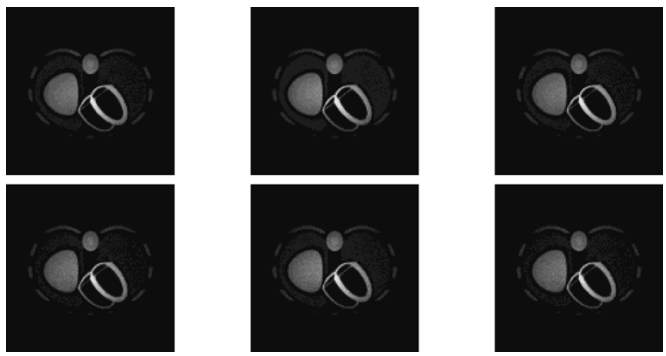

Fig. 11. Ensemble variance (top row) and sample variance (bottom row) images of the MCAT simulation for the RBI-EM (left), OS-EM (center), and RBI-SMART (right) algorithms. Both ensemble and sample variance images were generated for 16 full iterations of each algorithm. The projection data contained 500 000 total counts and were partitioned into 8 subsets.

of the ensemble statistical parameters for nonlinear iterative reconstruction operators, for which randomness due to quantum noise and object variability are both included.

Part I of this paper [1] gives a detailed treatment of a theoretical methodology to derive the approximate ensemble statistical properties of RBI reconstructed images. Their formulation relies on a low-noise approximation, which states that in order for the theory to be valid, the noise in the reconstructed image must be small compared to the mean image. Two obvious ways in which the low-noise approximation can be violated is for the projection data to be low-count or to consider high iteration numbers. In the former case, the relative signal-to-noise ratio is low and so the noise may not be small compared with the mean, even at the outset of the reconstruction process. In the latter case, as iteration number increases, noise variance in the reconstructed image tends to increase and, thus, the noise is not small compared with the mean.

In this work, we have attempted to confirm the predictions of the theory as elucidated in [1] Overall, we have shown the approximate ensemble statistical properties to agree well with their respective large sample estimates for a wide variety of noise levels and subset orderings, for each of the three algorithms. It was also demonstrated that the theoretical formulation breaks down when the low-noise approximation is violated, i.e. with low-count data or when iterating too far. However, there is some evidence that in certain low-count scenarios, the methodology might be useful for approximating the ensemble statistical properties of the three algorithms. The breakdown of the methodology is most apparent for large iteration numbers, which in practice, would not be used.

V. Conclusion

In this paper, we evaluated a methodology that predicts the ensemble statistical properties of the RBI-EM reconstruction algorithm, OS-EM, and RBI-SMART. This evaluation was accomplished using Monte Carlo methods to calculate the sample statistical properties of each algorithm and then comparing the results with the aforementioned theoretical formulations. We also investigated the validity of the methodology,

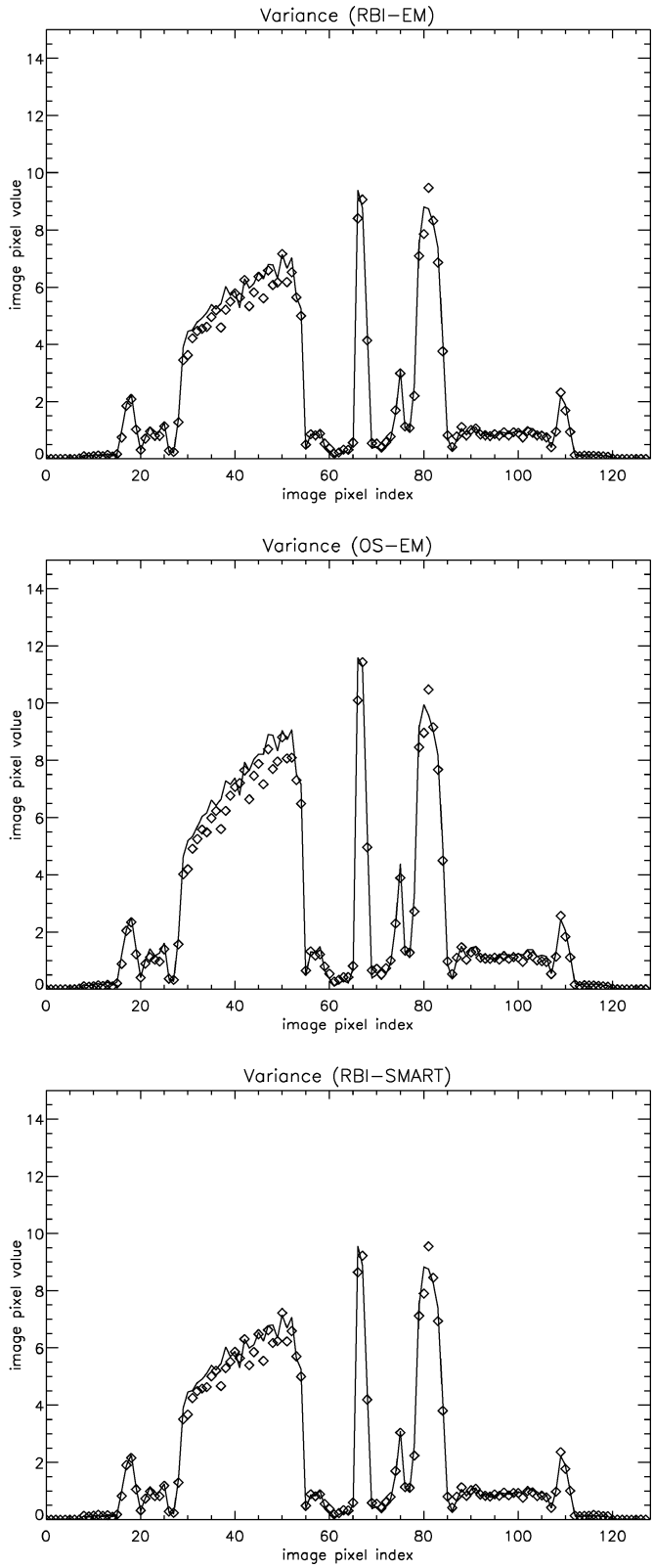

Fig. 12. Central profiles along the x-axis through the ensemble variance (solid lines) and sample variance ($\diamond$) images of the MCAT simulation, for the RBI-EM (top), OS-EM (center), and RBI-SMART (bottom) algorithms, as shown in Fig. 11.

which depends upon a low-noise approximation. We found that the methodology works quite well for a variety of noise levels, subset configurations and iteration points. However, the methodology breaks down in low-count situations and for large numbers of iterations. Thus, the theoretical methodology can be used to derive the ensemble statistical parameters for RBI reconstructed images, which in turn could be used for objective image quality assessment.

ACKNOWLEDGMENT

The authors would like to acknowledge many beneficial discussions on image noise and iterative algorithms with C. L. Byrne of the University of Massachusetts at Lowell, H. H. Barrett and Donald W. Wilson of the University of Arizona, C. K. Abbey of the University of California at Davis, H. Malcolm Hudson and Richard Larkin of Macquarie University, G. Gindi and W. Wang of SUNY Stony Brook, and J. A. Fessler of the University of Michigan.

REFERENCES

[1] E. J. Soares, C. L. Byrne, and S. J. Glick, "Noise characterization of block-iterative reconstruction algorithms: I. Theory," *IEEE Trans. Med. Imag.*, vol. 19, no. 4, pp. 261–270, Apr. 2000.

[2] H. H. Barrett, D. W. Wilson, and B. M. W. Tsui, "Noise properties of the EM algorithm: I. Theory," *Phys. Med. Biol.*, vol. 39, pp. 833–846, 1994.

[3] D. W. Wilson, B. M. W. Tsui, and H. H. Barrett, "Noise properties of the EM algorithm: II Monte Carlo simulations," *Phys. Med. Biol.*, vol. 39, pp. 847–871, 1994.

[4] J. A. Fessler, "Mean and variance of implicitly defined biased estimators (such as penalized maximum likelihood): applications to tomography," *IEEE Trans. Imag. Process.*, vol. 5, no. 3, pp. 493–506, Mar. 1996.

[5] W. Wang and G. Gindi, "Noise Analysis of Regularized EM Algorithms for SPECT," SUNY, Stony Brook, MY, Tech. Rep. MIPL-96-1, 1996.

[6] ——, "Noise Analysis of Regularized EM Algorithms for SPECT: Validation and Task Performance Application to Quantitation," SUNY, Stony Brook, MY, Tech. Rep. MIPL-96-3, 1996.

[7] K. J. Myers, "Visual perception in correlated noise," Ph.D. dissertation, Univ. Arizona (Opt. Sci.), Tucson, AZ, 1985.

[8] K. J. Myers and H. H. Barrett, "Addition of a channel mechanism to the ideal observer," *JOSA A*, vol. 4, pp. 2447–2457, 1987.

[9] H. H. Barrett, "Objective assessment of image quality: effects of quantum noise and object variability," *JOSA A*, vol. 7, no. 7, pp. 1266–1278, July 1990.

[10] H. C. Gifford, M. A. King, M. V. Narayanan, P. H. Pretorius, M. S. Smyczynski, and R. G. Wells, "Effect of block-iterative acceleration on Ga-67 tumor detection in thoracic SPECT," *IEEE Trans. Nucl. Sci.*, vol. 49, no. 1, pp. 50–55, Feb. 2002.

[11] H. H. Barrett and W. Swindell, *Radiological Imaging: The Theory of Image Formation, Detection and Processing*. New York: Academic, 1981.

[12] D. W. Wilson, B. M. W. Tsui, and J. A. Terry, "Non-stationary noise characteristics for SPECT images," in *Conf. Rec. 1991 IEEE Nuclear Science Symp. Medical Imaging Conf.*, G. T. Baldwin, Ed., Piscataway, NJ, 1991, pp. 1736–1740.

[13] X. D. Zhao, B. M. W. Tsui, G. K. Gregoriou, D. S. Lalush, J. Li, and R. L. Eisner, "Evaluation of corrective reconstruction methods using a 3-D cardiac-torso phantom and bull's-eye plots," *IEEE Trans. Nucl. Sci.*, vol. 41, no. 6, pp. 2831–2837, Dec. 1994.

[14] T.-S. Pan, D.-L. Luo, and M. A. King, "Design of an efficient 3-D projector and backprojector pair for SPECT," in *Proc. 1995 Int. Meeting Fully Three-dimensional Image Reconstruction in Radiology and Nuclear Medicine*, Aix-les-Bains, France, July 1995, pp. 181–185.

[15] R. L. Siddon, "Fast calculation of the exact radiological path for a three-dimensional CT array," *Med. Phys.*, vol. 12, no. 2, pp. 252–255, Mar./Apr. 1985.

[16] R. E. Walpole, R. H. Myers, S. L. Myers, and K. Ye, *Probability and Statistics for Engineers and Scientists*, 7th ed. Englewood Cliffs, NJ: Prentice-Hall, 2002.

[17] C. K. Abbey, "Assessment of Reconstructed Images," Ph.D. dissertation, Univ. Arizona (Appl. Math.), Tucson, AZ, 1998.