Title
On the Validity of Hypothesis Testing for Feasibility of Image Reconstructions

Author
Llacer, J.

Publication Date
1990

LBL-26632
Preprint

Lawrence Berkeley Laboratory
UNIVERSITY OF CALIFORNIA

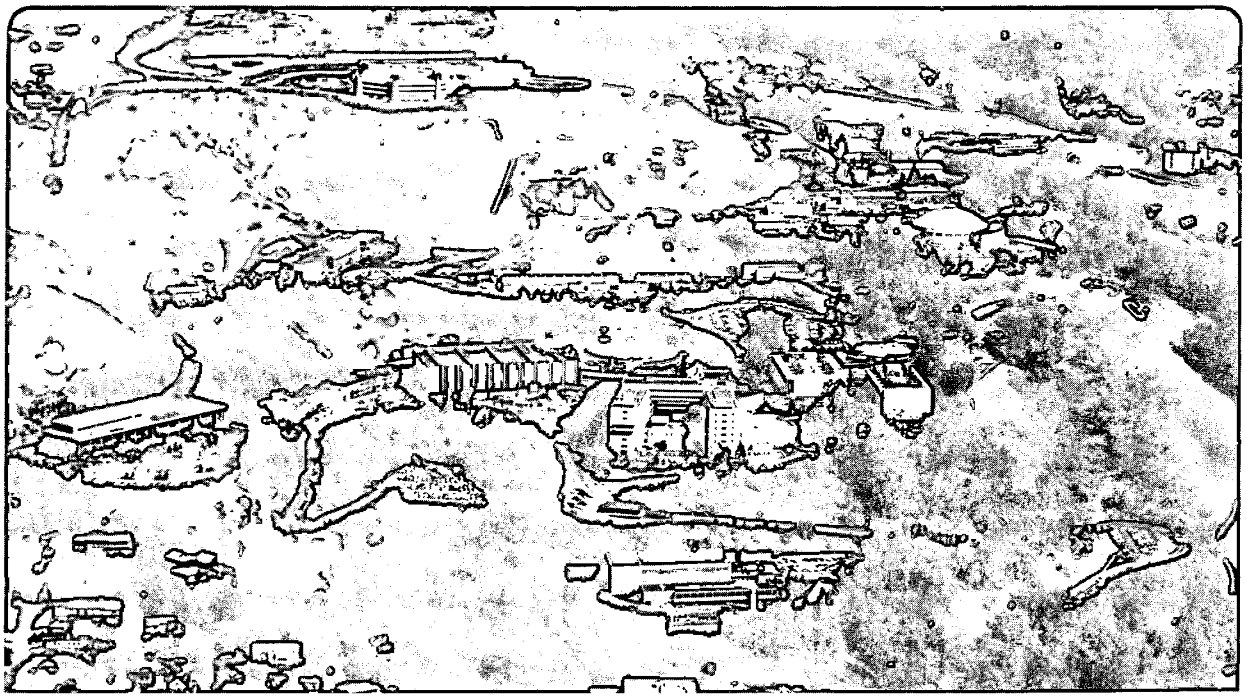
Engineering Division

Submitted to IEEE Transactions on Medical Imaging

On the Validity of Hypothesis Testing for Feasibility of Image Reconstructions

J. Llacer

January 1990



Prepared for the U.S. Department of Energy under Contract Number DE-AC03-76SF00098.

LBL-26632

On the Validity of Hypothesis Testing for Feasibility of Image Reconstructions.

Jorge Llacer, Senior Member, IEEE
Engineering Division
Lawrence Berkeley Laboratory
1 Cyclotron Road, Berkeley, CA, 94720

This work has been supported, in part, by a grant from the National Institutes of Health (CA-39501) and by the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

On the Validity of Hypothesis Testing for Feasibility of Image Reconstructions.

Jorge Llacer, Senior Member, IEEE
Lawrence Berkeley Laboratory, 1 Cyclotron Road, Berkeley, CA, 94720

Abstract - Feasible images have been defined as those images that could have generated the original data by the statistical process that governs the measurement. In the case of emission tomography, the statistical process of emission is Poisson and it has been shown that feasible images resulting from the maximum likelihood estimator (MLE) and Bayesian methods with entropy "priors" can be of high quality. Tests for feasibility have been described that are based on one critical assumption: the image that is being tested is independent of the data, although the reconstruction algorithm has used those data in order to obtain the image. This fact could render the procedure unacceptable unless it is shown that its effects on the results of the tests are negligible. This paper presents experimental evidence showing that images reconstructed by the MLE and stopped before convergence do indeed behave as if independent of the data. The results justify the use of hypothesis testing in practice, although they leave the problem of analytical proof still open.

1. INTRODUCTION

With increasing interest in image reconstruction by the Maximum Likelihood Estimator (MLE) method first proposed for Positron Emission Tomography (PET) by Shepp and Vardi [1], there has been an awareness of the fact that reconstructions by that method converge to a poor representation of the radioisotope distribution that generated the projection data. There is a range of iterations during the reconstruction, however, in which the images are visually good, with an apparently reasonable compromise between sharpness and noise in the regions of high activity.

The present author and co-worker (E. Veklerov) became aware of the fact that those effects are a necessary consequence of posing the reconstruction problem as one of maximizing the probability that the reconstructed image would have given the noisy data, as discussed in [2]. As a result of that observation, Veklerov and Llacer published a stopping rule in [3] based on statistical hypothesis testing that, in effect, acts as a constraint on the class of images that are acceptable as the outcome of a reconstruction. Those images have been called "feasible" in [4] and they are images that, if they were true radioisotope distributions, could have generated the projection data by the statistical process that governs the measurement. The implications of selecting reconstructions from the set of feasible images and several ways of obtaining such images have been examined in detail in [4 - 8]. Although there is little doubt that some feasible images, selected by stopping rules or otherwise, appear to be excellent reconstructions of data with Poisson noise, there is a problem in attempting to determine feasibility in terms of statistical hypothesis testing, as we have done in [3] and [4], Hebert et al in [9], and numerous authors in the field of astronomy in [10 - 15]. In the field of communication theory, feasibility concepts were described in [16,17].

The problem stems from the fact that the image which is being tested for feasibility has been obtained from the data against which it is being tested, i.e., the image and the data are not statistically independent. Under those conditions, hypothesis testing has some limitations which have to be taken in consideration. This paper will discuss the theoretical difficulty outlined above, describe a series of tests carried out with one hundred independent data sets and show that, in the region of feasibility, the hypothesis testing function H of Veklerov and Llacer [3] appears χ^2 distributed with the same number of degrees of freedom as in the case of image and data sets that are statistically independent. It is felt that the results justify the practical use of feasibility tests by hypothesis testing, although the question of analytical justification remains open.

2. BACKGROUND MATERIAL

The notation introduced by Shepp and Vardi in [1] will be used throughout this paper:

$n^*(d), \quad (d = 1,\ldots,D)$	- the projection data or the number of coincidences detected in tube d;
$\lambda(b), \quad (b = 1,\ldots,B)$	- the emission density;
$p(b,d)$	- the transition matrix;
$\lambda^*(d) = \sum_{b=1}^{B} \lambda(b)p(b,d)$	- the forward projections,

where B and D are the number of pixels and the number of projections, respectively. We shall consider specifically the case of emission tomography (PET or SPECT) in which disintegration data follow Poisson statistics.

Llacer and Veklerov have given the definition of a feasible image in [4], as follows:

Definition: The image $\lambda(1), \lambda(2),...,\lambda(B)$ is said to be feasible with respect to data $n^*(1), n^*(2),...,n^*(D)$, if and only if the statistical hypothesis that $n^*(1), n^*(2),...n^*(D)$ are a Poisson sample with means $\lambda^*(1), \lambda^*(2),...,\lambda^*(D)$, respectively, can be accepted (not rejected) at a given significance level.

An implementation of a test for feasibility has been described in [3]. It assumes independence between $\lambda^*(d)$ and $n^*(d)$. With this simplifying assumption, it is possible to define a simple hypothesis test with a well defined number of degrees of freedom. The first step of the test consists in scaling the differences $\lambda^*(d) - n^*(d)$ for each data pair to a new variable $x(d)$ which is uniformly distributed between 0 and 1 if the data are Poisson with respect to their means. Next, a histogram with N bins is generated with the values $x(d)$ for all the data pairs. Then we test the hypothesis that $x(d)$ is uniformly distributed between 0 and 1 by Pearson's procedure with N-1 degrees of freedom. The histogram testing function H is defined as:

$$H = \sum_{i=1}^{N} \frac{\left(h_i - \frac{D}{N}\right)^2}{\frac{D}{N}} \qquad (1)$$

where h_i is the observed frequency of $x(d)$ falling in bin i and $\frac{D}{N}$ is the expected frequency if $x(d)$ is uniformly distributed. The above test has been found to be too stringent when dealing with real tomographic data. Under those conditions, we have had to devise a less restricitive test, described in [4]. It will be shown below that the assumption of independence implicit in the test of [3] can be justified in practice by the negligible effects caused by the dependence of the feasible image on the data.

It is felt that it is important to show that those dependence effects are negligible. Cramer has shown in [18] that the χ^2 test, when certain parameters that describe the function against which we are testing are estimated from the sample, can only be used to test the hypothesis that the data are drawn from the defined probability distribution with some unspecified values of those parameters. If the method of estimating the parameters is an asymptotically normal and asymptotically efficient estimator, the χ^2 test can still be carried out as if the distribution were completely specified, but with a number of degrees of freedom reduced by the number of parameters estimated. In our case, when the estimating iterative procedure is stopped before convergence, the parameters $\lambda^*(d)$ have been derived from the data $n^*(d)$ by an estimator of unknown characteristics and it would appear that we can only test the hypothesis that the data were obtained from a Poisson distribution with some values of $\lambda^*(d)$ unknown to us. Unless the data pairs ($\lambda^*(d)$, $n^*(d)$) are found to behave as if they were statistically independent, the test of feasibility by hypothesis testing could be meaningless.

In this context it may be important to remark that the Poisson nature of the positron emission or detection data is often destroyed by multiplying the measured data by correction factors and/or subtracting random background information before initiating a reconstruction by the MLE. Tests for feasibility of images thus reconstructed will always fail. However, if the corrections are applied to the matrix elements $p(b,d)$ and the measured background is not subtracted, but it is used as one more column of that matrix in order to estimate the background in the image data, the Poisson nature of the detected data is not destroyed and the test for their Poisson characteristics with respect to a reconstructed image can be meaningful.

3. DESCRIPTION OF EXPERIMENTS.

The validity of the "independence assumption" for reconstructed images has been demonstrated by generating one hundred independent data sets from the same source distribution, testing the data sets for their Poisson nature, obtaining reconstructions by the MLE method, examining the distribution of the hypothesis testing function H at certain stages of the iterative process and showing that the distributions of H values in the region of iteration numbers which we have called feasible can, indeed, be fitted to a χ^2 distribution with a high degree of confidence.

3.1 Preparation of data

Figure 1 shows the "parent" distribution with 200 million counts, used to generate the one hundred independent realizations of the data set. The geometry of the UCLA ECAT-III tomograph [19] was used in generating the matrix $p(b,d)$ and the simulated data. An image plane of 64 x 64 pixels was used for all the computations. Each independent data set contained 500,000 counts. The Poisson character of the independent data sets was verified successfully by standard methods and by the feasibility test of [3]. The results of the latter test indicated that the independent data sets were, indeed, feasible with respect to the "parent" distribtuion scaled down to 500,000 counts.

3.2 Analysis of reconstruction results.

Once the quality of the data sets was verified, the following operations were carried out:

1. - Each data set was reconstructed by the MLE method of [1], with images and projections of those images being stored at iterations 15,21,27,36,60 and 150.

2. - The feasibility test of [3] was applied to each projection vector $\lambda^*(d)$ with respect to the corresponding data $n^*(d)$ and the value of the H function of (1) for $N = 10$ bins was saved.

3. - For each of the iteration numbers, the one hundred values of the H function were placed in a 50 bin histogram and each histogram was subsequently normalized to unity area.

4. - Each of the histograms was compared to a χ^2 distribution with N-1 degrees of freedom by carrying out Pearson's test for goodness of fit.

Figures 2a through 2f show the reconstructions at the above indicated iteration numbers for one of the data sets and Figs. 3a through 3f show the corresponding normalized histograms of the H function for the one hundred data sets, compared to a χ^2 distribution with 9 degrees of freedom.

The values of χ^2 printed inside the Figs. 3 correspond to testing the following null hypothesis: *The experimental data points for the H function values follow a χ^2 distribution.* For approximately 25 to 30 degrees of freedom (the number of bins in the histograms containing values substantially different from zero in the region of feasibility), the critical values for rejection of the null hypothesis are ≈ 47 for 99% confidence level, ≈ 37 for 90% and ≈ 18 for 10% confidence level. The values of χ^2 for the cases of iterations 15, 60 and 150 indicate that we can reject the null hypothesis with assurance, but for iterations 21, 37 and 36 the hypothesis cannot be rejected even at confidence levels near 10%. Although we have no assurance that the histograms of H at the region of intermediate iterations are χ^2 distributed, the above data provide ample indication that they may be so.

4. DISCUSSION AND CONCLUSIONS

Leahy et at. in [9] describe the use of a χ^2 feasibility test that checks that the second moment of the distribution of the $(\lambda^*(d), n^*(d))$ data pairs corresponds to that of a Poisson distribution. The same test has also been discussed in [4] and [7, 8]. In all cases it has been found that, in the region of acceptable images reconstructed by the MLE, the number of degrees of freedom has been D - 1, rather than (D - B -1). Leahy concludes that, during the iterative process, the number of degrees of freedom must range between those two numbers, but it is still close to D - 1 in the region of acceptability. The more demanding feasibility test described in [3], by binning and rescaling the residuals ($\lambda^*(d) - n^*(d)$) into a relatively small number N of bins may, in effect, change the nature of the "goodness of fit" problem sufficiently that the actual number of degrees of freedom of the original problem is lost. Similarly, the fact that the parameters were estimated from the data becomes irrelevant. Under those conditions, the parameter *H* of the test of [3] could very well behave as if the data and the projection of the reconstruction were statistically independent and χ^2 distributed with N - 1 degrees of freedom, as we have found in our experiments. It is hoped that these results will stimulate further work and that an analytical understanding will follow in the future.

Like most statistical concepts, the feasibility of an image is not a tight concept. There is a region of feasibility and a worker can pick one of the images in that region, or near it, and know that there is a high probability that, if that image were a true radiation field, it could have given the original data by a Poisson process. In that context, it is interesting to note the differences in the images at the entrance of the feasibility region (in the vicinity of iteration 21) and those toward the exit (in the vicinity of iteration 36). In the first case, the image has not developed full contrast, while in the second case it has. Towards the exit of the region the images also begin to develop what could be considered excessive noise in the areas of high activity. The limitations of the printing process may not quite show those effects in the images of Fig. 2. As the author and co-workers have indicated in [6 - 8] and [20], there are many other ways of obtaining feasible images, not all of which are acceptable as representations of the original source. Feasibility is certainly not a sufficient condition for accepting a reconstruction, but it is logical to make it a necessary condition which can assist in defining parameters for a reconstruction which are otherwise arbitrary

5. ACKNOWLEDGMENTS.

The author would like to acknowledge stimulating discussions with E. Veklerov, J. Nunez and Deborah Nolan, and thank Harrison Barrett for a thorough analysis and thought-provoking criticism of some of the results of our group. This work has been supported by a grant from the National Cancer Institute (CA-39501) and the U.S. Department of Energy under Contract NO. DE-AC03-76SF00098.

REFERENCES

[1] L.A. Shepp and Y. Vardi, "Maximum likelihood reconstruction for emission tomography", IEEE trans. Med. Imaging, vol MI-1, pp. 113-122, 1982.

[2] J. Llacer and E. Veklerov, "The Maximum Likelihood Estimator Method of Image Reconstruction: Its Fundamental Characteristics and Their Origin", Proc. of the Xth Information Processing in Medical Imaging, (IPMI), International Conference, Utrecht, The Netherlands, pp 201-215, Plenum Pub. Co., (1987).

[3] E. Veklerov and J. Llacer, "Stopping Rule for the MLE Algorithm Based on Statistical Hypothesis Testing", IEEE Trans. Med. Imaging, vol MI-6, No. 4, pp. 313-319, 1987.

[4] J. Llacer and E. Veklerov, "Feasible Images and Practical Stopping Rules for Iterative Algorithms in Emission Tomography", IEEE Trans. Med. Imaging, vol MI-8, No,. 2, pp. 186-193, 1989.

[5] J. Llacer and E. Veklerov, "The Use of a Stopping Rule in Iterative Image Reconstruction", to be published in the Proceedings of the 1988 AMS-IMS-SIAM Joint Conference on Spatial Statistics and Imaging, Brunswick, Maine, June 1988, Report LBL-25314.

[6] J. Llacer, E. Veklerov and J. Nunez, "The Concept of Causality in Image Reconstruction", to be published in the Proceedings of the NATO Advanced Study Institute on Formation, Handling and Evaluation of Medical Images, Sept. 1988, Report LBL-25877.

[7] J. Nunez and J. Llacer, "Maximum Entropy and the Concept of Feasibility in Tomographic Image Reconstruction", Proc. of SPIE Medical Imaging III Conference, Newport Beach, California, Vol. 1090, pp. 359-372, 1989.

[8] J. Nunez and J. Llacer, "A Fast Bayesian Reconstruction Algorithm for Emission Tomography with Entropy Prior Converging to Feasible Images", accepted for publication in IEEE Trans. Med. Imaging, June 1990.

[9] T. Hebert, R. Leahy and M. Singh, "Fast MLE for SPECT using an intermediate polar representation and stopping criterion", IEEE Trans. Nucl Sci., NS-35, No. 1, pp. 615-619, 1988.

[10] S.F. Gull and G.J. Daniel, "Image Reconstruction from Incomplete and Noisy Data", Nature, vol. 272, pp 686-690, 1978.

[11] J. G. Ables, "Maximum Entropy spectral analysis", Astron. Astrophys. Suppl. Ser., vol. 15, pp 383-393, 1974.

[12] J. Skilling and R.K. Bryan, "Maximum Entropy Reconstruction: general algorithm", Monthly Notices of the Royal Astronomical Soc., vol. 211, pp 111-124, 1984.

[13] R. Narayan and R. Nityananda, "Maximum Entropy Image Restoration in Astronomy", Ann. Rev. Astron. Astrophys., vol. 24, pp 127-170, 1986.

[14] J. Reiterer and J. Pfleiderer, "Improvement on MEM-deconvolution by an additional constraint", Astron. Astrophys., vol. 166, pp 381-392, 1986.

[15] A. Tarantola, Inverse Problem Theory. Elsevier, 1987.

[16] H.J. Trusell, "Convergence criteria for iterative restoration methods", IEEE Trans. ASSP, Feg. 1983.

[17] H.J. Trusell and N.R. Civanlar, "The feasible solution in signal restoration", IEEE Trans. ASSP, April 1984.

[18] H. Cramer, Mathematical Methods of Statistics, Princeton University Press, 1946.

[19] E.J. Hoffman, A.R. Ricci, L.M.A.M. Van der Stee and M. Phelps, "ECAT-III, Basic Design Considerations", IEEE Trans. Nucl. Sci., NS-30, No. 1, pp. 729-733, 1983.

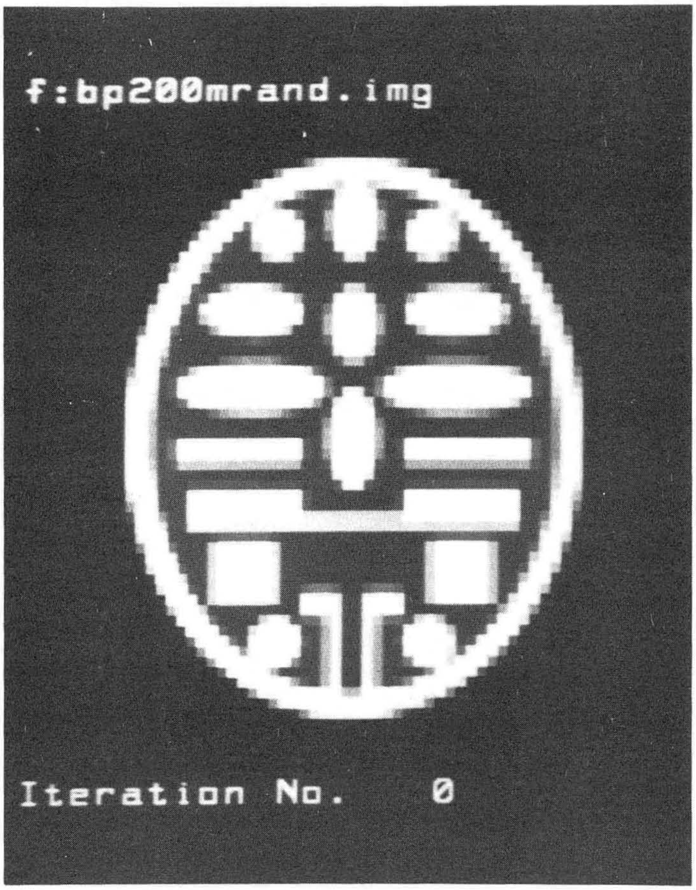
[20] E. Veklerov and J. Llacer, "The feasibility of Images Reconstructed with the Method of Sieves", to be presented in IEEE Nucl. Sci. Symp., San Francisco, January 1990, Report LBL-27088.

FIGURE CAPTIONS.

Fig. 1. Mathematical brain phantom used for the simulations of this paper. The image contains 200 million counts.

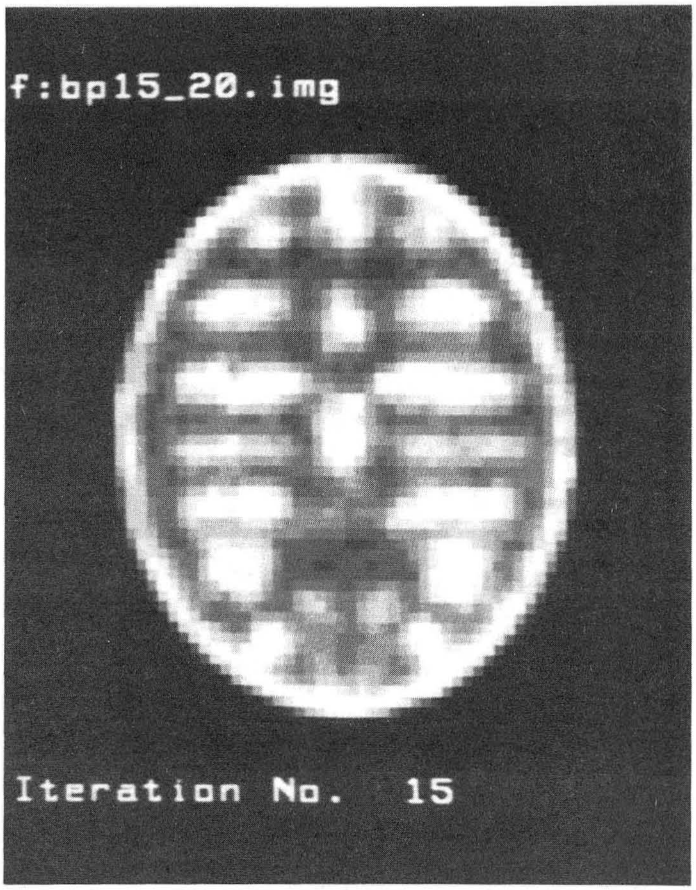
Fig. 2. Reconstructed images for one of the 200,000 count data sets. a) at 15, b) at 21, c) at 27, d) at 36, e) at 60 and f) at 150 iterations.

Fig.3. Distribution of *H* parameter values from the feasibility tests of [3] for 100 reconstructed images at different iteration numbers, a) at 15, b) at21, c) at 27, d) at 36, e) at 60 and f) at 150 iterations. A χ^2 distribution with 9 degrees of freedom is shown for comparison. The values of χ^2 printed correspond to the Pearson's goodness of fit test between the measured data and the χ^2distribution. Critical value for 90% confidence level is approximately 37.



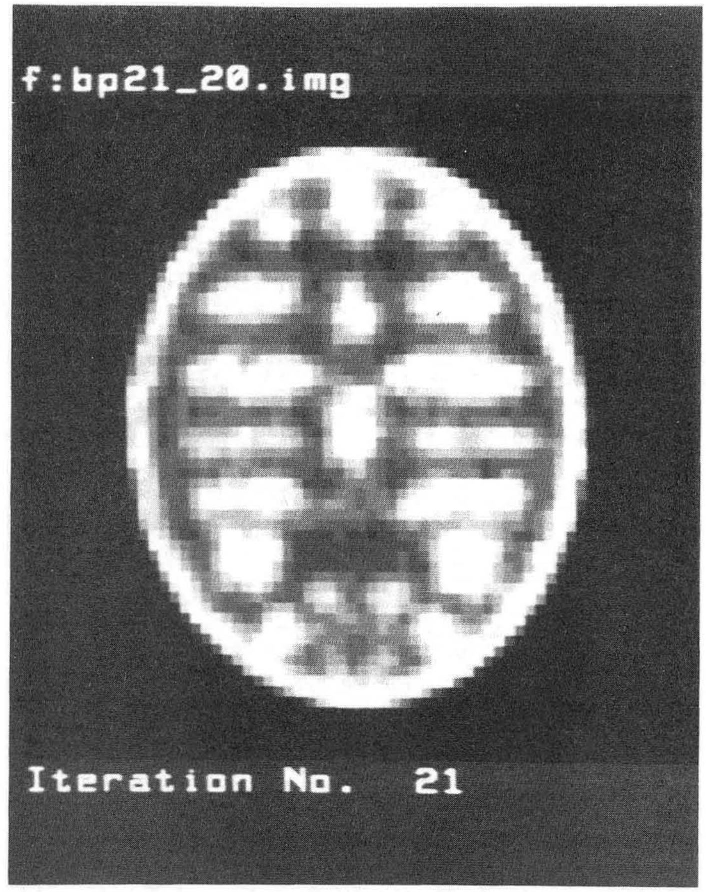
XBB 906-4765

Figure 1



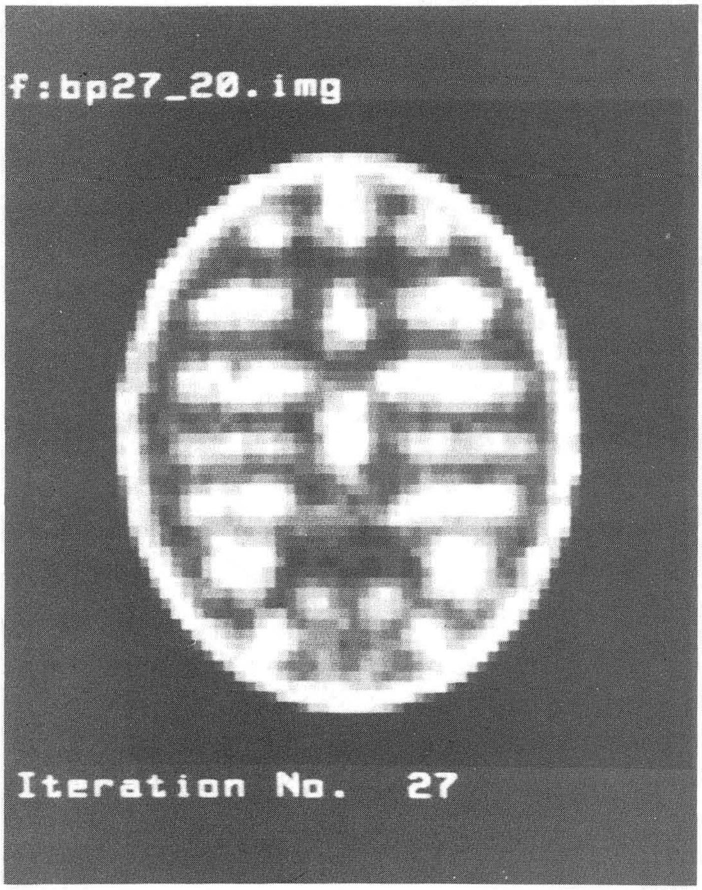
XBB 906-4766

Figure 2a



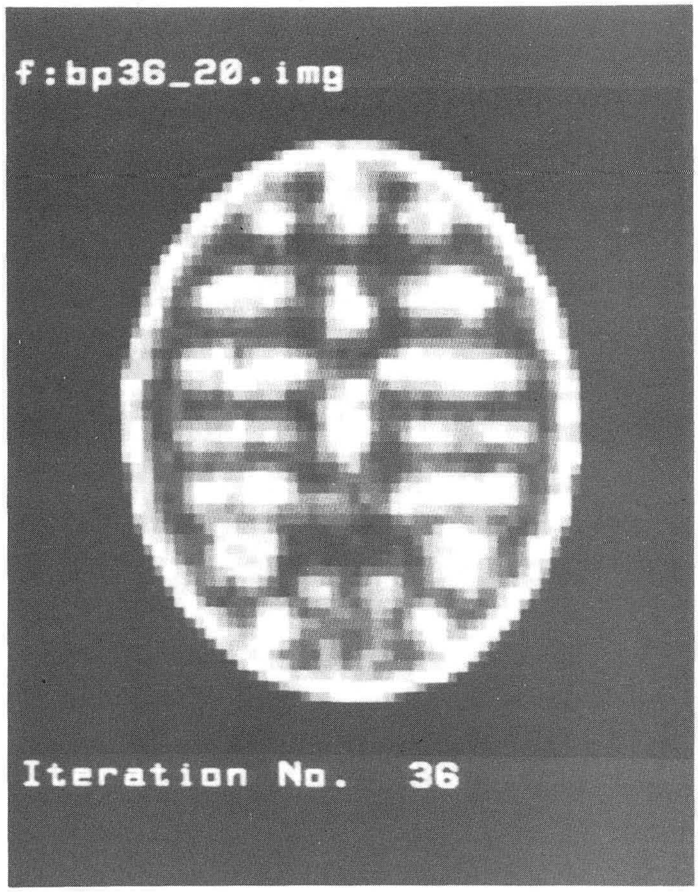
XBB 906-4767

Figure 2b



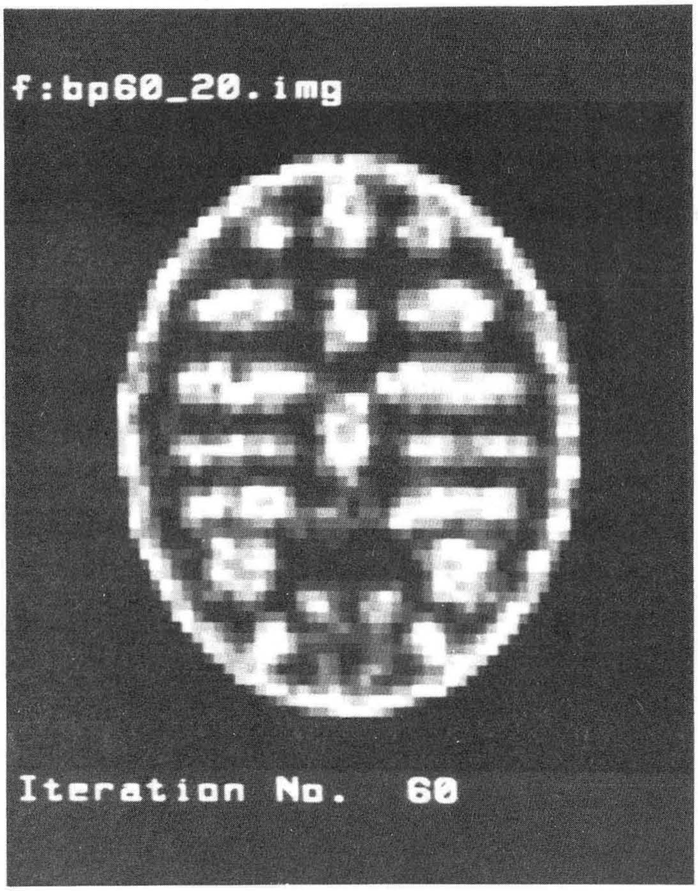
XBB 906-4768

Figure 2c



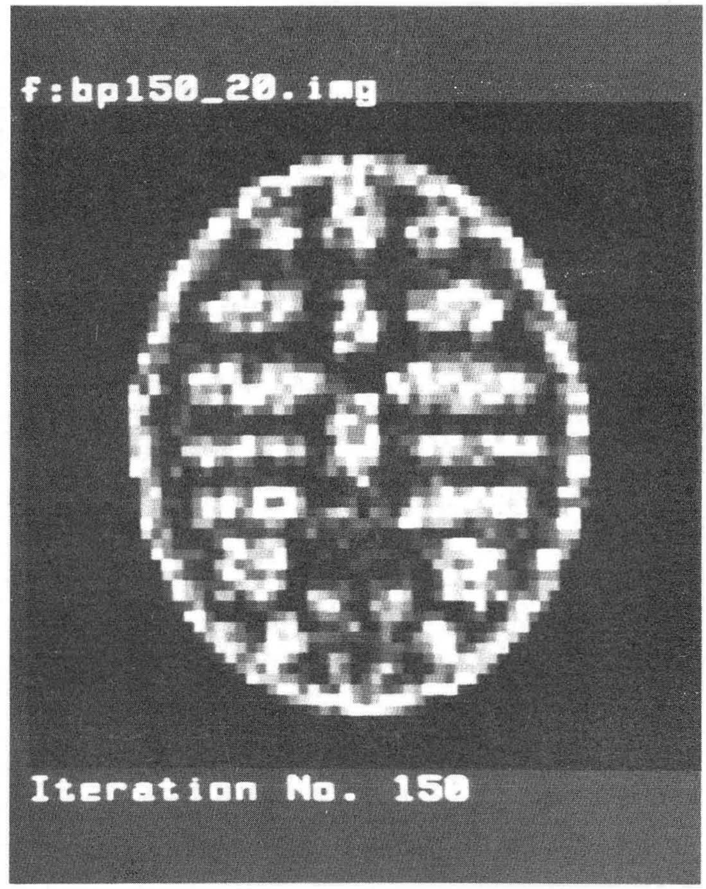
XBB 906-4769

Figure 2d



XBB 906-4770

Figure 2e



XBB 906-4771

Figure 2f

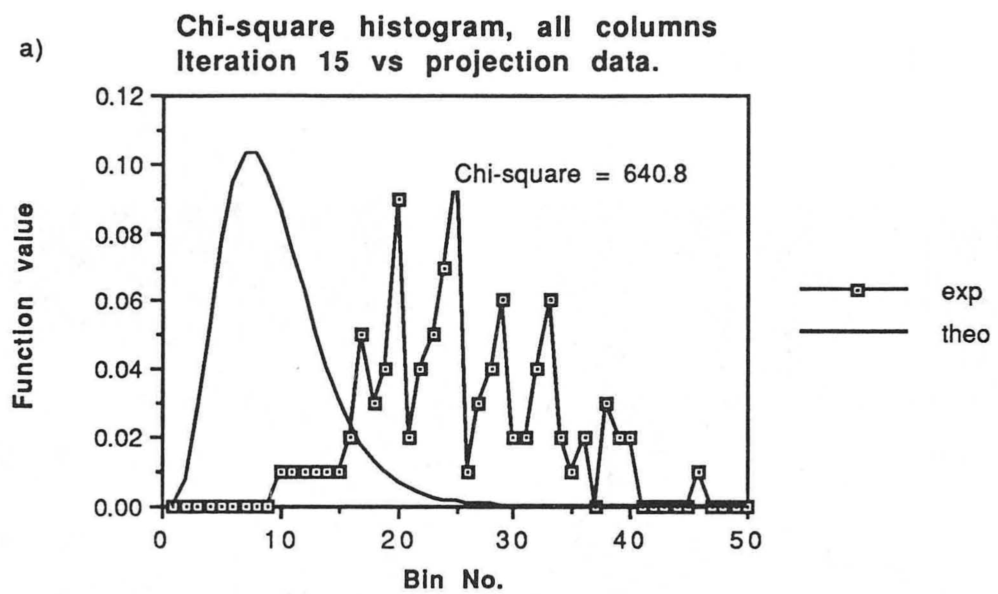


Figure 3a

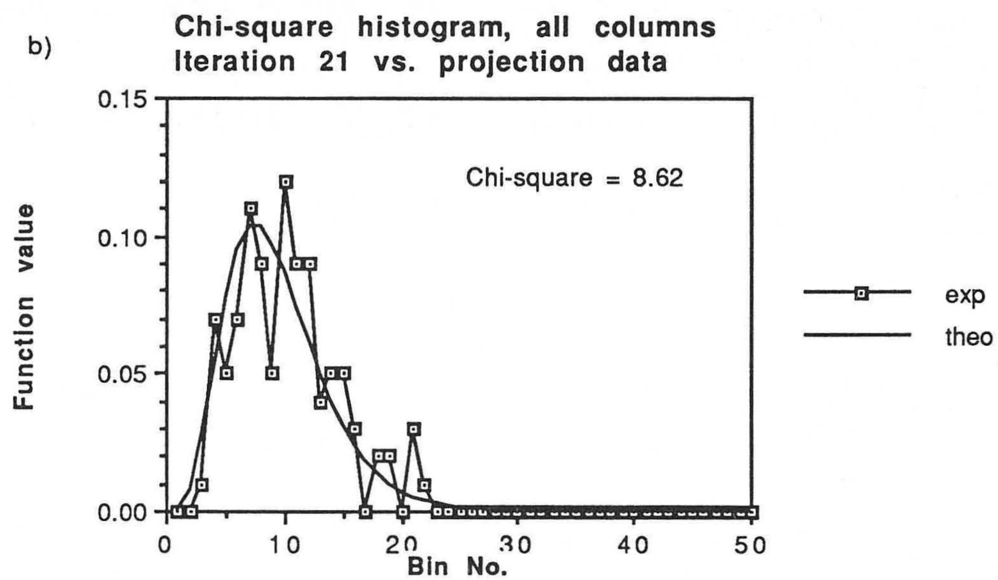


Figure 3b

c)

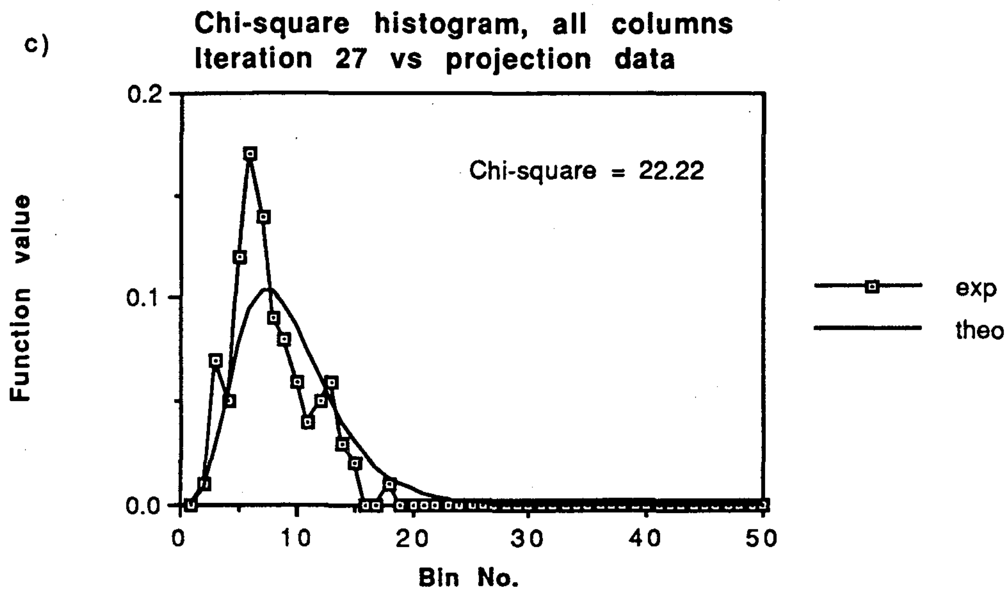


Figure 3c

d)

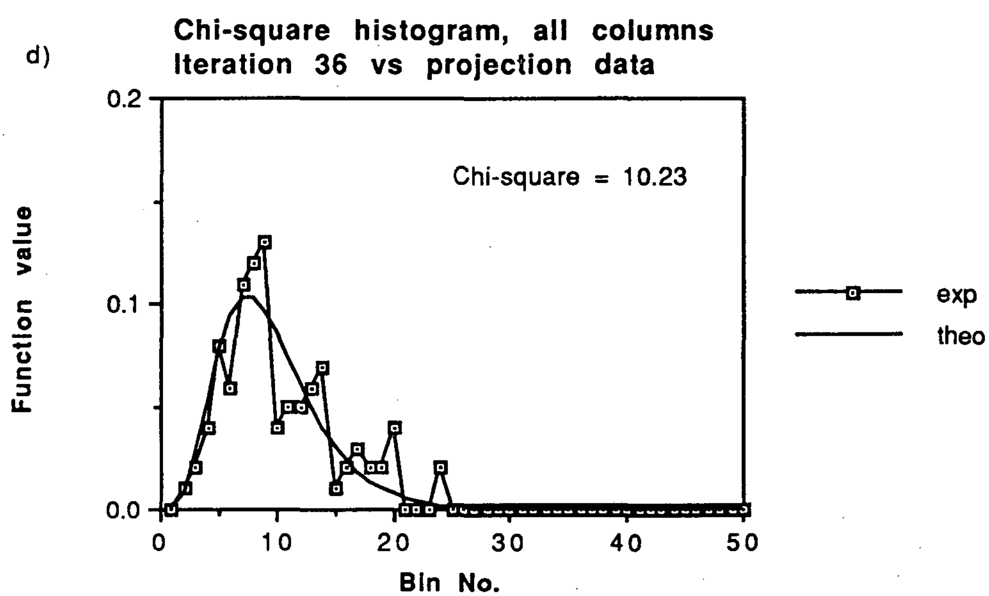


Figure 3d

e)

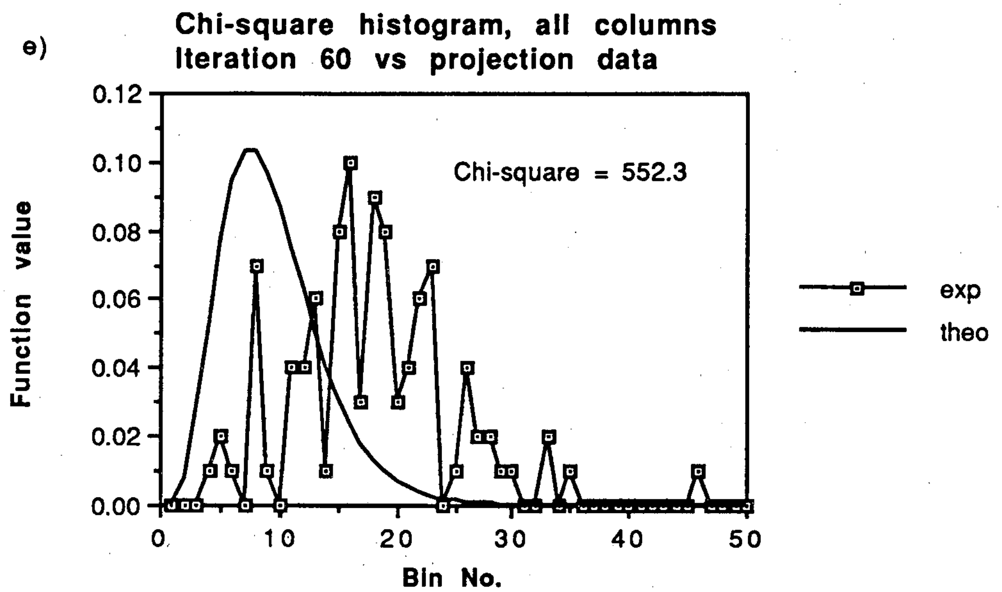


Figure 3e

f)

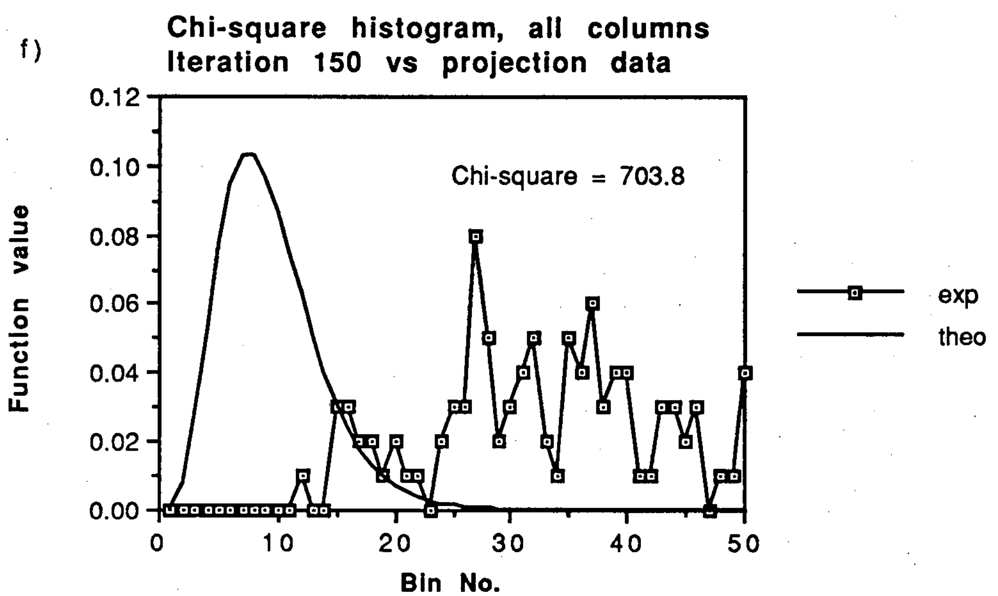


Figure 3f

LAWRENCE BERKELEY LABORATORY
TECHNICAL INFORMATION DEPARTMENT
1 CYCLOTRON ROAD
BERKELEY, CALIFORNIA 94720