# SIMULTANEOUS CONFIDENCE INTERVALS FOR IMAGE RECONSTRUCTION PROBLEMS

*Yong Zhang*, Alfred O. Hero III**, and W. L. Rogers****

* Bio-Engineering Program, **Dept. of EECS, and ***Division of Nuclear Medicine
The University of Michigan, Ann Arbor, MI 48109

## ABSTRACT

We provide a methodology for specifying a set of simultaneous $(1-\alpha)\%$ confidence intervals on the intensity of each image pixel for emission and transmission tomography. These intervals give a $(1-\alpha)\%$ confidence region which, given a specific family of noise distributions, e.g. Gaussian or Poisson, is guaranteed to contain the actual image with probability at least $1-\alpha$. This region is a "set estimate" of the image which can be used to study confidence levels of popular image reconstructions such as filtered back projection, weighted-least-squares, and maximum likelihood. Alternatively, the set estimate can be used as a feasibility region from which particular image estimates can be selected based on additional criteria. A simulation for parallel ray projection geometries in emission tomography is given.

## 1. INTRODUCTION

Tomographic reconstruction can be stated in terms of estimating an image intensity $\underline{\lambda} \in IR^p$ from $N$ projection measurements $\underline{Y} = A\underline{\lambda} + \underline{e}$ where $A$ is an $N \times p$ tomographic system matrix and $\underline{e}$ represents errors in the linear model $\overline{\underline{Y}} = A\underline{\lambda}$ due to noise or system mismodeling. A *point estimator* of $\underline{\lambda}$ is a point function $\hat{\underline{\lambda}} = \hat{\underline{\lambda}}(\underline{Y})$. The maximum likelihood via EM (MLEM), weighted-least squares (WLS), algebraic reconstruction technique (ART), and filtered back projection (FBP) are examples of point estimation strategies. While some of these point estimators may derived based on some statistical optimality criterion, for a given realization a point estimator does not provide any information about its statistical confidence or about its consistency with properties of the projection noise distribution. Such properties may be strongly parametric characterizations, e.g. known Poisson or Gaussian noise statistics, or they may be weaker non-parametric characterizations, e.g. upper bounds on the $\alpha$ and $1-\alpha$ quantiles of the noise distribution ($\alpha \in [0,1]$).

In this paper the quantiles are used to develop a confidence region on the true image given the measured data set using methods of bounded error estimation. Using our methodology we can specify a *set estimator* which is guaranteed to contain the true image with probability at least $1-\alpha$. An ellipsoid parallel cuts (EPC) algorithm is presented for constructing confidence regions for image reconstruction which uses the QR decomposition to improve runtime efficiency. Simulations of a simple PET phantom are performed and it is determined that: 1) the centroid of the $(1-\alpha)\%$ confidence region is a point estimator which is competitive with the unregularized MLEM and converges in a finite number of iterations; 2) Llacer's feasible set of images is closely related to our confidence regions.

[1]This research was supported in part by the National Science Foundation under grant BCS-9024370 and the National Cancer Institute under grant R01-CA-54362-02

## 2. BOUNDED ERROR ESTIMATION

Bounded error estimation has a history in systems identification [1], robust estimation and prediction [2], state estimation and filtering [3], and signal processing [4]. Assume a nominal linear measurement model: $\overline{Y}_k = \underline{\phi}^T(k)\underline{\theta}, k = 1, ..., N$, where $\overline{Y}_k$ is the model output, $\underline{\phi}^T(k)$ is a vector specific to the system, e.g. the $k$-th row of the tomographic system matrix $A$, $\underline{\theta} \in IR^p$ is the parameter vector to be estimated, e.g. $\underline{\lambda}$, and $k$ is the measurement index. If it is known that the error $e(k) = Y_k - \overline{Y}(k)$ is bounded within $[e_{min}(k), e_{max}(k)]$:

$$e_{min}(k) \leq Y_k - \phi^T(k)\underline{\theta} \leq e_{max}(k), \qquad k = 1, ..., N, \qquad (1)$$

then the set of all values of $\underline{\theta}$ consistent with (1) is given by the intersection $\Lambda = \cap_{k=1}^N H_k$ of the hyperslabs: $H_k = \{\underline{\theta} : Y_k - e_{max}(k) \leq \underline{\phi}^T(k)\underline{\theta} \leq Y_k - e_{min}(k)\}$, $k = 1, ..., N$. The Ellipsoid Parallel Cuts (EPC) algorithm [3, 1] finds a sequence of successively smaller ellipsoids $\{E_k\}_{k=1}^N$ containing $\Lambda$:

$$E_k = E(\underline{\theta}_k, \Sigma_k) = \left\{ \underline{\theta} : (\underline{\theta} - \underline{\theta}_k)^T \Sigma_k^{-1} (\underline{\theta} - \underline{\theta}_k) \leq 1 \right\} \qquad (2)$$

where $\underline{\theta}_k$ is the centroid and $\Sigma_k$ is a positive definite concentration matrix defining principal and minor axes of the ellipsoid. As long as $N \geq p$, after $N$ steps the EPC algorithm yields the minimal volume ellipsoid containing $\Lambda$. (Fig. 1)

**INITIALIZATION**: $\Sigma_0 = r^2 \times I_{p \times p}$; $\underline{\theta}_0 = \underline{0}$;
**FOR** $k = 1, \ldots N$:

$$\alpha_k^+ = \frac{Y_m(k) - \underline{\phi}^T(k)\underline{\theta}_{k-1} - e_{max}(k)}{\sqrt{\underline{\phi}^T(k)\Sigma_{k-1}\underline{\phi}(k)}}$$

$$\alpha_k^- = \frac{\underline{\phi}^T(k)\underline{\theta}_{k-1} - Y_m(k) - e_{min}(k)}{\sqrt{\underline{\phi}^T(k)\Sigma_{k-1}\underline{\phi}(k)}}$$

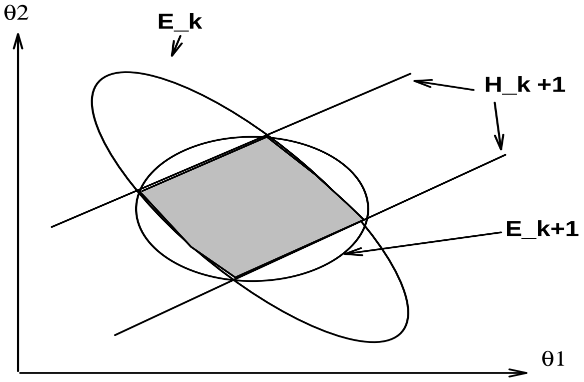


Figure 1. EPC finds minimal volume ellipsoid containing $\cap_{k=1}^{q} H_k$, $q = 1, \ldots, N$.

**IF** $\alpha_k^+ > 1$ or $\alpha_k^- > 1$, null intersection.
  **THEN** ignore $Y_k$, or adjust $e_m(k)$ and restart EPC.
**ELSE** $\alpha_k^+ = \max(\alpha_k^+, 1)$ and $\alpha_k^- = \max(\alpha_k^-, 1)$
  **IF** $\alpha_k^+ \alpha_k^- \geq 1/p$
  The intersection is the same as the current ellipsoid.
    **THEN** $E_k = E_{k-1}$ and $\underline{\theta}_k = \underline{\theta}_{k-1}$
  **ELSE**
    **IF** $\alpha_k^+ = \alpha_k^-$
      **THEN** Apply centrally symmetric parallel-cut as:

$$\underline{\theta}_k = \underline{\theta}_{k-1}$$
$$\Sigma_k = \beta_k \left( \Sigma_{k-1} - \frac{\zeta_k}{\underline{\phi}^T(k)\Sigma_{k-1}\underline{\phi}(k)} \Sigma_{k-1}\underline{\phi}(k)\underline{\phi}^T(k)\Sigma_{k-1} \right)$$

where

$$\beta_k = \frac{p(1-\alpha_k^2)}{p-1}$$
$$\zeta_k = 1 - p\alpha_k^2(1-\alpha_k^2)$$

**ELSE** Apply normal parallel cuts as:

$$\underline{\theta}_k = \underline{\theta}_{k-1} + \frac{\sigma_k(\alpha_k^+ - \alpha_k^-)}{2\sqrt{\underline{\phi}^T(k)\Sigma_{k-1}\underline{\phi}(k)}} \Sigma_{k-1}\underline{\phi}(k)$$
$$\Sigma_k = \delta_k(\Sigma_{k-1} - \frac{\sigma_k}{\underline{\phi}^T(k)\Sigma_{k-1}\underline{\phi}(k)} \Sigma_{k-1}\underline{\phi}(k)\underline{\phi}^T(k)\Sigma_{k-1})$$

where

$$\delta_k = \frac{p^2}{p^2-1}(1 - \frac{(\alpha_k^+)^2 + (\alpha_k^-)^2 - \rho_k/p}{2})$$
$$\sigma_k = \frac{1}{p+1}(p + \frac{2}{(\alpha_k^+ - \alpha_k^-)^2}(1 - \alpha_k^+\alpha_k^- - \rho_k/2)),$$
$$\rho_k = \sqrt{4(1-(\alpha_k^-)^2)(1-(\alpha_k^-)^2) + p^2((\alpha_k^+)^2 - (\alpha_k^-)^2)^2)}.$$

Taking advantage of symmetries in the EPC calculations, the EPC algorithm requires on the order of $N(4p^2+28p+21)$ flops. The principal bottleneck in EPC is the sequence of $N$ matrix-vector multiplies of the form $\Sigma_k\underline{\phi}(k)$, $k = 1, \ldots, N$. By using the QR decomposition we obtain the equivalent measurement equation: $Q^T\underline{Y} = R\underline{\lambda} + Q^T\underline{e}$. This effectively rotates components of the noise $\underline{e}$ lying outside the range space of $A$ into the lower $N-p$ elements of $Q^T\underline{e}$, hence reducing the number of non-zero rows $\underline{\phi}_k^T$ from $N$ to $p$. Discarding the last $N-p$ zero rows of the above measurement equation, we obtain a $p \times p$ system of measurement equations so that the QR-EPC algorithm converges in only $p$ steps instead of $N$. We conclude that the total online cost of QR-EPC becomes $(4p^3 + 25p^2 + 21p + 8Np)$ flops. Since the matrix $Q$ is needed, there is also an increase in the memory storage requirements for QR-EPC as opposed to regular EPC. Storage of the matrix $Q$ requires $O(Np)$ Bytes if Householder reflections are used, whereas it requires only $O(\rho Np)$ bytes if Given's rotations are used, where $\rho \in [0,1]$ is the sparsity factor of $A$. For comparison the general (no exploitation of sparse $A$) MLEM algorithm needs $N(2p^2+2p)$ flops per iteration and may in some cases require several thousand iterations to converge. In particular, if $N$ is twice as large $p$, the time to compute all $p$ iterations of QR-EPC will be comparable to the time to compute a single iteration of general MLEM.

## 3. CONFIDENCE REGIONS

When $\underline{e}_{max}$ and $\underline{e}_{min}$ are selected to correspond to specific quantiles of the projection noise distribution the consistency set $\Lambda$ can often be manipulated to yield a $(1-\alpha)\%$ confidence region for $\underline{\theta}$. For ECT, the projection data $Y_1, \ldots, Y_N$, are distributed as independent Poisson random variables with rates $E[\underline{Y}] = A\underline{\lambda}$. It can be shown [5] that for a Poisson variable $Y_k$, a $(1-\beta)\%$ confidence interval for the rate $E[Y_k]$ is:

$$[L_{min}(k), L_{max}(k)] = \left[\frac{1}{2}\chi^2_{\beta/2}(2Y_k), \frac{1}{2}\chi^2_{1-\beta/2}(2[Y_k+1])\right] \quad (3)$$

where $\chi^2_\xi(\nu)$ is the $\xi$th percentile of the chi-square distribution with $\nu$ degrees of freedom. Using this fact a $(1-\alpha)\%$ confidence rectangle for $A\underline{\lambda}$ is obtained as $[\underline{L}_{min}, \underline{L}_{max}] \stackrel{def}{=} \bigtimes_{k=1}^{N}[L_{min}(k), L_{max}(k)]$, where in $L_{min}(k)$ and $L_{max}(k)$ (3) we have set $\beta = 1-(1-\alpha)^{\frac{1}{N}}$. From this we obtain a $(1-\alpha)\%$ confidence region for $\underline{\lambda}$:

$$\begin{aligned} \Lambda_{1-\alpha} &= \{\underline{\lambda} : \underline{L}_{min} \leq A\underline{\lambda} \leq \underline{L}_{max}\} \\ &= \{\underline{\lambda} : \underline{e}_{min} \leq \underline{Y} - A\underline{\lambda} \leq \underline{e}_{max}\} \end{aligned} \quad (4)$$

where $\underline{e}_{min} \stackrel{def}{=} \underline{Y} - \underline{L}_{max}$, $\underline{e}_{max} \stackrel{def}{=} \underline{Y} - \underline{L}_{min}$. The $(1-\alpha)\%$ confidence region (4) is in the form of a bounded error (1) to which the QR-EPC algorithm can be directly applied to find the minimum ellipsoid $\Lambda_{1-\alpha}$ containing $\Lambda$.

If a point estimator $\hat{\underline{\theta}}$ is outside of the ellipsoid region $\Lambda_{1-\alpha}$ then the Euclidean distance between $\hat{\underline{\theta}}$ and the set $E_p$ is simply:

$$\delta(\hat{\underline{\theta}}) = (\hat{\underline{\theta}} - \underline{\theta}_p)^T \Sigma_p^{-1} (\hat{\underline{\theta}} - \underline{\theta}_p) - 1. \quad (5)$$

The distance $\delta(\hat{\underline{\theta}}) + 1$ is called the *EPC distance*, which is a weighted distance between $\hat{\underline{\theta}}$ and the centroid $\underline{\theta}_p$ of $E_p$ where the weight matrix is equal to $\Sigma_p^{-1}$. When the EPC algorithm is implemented on the noisy projections data, this distance measure can be used as a measure of closeness of an image reconstruction $\hat{\underline{\theta}}$ to a $(1-\alpha)\%$ confidence region for $\underline{\theta}$.

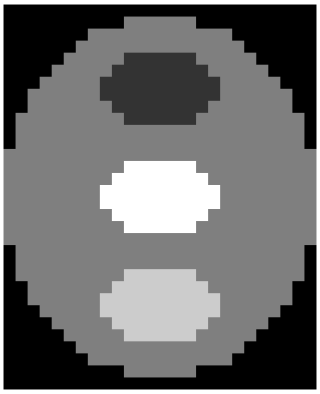

Figure 2. Phantom image used for simulations

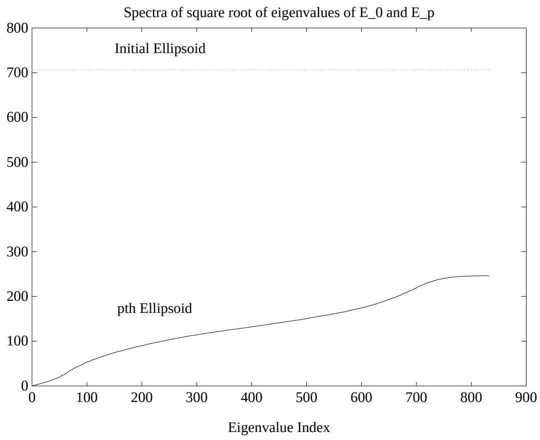


Figure 3. The square root of the eigenvalues of initial ellipsoid and final ellipsoid.

## 4. NUMERICAL RESULTS

A simple 4 level phantom (Fig. 2) with intensity $\underline{\lambda}$ sampled over $26 \times 32$ pixels was projected onto 42 detector bins at 90 equally distributed angles over the range of 180 degrees using strip integrals to form 3780 mean projections $A\underline{\lambda}$. Using these mean projections as Poisson rates, the projection data was generated as 3780 independent Poisson random variables. The total number of counts collected was approximately 1.96 million.

The EPC algorithm was initialized with a spheroid $E_0$ with radius $r_0 = 7.07 \times 10^2$. The QR-EPC algorithm was implemented with two values 0.9 and 0.95 of $1-\alpha$ and iterated until convergence occurred at $p = 832$ iterations. The square root of the eigenvalues of the final ellipsoid matrix $\Sigma_p$ for $1-\alpha = 0.95$, which represents the half axes of the ellipsoid, is shown in Fig. 3. We see that the size of the final ellipsoid is significantly reduced as compared to the initial spheroid axes.

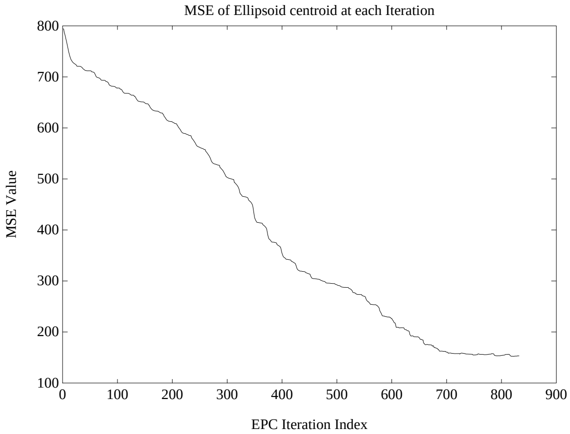


Figure 4. Mean square error as a function of EPC iteration index. Total number of counts is 1.94M. Error bounds used in EPC correspond to $95\%$ confidence region.

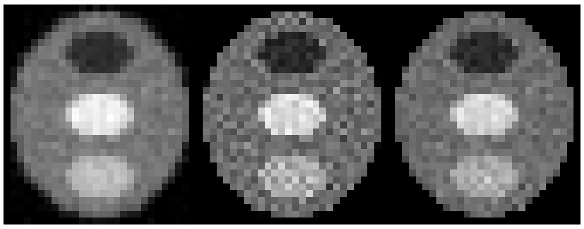

Figure 5. Left: FBP, Middle: MLEM at 500th iteration, Right: centroid of the final ellipsoid. Error bounds in EPC correspond to $95\%$ confidence region.

The sequence of centroids $\underline{\theta}_k = \underline{\theta}_k(\underline{Y})$, $k = 1, \ldots, p$, generated by the EPC algorithm of Section II, is a sequence of image reconstructions corresponding to the geometric centers of the ellipsoids. The mean squared error $\|\underline{\theta}_k - \underline{\theta}\|^2$ between this sequence and the true image is plotted in Fig. 4 and decreases monotonically as a function of EPC iteration number $k$. The non-iterative filtered backprojection (FBP) and the iterative maximum likelihood EM (MLEM) image reconstruction were implemented and compared to the final EPC centroid image reconstruction $\underline{\theta}_p$ Fig. 5. The MLEM reconstruction $\hat{\underline{\theta}}_k^{ML}$ is known to become quite noisy after a large number of iterations $k$ so in the figure it was stopped at the iteration (500) which provided the minimum MSE - an evidently unrealizable stopping rule. Note that the EPC centroid appears to reconstruct the high frequencies better than the FPB algorithm without as much granular noise degradation as the MLEM reconstruction.

Llacer and Veklerov [6] introduced the concept of feasible sets of images to obtain a statistically based stopping rule for the MLEM algorithm. For each iteration of MLEM they computed a statistic, called the $H$-statistic in [6], which measured the closeness of the MLEM iterate to a Poisson distributed image reconstruction. In [6] it was proposed to stop iterating the MLEM algorithm at an iteration for which the $H(\hat{\underline{\theta}}_k^{ML})$ falls below a threshold, a set of iterates called a "set of feasible images", where this threshold was specified to give a false alarm probability $\alpha$. Various stopping rules have been proposed, such as the first threshold downcrossing iteration, the minimum $H$ iteration, or the first threshold upcrossing iteration. Here we compare this with a stopping rule based on the distance $\delta(\hat{\underline{\theta}}_k^{ML})$ between each MLEM iterate and the boundary of the $(1-\alpha)\%$ confidence region. Figure 6 indicates that MLEM enters the 95% confidence region (represented by the "1" line) at the 34th iteration remaining within the confidence region until the 113th iteration, the minimum value of $\delta(\hat{\underline{\theta}}_k^{ML})$ being attained at the 54th iteration. Comparing this result with Llacer's technique applied to MLEM, the $H(\hat{\underline{\theta}}_k^{ML})$ curve Fig. 7 indicates that MLEM enters the 95% feasible set after the 38th iteration, remaining within this feasible set until the 106th iteration, the minimum value of $H(\hat{\underline{\theta}}_k^{ML})$ being attained at the 53rd iteration. The results for 90% confidence levels are analogous. This remarkable similarity suggests that Llacer's estimator-dependent definition of feasible set is closely related to our estimator-independent definition of $(1-\alpha)\%$ confidence region. Figure 8 displays the MLEM reconstructions at the iterations where it enters and leaves the 95% confidence region, and at the iteration where the distance $\delta(\hat{\underline{\theta}}_k^{ML})$ reaches its minimum value.

## REFERENCES


[1] E. Fogel and Y. Huang, "On the value of information in system identification - bounded noise case," *Automatica*, vol. 18, pp. 229–238, 1982.

[2] M. Milanese and R. Tempo, "Optimal algorithms theory for robust estimation and prediction," *IEEE Trans. Automatic Control*, vol. AC-30, pp. 730–738, 1985.

[3] F. C. Schweppe, "Recursive state estimation: unknown but bounded errors and system inputs," *IEEE Trans. Automatic Control*, vol. AC-13, pp. 22–28, 1968.

[4] P. L. Combettes and H. J. Trussell, "The use of noise properties in set theoretic estimation," *Signal Processing*, vol. SP-39, pp. 1630–1641, 1991.

[5] E. B. Manoukian, *Modern Concepts and Theorems of Mathematical Statistics*, Springer-Verlag, New York N.Y., 1986.

[6] J. Llacer and E. Veklerov, "Feasible images and practical stopping rules for iterative algorithms in emission tomography," *IEEE Trans. on Medical Imaging*, vol. 8, no. 2, pp. 186–193, 1989.


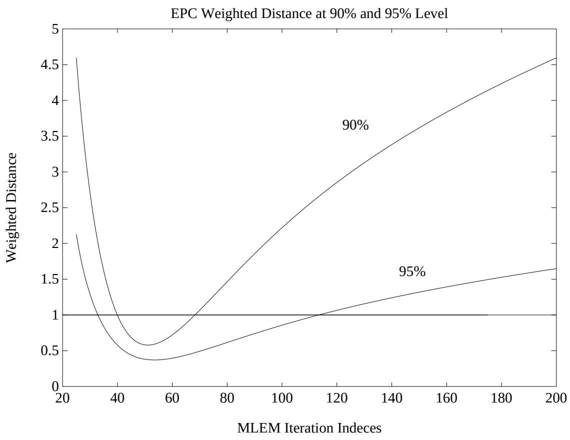


Figure 6. EPC weighted distance shows when the MLEM estimate enters and leaves the 90% or 95% confidence region.

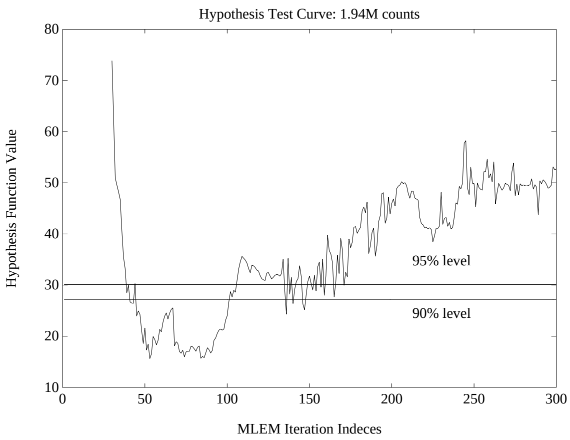


Figure 7. Llacer's $H$ curve for MLEM with $1-\alpha = 90\%$ and 95%.

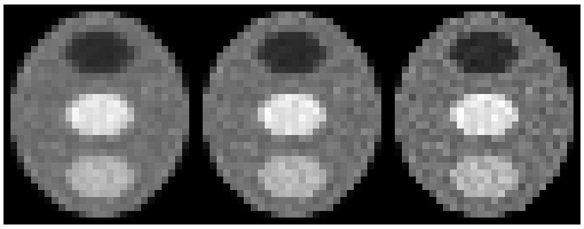

Figure 8. MLEM stopped at first 95% level downcrossing of $\delta$, minimum $\delta$, first 95% level upcrossing of $\delta$.