Noise analysis of Maximum a Posteriori Expectation Maximization (MAP-EM) algorithm using Gamma Root Priors (GRPs) for Positron Emission Tomography (PET)"

Farah Adeebah Khokhar

Abstract

To model the propagation of photon noise theoretically through reconstruction algorithms using PET data is important to evaluate the reconstructed image quality as a function of parameters of the algorithm. The linearization used to model theoretically the propagation fame an image and a covariance matrix from one iteration to the other using iterative ML-EM (maximum-likelihood expectation maximization). Our case analysis of MAP (maximum a posteriori)-EM algorithms analyzes the EM approach includes prior terms. The MAP-EM algorithm using an independent gamma prior is a special case. For this, we use a Monte Carlo process to justify our theory. The comparison of theoretical evaluations of mean and variance with sample approximations at each iteration shows that our theoretical analysis works well in real where the noise in the reconstructed images do not presume maximal values.

Introduction

The Maximum-A-posteriori (MAP) in anticipation of an increase in the emission computed tomography image reconstruction methods that help to solve basic circumstances. The important application of this method is to gain greater imaging consequences on diagnosis of a particular disease application. The three dimensional functional information of the patient due to the effect of any radiopharmaceutical is observed through the "Emission Computed Tomography (E.C.T) either Positron Emission Tomography (PET) or Single Positron Emission Tomography (SPECT). To study and measure the data of the patient's body, a simple process is needed about which a comparable quantity of radionuclide, "Fluorine- 18 (F-18)" is injected into the subject that fit in the definite organ. The reconstructed images that are formed through the distribution of radionuclide around the body of a patient are assumed to be very useful in finding the volume of radionuclide in other organs of the patient [1, 2].

The pair of oppositely moving positrons emitting through the radionuclides such as F^{18}, used in PET travel through short distance before annihilation process. The event created through this pair is then detected through the line joining to the pair of detectors in any manner known as the line-of-response (LOR) [1, 2]. The coincident event added in the list of detected events if it occurs at the same time approximate (12-20 ns). The collective line-of-responses (LORs) through different angles move along the radius to get the reconstructed image through different reconstruction algorithms [2, 3].

Image Reconstruction is an ill-posed problem

The iteration based noise analysis using PET data, by making a comparison of ideal Shepp Logan Phantom through Maximum Likelihood Expectation Maximum (MLEM), the Maximum Likelihood Expectation through prior (MLEP/ PLEM) and Maximum-A- Posteririeo (MAP) tomographic reconstruction algorithms through "Gamma Root Prior" and will show that how average mean, theoretical mean and average variance and theoretical variance varies with iteration. The noise in the data with its incomplete nature makes the mathematical image reconstruction problem ill-posed and that is the reason of using priors because the prior knowledge will set the previous improper data to the proper. If we use iterative methods without regularization, they cannot compensate for ill-conditioning of the reconstruction problem, thus, produce reconstruction based noise and the only reason is that they use information is limited to the data only and hence, do not use any priori information about the object, however, most importantly, subject of their convergence and low frequency is superimposed by the high frequency noise if large number of iterations is used and make their resolution low, hence, that is the reason iterative methods with regularization is needed to improve the resolution of the reconstruction data better.

Priors

The ill-posed problems are solved through the reconstruction algorithm. The inversion methods are used to solve the inverse operator. The regularized image reconstruction MAP methods that are based on "Gamma Root Priors (GRPs)" have better resolution control with lower blurring nature and better convergence speed. The reconstruction algorithms are used for the study and reduction for the noise using different parameters through the specific data used. The "MLEM (Maximum Likelihood Estimation) algorithm" using priors measures the average and variance

image in accordance to each iteration. The MAP-EM (Maximum a Posteriori Estimation) algorithm incorporating the "Independent Gamma Prior". We used a Monte Carlo (MC) methodology to validate our theoretical analysis to compare the estimation of mean and variance with ideal estimation. Through this comparison it would be shown that this theory works well practically and will also compare our noise propagation results through "Maximum Likelihood Expectation Maximization (MLEM)" and "Maximum Likelihood Expectation through prior (PLEM)".

Theoretical Data of PET Acquisition

The data acquisition in tomography systems is assumed to be ideal. Let "$f(x,y)$", be the estimated number of photons at any location (x,y) of the transverse pixel in the field of view (FOV). Similarly "2-D projection" of "f"on detector bin is a function of $p_\phi(r)$indicating object intensity radially at an angle "ϕ"in "polar coordinates";

$$\boldsymbol{p_\phi(r) = \int_{line(r,\phi)} g(x,y)ds}$$

$$\boldsymbol{p_\phi(r) = \iint_{-\infty}^{\infty} g(x,y)\,\delta(xcos\phi + ysin\phi - r)dxdy} \qquad (1)$$

Where g is a vector of emission data of N × 1 vector. Such whole collection of all projections around the object $\{p_\phi(r): \phi\epsilon[0,\pi], \mathrm{r}\,\epsilon -\infty,\infty\}$ is known as Radon Transform [28]. Theoretically, given the Poisson nature of emission process, $\boldsymbol{y = \mathcal{A}g}$ and this called direct inversion method because $g = \boldsymbol{\mathcal{A}^{-1}y}$ and $\boldsymbol{\mathcal{A}^{-1}}$is ill conditioned due to the noise in the projection, thus

$$\boldsymbol{y = \mathcal{A}g + r} \qquad (2)$$

And $$\boldsymbol{r = y - \mathcal{A}g}$$

Since $\boldsymbol{\mathcal{A}}$ is a sparse, thus iterative methods sound good to compute its inverse.

Where the conditional mean of object dependent noise is;

$$E[\boldsymbol{r}|\boldsymbol{g}] = E(\boldsymbol{y - \mathcal{A}g}|\boldsymbol{g}] = E[\boldsymbol{y}|\boldsymbol{g}] - \boldsymbol{\mathcal{A}g} = \boldsymbol{0}$$

Where,"$M \times 1$"a matrix with all element is equals to zero. The covariance matrix of "**R"** is diagonal matrix with mth diagonal element denoting the variance of$\boldsymbol{y}_m = [\boldsymbol{\mathcal{A}g}]_m$. Thus

$$Cov_{R\,|\,g} = Cov_{X\,|\,g} = diag(\boldsymbol{\mathcal{A}g})$$

Where

$Cov_{R\,|\,g} \equiv E[\boldsymbol{rr^T}|\boldsymbol{g}]$, andconditional covariance matrix of **X**, the notation $diag(\boldsymbol{\mathcal{A}g})$ is a diagonal matrix equal to the element of vector$[\boldsymbol{\mathcal{A}g}]_m$. However, the following conventions are used for convenient;

- The capital letters in the above equations denote the random vectors corresponding with their random variables
- The small alphabetic letters, however, denote deterministic vectors with corresponding with their vector elements
- The italic style letters denote matrices corresponding to their matrix elements.
- While using convenient "Hadamard" notations [50], **cd** and **c/d** are vectors whose nth components are $c_n d_n$ and c_n/d_n
- Dot and matrix products are defines as $\boldsymbol{cd}$ and $\boldsymbol{\mathcal{A}c}$.

ML-EM Algorithm

The "Emission Computed Tomography" data can be written linearly as follows;

$$\boldsymbol{y} \sim Poisson\,[\boldsymbol{\mathcal{A}g} + \boldsymbol{r}]$$

Here system matrix $\boldsymbol{\mathcal{A}}$ and background noise **r,** are non-negative constants. If the background noise is neglected, thus any m[th] projection $\mathbf{y_m}$ may be represented with mean as;

$$\boldsymbol{y_m} = \sum_{n=1}^{N} \boldsymbol{\mathcal{A}}_{mn}\, \boldsymbol{g}_n \qquad (3)$$

Where $\boldsymbol{\mathcal{A}}_{mn}$ be the mnth elements of the system matrix, thus with constant projection vector$\boldsymbol{y}$, likelihood function written as;

$$likelihood = P_y(\boldsymbol{y}|\boldsymbol{g}) = \prod_{\boldsymbol{m=1}}^{\boldsymbol{M}} \boldsymbol{e}^{\boldsymbol{y_m}} \frac{(\boldsymbol{y_m})^{\boldsymbol{y_m}}}{\boldsymbol{y}!} \qquad (4)$$

Recognizing that projections $\boldsymbol{y}$ consists of independent observations $\{\boldsymbol{y}_1, \boldsymbol{y}_2 \ldots\ldots\ldots\}$, we seek that value$\boldsymbol{g}$ which maximizes

$$\prod_{\boldsymbol{y_m} \in \boldsymbol{y}} P_X(\boldsymbol{y}_m|\boldsymbol{g})$$

MAP-EM algorithm

The Bayes' maximum a posteriori estimate (MAP), denoted $\hat{\boldsymbol{g}}$ for the parameter set $\boldsymbol{g}$, using Bayes' rule

$$posterior = \frac{likelihood\ .\ prior}{projection}$$

$$P_y(\boldsymbol{g}|\boldsymbol{y}) = \frac{P_y(\boldsymbol{y}|\boldsymbol{g})P_F(\boldsymbol{g})}{P_y(\boldsymbol{y})} \qquad (5)$$

The probability density function $P_F(\boldsymbol{g})$of object known as prior probability density.

Here $P_y(\boldsymbol{y})$ being normalizing factor is assumed as constant.

$$P_y(\boldsymbol{y}) = \int P_y(\boldsymbol{y}|\boldsymbol{g})P_F(\boldsymbol{g})\, d\boldsymbol{g}$$

Whereas, a maximum-a-posteriori estimate of the unknown object after considering projection data and prior maximizes the probability density function

$$\hat{\boldsymbol{g}} = arg\max_{\boldsymbol{g}}\{\frac{P_y(\boldsymbol{y}|\boldsymbol{g})P_F(\boldsymbol{g})}{P_y(\boldsymbol{y})}\}$$

In image reconstruction problems the prior information is valuable as data in case of noisy data or imperfect data but useless if data is complete. [12]

Taking log of the MAP density function, dropping the denominator

$$\hat{\boldsymbol{g}} = \mathbf{max}_{\boldsymbol{g}} \log\big(P_y(\boldsymbol{y}|\boldsymbol{g}\big) + \boldsymbol{log}P_F(\boldsymbol{g})\} \qquad (6)$$

Placing derivative with respect to pixel $\boldsymbol{g_n}$equal to zero to find iterative update equation

$$\frac{\partial Z(\boldsymbol{F})}{\partial \boldsymbol{g}} = \mathbf{0}$$

Let; $$Z(\boldsymbol{g}) = \log\big(P_y(\boldsymbol{y}|\boldsymbol{g}\big) + \boldsymbol{log}P_G(\boldsymbol{g})\} \qquad (7)$$

The solution must satisfy the Kuhn-Tuckers optimally conditions of the problem

$$\frac{\partial Z}{\partial \boldsymbol{f_n}} \geq 0, \forall n$$

$$\boldsymbol{g_n}\frac{\partial Z}{\partial \boldsymbol{f_n}} = 0, \forall n$$

Solving Z with respect to equation. (8), substituting it in 2nd condition

$$g_n\left\{\sum_{m=1}^{M}\boldsymbol{\mathcal{A}}_{mn} - \sum_{m=1}^{M}\frac{x_m}{\boldsymbol{\mathcal{A}}_{mn}\boldsymbol{g}_n}\boldsymbol{\mathcal{A}}_{mn} + \frac{\partial}{\partial \boldsymbol{g_n}}\boldsymbol{log}P_G(\boldsymbol{g})\right\} = 0 \qquad (8)$$

Using "EM method", maximizing equation. (6), is equaling iteratively maximizing a function$Z(\boldsymbol{g}|\hat{\boldsymbol{g}}^{\boldsymbol{k}})$, gives the previous estimate,

$$\partial Z(\boldsymbol{g}|\hat{\boldsymbol{g}}^{k})^{MAP-EM} = -\boldsymbol{u}^T\boldsymbol{g} + (\boldsymbol{v}^{\boldsymbol{k}})^{\boldsymbol{T}}log\boldsymbol{g} + logP_G(\boldsymbol{g})$$

Where, $\boldsymbol{u} \equiv \boldsymbol{\mathcal{A}}^T \mathbf{1}$ is a $M \times 1$ sensitivity vector equals to unity;

By taking derivative of above relation "MAP-EM algorithm" is derived as

$$\frac{\partial Z(\boldsymbol{g}|\hat{\boldsymbol{g}}^k)^{MAP-EM}}{\partial \boldsymbol{g}} = -u + \frac{\hat{v}^k}{\boldsymbol{g}} + \frac{\partial log P_G(\boldsymbol{g})}{\partial \boldsymbol{g}} = 0 \qquad (9)$$

In case of no prior, we left with solution $u + \frac{\hat{v}^k}{\boldsymbol{g}} = 0$and this leads to "ML-EM algorithm"

$$\hat{\boldsymbol{g}}^{k+1,ML-EM} = \frac{\hat{v}^k}{\boldsymbol{u}} = \frac{\hat{\boldsymbol{g}}^k}{u}\left\{\boldsymbol{\mathcal{A}}^T\left[\frac{y}{\boldsymbol{\mathcal{A}}\hat{\boldsymbol{g}}^k}\right]\right\} \qquad (10)$$

Green et al [13] suggested "One-Step-Late (OSL) algorithm" to use [49], thus rearranging equation we will get

$$\hat{\boldsymbol{g}}^{k+1,OSL} = \frac{\boldsymbol{g}_n^k}{\left\{\sum_{m=1}^M \boldsymbol{\mathcal{A}}_{mn} + \frac{\partial}{\partial \boldsymbol{g}_n} \boldsymbol{log} P_G(\boldsymbol{g})\right\}} \sum_{m=1}^M \frac{x_m}{g} \boldsymbol{\mathcal{A}}_{mn} \quad (11)$$

$$\boldsymbol{g}_n \geq 0, for\ all\ n;$$

In vector notation can be written as

$$\hat{\boldsymbol{g}}^{k+1,OSL} = \frac{1}{u + \frac{\partial}{\partial \boldsymbol{g}_n} \boldsymbol{log} P_G(\boldsymbol{g})\Big|_{\boldsymbol{g}=\boldsymbol{g}^k}} \left[\frac{\boldsymbol{y}}{\boldsymbol{\mathcal{A}}\boldsymbol{g}^k}\right] \boldsymbol{\mathcal{A}}^t \boldsymbol{g}^k \qquad (12)$$

Theoretical Noise Propagation

The theoretical noise propagation according to the "Barrett et al 1994" linearize the statistics of noise propagation in equation (13) and (20) including the effects of prior terms. The noise at iteration"*k+1"*is related to noise at previous iteration k, thus in iterative reconstruction algorithm we decompose each random vector by ignore the quadratic noise terms. The two "Taylor series" expansions are;

$$\log\left(\boldsymbol{a}^k + \boldsymbol{R}_g^k\right) = log\boldsymbol{a}^k + log\left(\mathbf{1} + \frac{\boldsymbol{R}_f^k}{\boldsymbol{a}^k}\right) \cong log\boldsymbol{a}^k + \frac{\boldsymbol{R}_f^k}{\boldsymbol{a}^k} \qquad (A)$$

$$\left(\boldsymbol{a}^k + \boldsymbol{R}_g^k\right)^{-1} \cong \frac{1}{\boldsymbol{a}^k}\left(1 + \frac{\boldsymbol{R}_f^k}{\boldsymbol{a}^k}\right)^{-1} = \frac{1}{\boldsymbol{a}^k}\left(1 - \frac{\boldsymbol{R}_f^k}{\boldsymbol{a}^k}\right) \qquad (B)$$

These approximations based on that $|signal| \gg |noise| or \frac{\boldsymbol{R}_f^k}{\boldsymbol{a}^k} \ll 1$. Here $\boldsymbol{H} = \boldsymbol{\mathcal{A}}\boldsymbol{g} + \boldsymbol{R}$ is considered to be random detector measurement variable for which$\boldsymbol{y}$ is noisy projection. Where random vector $\boldsymbol{G}^k = \boldsymbol{a}^k + \boldsymbol{R}_g^k$ is decomposed into mean image or expectation $\boldsymbol{a}^k \equiv \boldsymbol{E}\left[\boldsymbol{G}^k \middle| \boldsymbol{g}\right]$

and $\boldsymbol{R_g^k}$ is noise terms or the deviation of $\boldsymbol{G^k}$ from its mean$\boldsymbol{a^k}$ where deviation has zero mean itself.

MAP-EM with Independent Gamma Root Prior

The probability density function (pdf) of gamma prior is defined as

$$P_{\boldsymbol{F}}(\boldsymbol{g}) = \prod_n \frac{1}{\Gamma(\alpha_n)} ({}^{\alpha_n}/_{\beta_n})^{\alpha_n} g_n^{\alpha_n - 1} \exp(-\frac{\alpha_n g_n}{\beta_n})$$

Where β_n is mean image, ${\alpha_n}^{-\frac{1}{2}}$ is the variance co-efficient respectively. Taking log and derivative on both sides and dropping terms independent of$\boldsymbol{g}$,

$$\frac{\partial log P_{\boldsymbol{F}}(\boldsymbol{g})}{\partial \boldsymbol{g}} = \frac{\partial}{\partial \boldsymbol{g}} \sum_n [\,(\alpha_n - 1) log g_n - \alpha_n g/\beta_n] \; = \frac{(\alpha_n - 1)}{\boldsymbol{g}} - ({}^{\alpha_n}/_{\beta_n})$$

Putting these values in equation. (9) and substituting $\hat{v}^k \equiv [\frac{y}{\boldsymbol{\mathcal{A}^T g^k}}]\boldsymbol{\mathcal{A}^T g^k}$we get

$$(\hat{\boldsymbol{g}}^{\boldsymbol{k+1}})^{MAP-EM} = \frac{\hat{\boldsymbol{g}}^{\boldsymbol{k}}}{u + {}^{\alpha_n}/_{\beta_n}} \left[\boldsymbol{\mathcal{A}}^T \left[\frac{y}{\boldsymbol{\mathcal{A}}^T \hat{\boldsymbol{g}}^{\boldsymbol{k}}}\right] + \frac{(\alpha_n - 1)}{\hat{\boldsymbol{g}}^{\boldsymbol{k}}}\right] \qquad (20)$$

Where $\boldsymbol{g}$and $\boldsymbol{d}$ are "$N \times 1$" vectors with elements$(\alpha_n - 1)$and ${}^{\alpha_n}/_{\beta_n}$.

Taking log we can write above equation as

$$log \hat{\boldsymbol{G}}^{\boldsymbol{k+1}} = log \hat{\boldsymbol{G}}^{\boldsymbol{k}} + \log \boldsymbol{\mathcal{A}}^T \left[\frac{\boldsymbol{H}}{\boldsymbol{\mathcal{A}}^T \hat{\boldsymbol{G}}^{\boldsymbol{k}}}\right] + \log \left[\frac{\boldsymbol{g}}{\hat{\boldsymbol{G}}^{\boldsymbol{k}}}\right] - \log(\boldsymbol{u} + \boldsymbol{d}) \; (21)$$

We get solutions of first and 2nd term of above equation from "median root prior" section and third term will be

$$\left[\frac{\boldsymbol{g}}{\hat{\boldsymbol{F}}^{\boldsymbol{k}}}\right] = \left[\frac{\boldsymbol{g}}{\boldsymbol{a^k} + \boldsymbol{a^k R_y^k}}\right] \cong \frac{\boldsymbol{g}}{\boldsymbol{a^k}}(1 - \boldsymbol{R_y^k}) = \left[\frac{\boldsymbol{g}}{\boldsymbol{a^k}} - \frac{\boldsymbol{g R_y^k}}{\boldsymbol{a^k}}\right]$$

Now combining solution of all three terms and the small quantity of quadratic terms that are$\boldsymbol{R_y^k}$and $\boldsymbol{R}$ will be dropped

$$log \boldsymbol{a^{k+1}} + \boldsymbol{R_y^{k+1}} = log \boldsymbol{a^k} + \boldsymbol{R_y^k} + \log \left\{\boldsymbol{\mathcal{A}^t} \left[\frac{\boldsymbol{\mathcal{A}f}}{\boldsymbol{\mathcal{A}a^k}}\right] + \frac{\boldsymbol{g}}{\boldsymbol{a^k}} + \boldsymbol{\mathcal{A}^t} \left[\frac{\boldsymbol{R}}{\boldsymbol{\mathcal{A}a^k}}\right] - \boldsymbol{\mathcal{A}^t} - \frac{\boldsymbol{g R_y^k}}{\boldsymbol{a^k}}\right\}$$

Using Taylor series expansion (A) to yield

$$= log\boldsymbol{a^k} + \boldsymbol{R_y^k} + \log\left\{\boldsymbol{\mathcal{A}^t}\left[\frac{\boldsymbol{\mathcal{A}f}}{\boldsymbol{\mathcal{A}a^k}}\right]\right\} + \frac{\boldsymbol{g}}{\boldsymbol{a^k}} + \frac{\boldsymbol{\mathcal{A}^t}\left[\frac{\boldsymbol{R}}{\boldsymbol{\mathcal{A}a^k}}\right] - \boldsymbol{\mathcal{A}^t}\left[\frac{[\boldsymbol{\mathcal{A}f}][\boldsymbol{\mathcal{A}}(\boldsymbol{a^k R_y^k})]}{[\boldsymbol{\mathcal{A}a^k}]}\right] - -\frac{\boldsymbol{gR_y^k}}{\boldsymbol{a^k}}}{\boldsymbol{\mathcal{A}^t}\left[\frac{\boldsymbol{\mathcal{A}f}}{\boldsymbol{\mathcal{A}a^k}}\right] + \frac{\boldsymbol{g}}{\boldsymbol{a^k}}}$$

Inserting the above solution in equation no. (15)

$$= log\boldsymbol{a^k} + \boldsymbol{R_y^k} + \log\left\{\frac{\{\boldsymbol{\mathcal{A}^t}\left[\frac{\boldsymbol{\mathcal{A}f}}{\boldsymbol{\mathcal{A}a^k}}\right] + \frac{\boldsymbol{g}}{\boldsymbol{a^k}}}{(\boldsymbol{u}+\boldsymbol{d})}\right\} + \frac{\boldsymbol{\mathcal{A}^t}\left[\frac{\boldsymbol{R}}{\boldsymbol{\mathcal{A}a^k}}\right] - \boldsymbol{\mathcal{A}^t}\left[\frac{[\boldsymbol{\mathcal{A}}][\boldsymbol{\mathcal{A}}(\boldsymbol{a^k R_y^k})]}{[\boldsymbol{\mathcal{A}a^k}]}\right] - \frac{\boldsymbol{gR_y^k}}{\boldsymbol{a^k}}}{\boldsymbol{u} + \frac{\boldsymbol{g}}{\boldsymbol{a^k}}}$$

Now by equating random and non-random terms, we get

$$log\boldsymbol{a^{k+1}} = log\boldsymbol{a^k} + \log\left\{\frac{\{\boldsymbol{\mathcal{A}^t}\left[\frac{\boldsymbol{\mathcal{A}f}}{\boldsymbol{\mathcal{A}a^k}}\right] + \frac{\boldsymbol{g}}{\boldsymbol{a^k}}}{(\boldsymbol{u}+\boldsymbol{d})}\right\} \qquad (21)$$

$$\boldsymbol{R_y^{k+1}} = \boldsymbol{R_y^k} + \frac{\boldsymbol{\mathcal{A}^t}\left[\frac{\boldsymbol{R}}{\boldsymbol{\mathcal{A}a^k}}\right] - \boldsymbol{\mathcal{A}^t}\left[\frac{\boldsymbol{\mathcal{A}}(\boldsymbol{a^k R_y^k})}{[\boldsymbol{\mathcal{A}a^k}]}\right] - \frac{\boldsymbol{gR_y^k}}{\boldsymbol{a^k}}}{\boldsymbol{u} + \frac{\boldsymbol{g}}{\boldsymbol{a^k}}} \qquad (22)$$

Here equation (16) predicts through running noise free "MAP-EM algorithm", the ensemble mean reconstruction $\boldsymbol{a^k}$ with **"kth"** iteration can be calculated on phantom"$\boldsymbol{f}$ ".

Equation (17) estimates the iteration *k* for noise in the log of reconstruction,

$$\boldsymbol{R_y^{k+1}} = \boldsymbol{P^k R} + (\boldsymbol{Q^k} - \boldsymbol{S^k})\boldsymbol{R_y^k} \qquad (23)$$

$$\mathbf{P}^{\boldsymbol{k}}\boldsymbol{R_y^k} = \frac{\boldsymbol{\mathcal{A}^t}\left[\frac{\boldsymbol{\mathcal{A}}(\boldsymbol{a^k R_y^k})}{[\boldsymbol{\mathcal{A}a^k}]}\right]}{\boldsymbol{u} + \frac{\boldsymbol{g}}{\boldsymbol{a^k}}}$$

$$\boldsymbol{Q^k R} = \frac{\boldsymbol{\mathcal{A}^t}\left[\frac{\boldsymbol{R}}{\boldsymbol{\mathcal{A}a^k}}\right]}{\boldsymbol{u} + \frac{\boldsymbol{g}}{\boldsymbol{a^k}}}$$

$$\boldsymbol{S^k R_y^k} = \frac{\boldsymbol{uR_y^k}}{\boldsymbol{u} + \frac{\boldsymbol{g}}{\boldsymbol{a^k}}}$$

And we can write these matrices explicitly

$$\mathbf{P}^{\boldsymbol{k}} = diag\left(\frac{1}{\boldsymbol{u} + \frac{\boldsymbol{g}}{\boldsymbol{a}^{\boldsymbol{k}}}}\right)\boldsymbol{\mathcal{A}}^{\boldsymbol{t}} diag\left(\frac{1}{\boldsymbol{\mathcal{A}}\boldsymbol{a}^{\boldsymbol{k}}}\right)\boldsymbol{\mathcal{A}} diag(\boldsymbol{a}^{\boldsymbol{k}})$$

$$\boldsymbol{Q}^{\boldsymbol{k}} = diag\left(\frac{1}{\boldsymbol{u} + \frac{\boldsymbol{g}}{\boldsymbol{a}^{\boldsymbol{k}}}}\right)\boldsymbol{\mathcal{A}}^{\boldsymbol{t}} diag\left(\frac{1}{\boldsymbol{\mathcal{A}}\boldsymbol{a}^{\boldsymbol{k}}}\right)$$

$$\boldsymbol{S}^{\boldsymbol{k}} = diag\left(\frac{u}{\boldsymbol{u} + \frac{\boldsymbol{g}}{\boldsymbol{a}^{\boldsymbol{k}}}}\right)$$

Experiment

For a 64 × 64 "Shepp Logan phantom", we use two different count levels that are low counts 80,000 and high counts 500,000 counts, theoretically adding the Poisson noise. The noise free and noisy data then compared relatively.

The sample "mean" of the L reconstructions at ith pixel, denoted as G_i^k ;

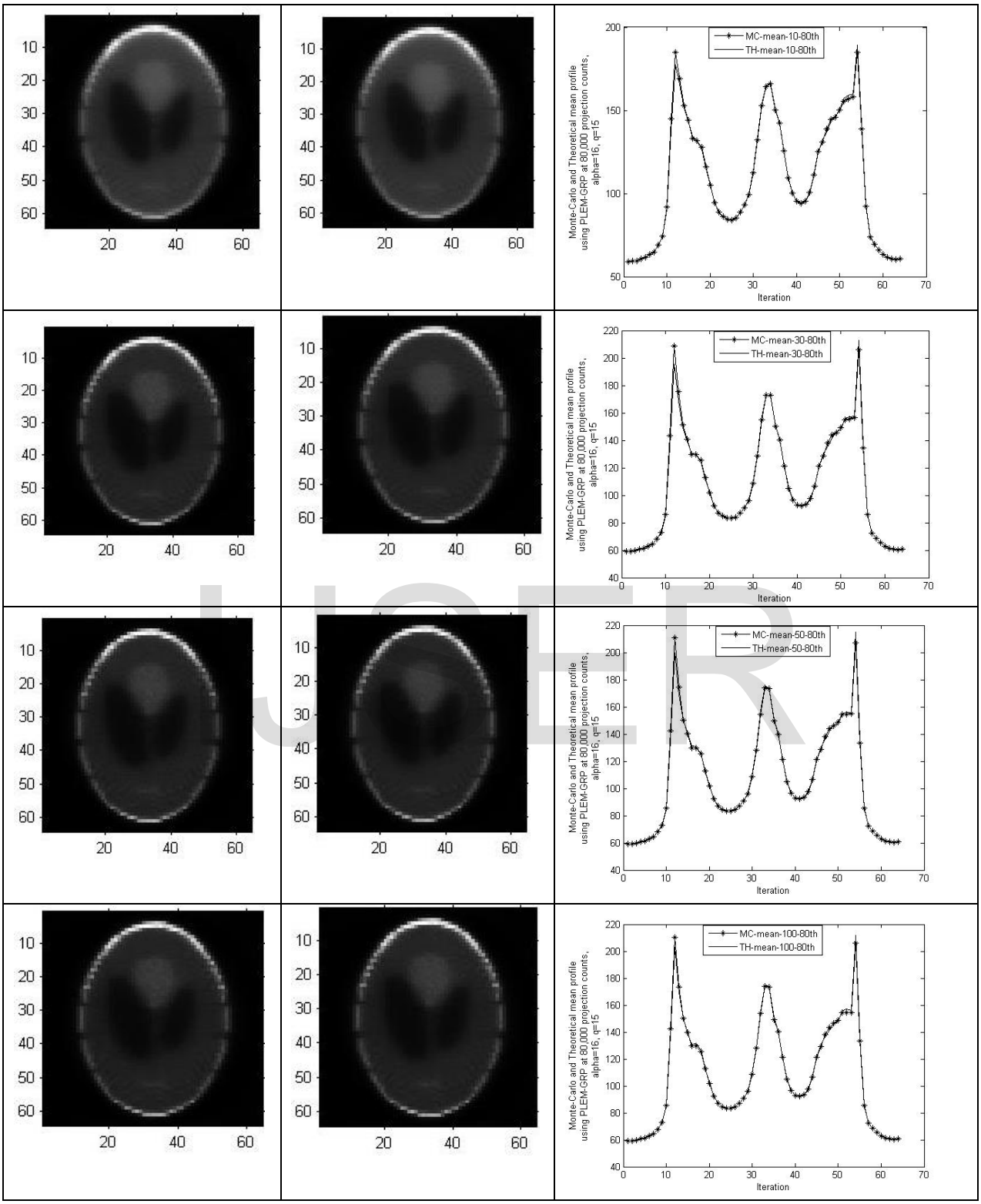
$$G_i^k = \frac{1}{L}\sum_{l=1}^{L} g_i^{k,l} \text{ for i = 1… N,}$$

Where the subscripts k and L denote the iteration number and the sample number respectively and $g_i^{k,l}$ is the object estimate at ith pixel at iteration k. The variance of the reconstruction between two different pixels ith and jth denoted as $[\mathrm{K}_G^k]ij$;

$$[\mathrm{K}_G^k]ij = \frac{1}{L-1}\sum_{l=1}^{L}(g_i^{k,l} - G_i^k)(g_j^{k,l} - G_j^k) \text{ for i, j = 1,…, N.}$$

Results and Discussions

For the "MAP-EM" using Independent Gamma Root Prior, we used a Shepp Logan phantom at 80,000 counts,$\alpha = 16$ and 500,000 counts, $\alpha = 30$, whereas, β is the weight of the pixel. The average is taken within the "disk of region (ROI)". For independent gamma prior, the "ROI" here is the disk itself. The description of general theoretical formula estimates the variance (K_δ) of reconstruction. This is performed mathematically depending on the low and high counts statistically corresponding to their respective variances.



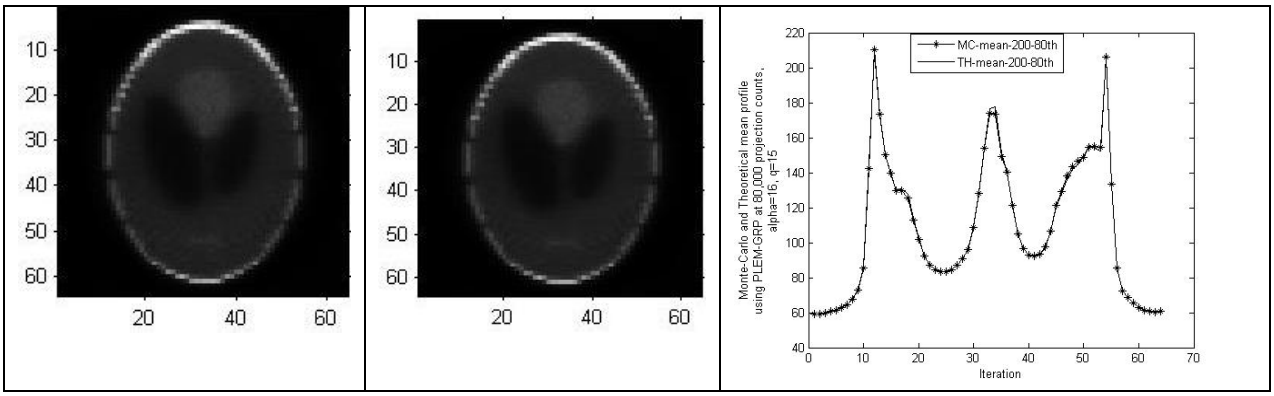
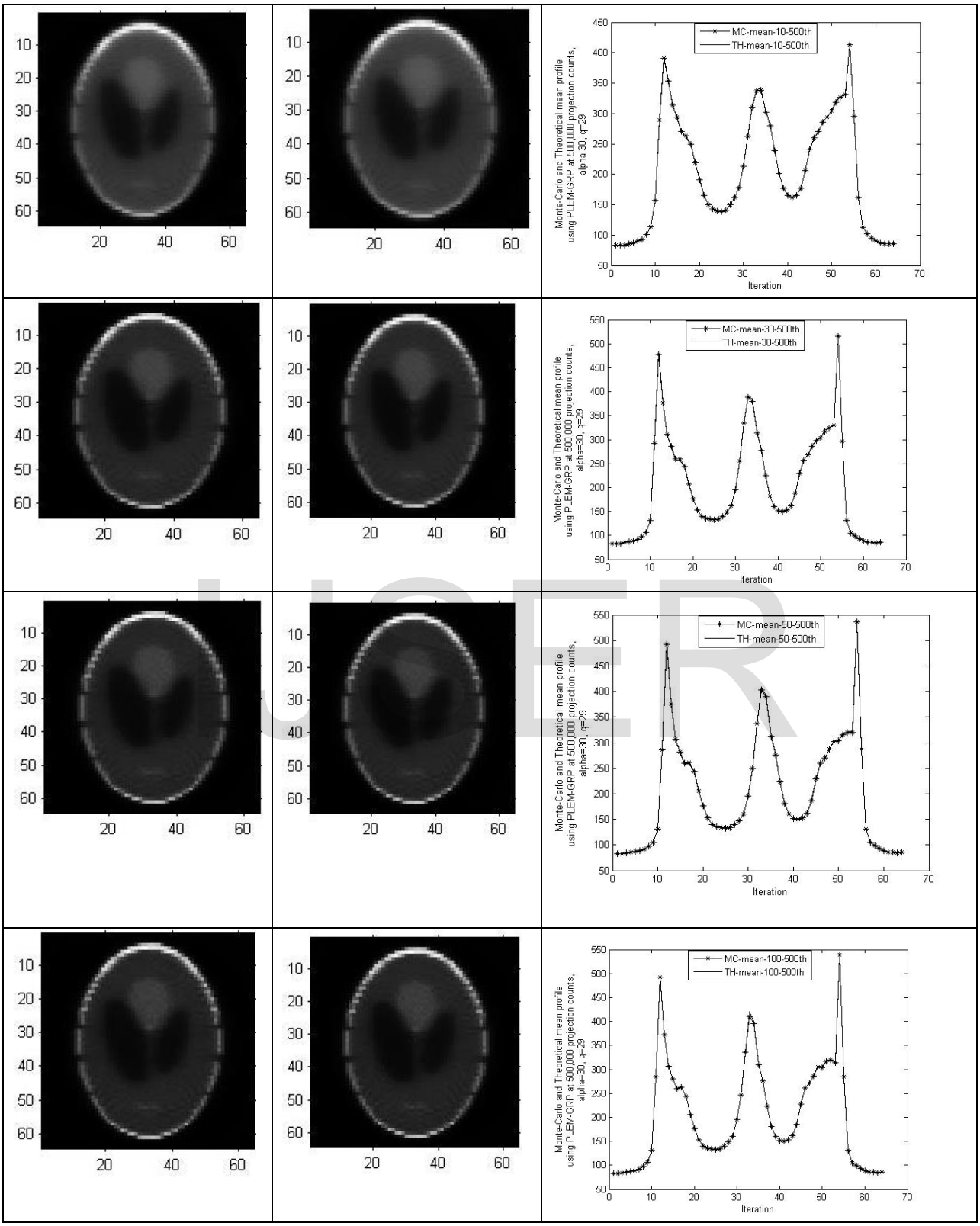


Figure 1: Average and theoretical mean corresponding to their profiles at each iteration (10, 30, 50, 100, and 200) using PLEM and MAP-EM algorithm through GRP at 80,000 projection counts, α = 16, q = 15 and β = value of the phantom.

The "Monte Carlo" and theoretical mean images of total 500 noisy reconstructions at different iterations such as 10,30, 50, 100 and 200 are shown in figure (1) whereas, the total projection counts are 80,000 counts. The top left corner of the figure (1) shows the noisy mean images of 500 noisy images at 80,000 total projection counts reconstructed through nonlinear "Penalized Maximum-Likelihood Expectation Maximization (PLEM)" algorithm using "Gamma Root Prior (GRP)" for iterations equal to "10, 30, 50, 100 and 200"respectively, whereas, the middle portion presents the theoretical mean images at" 80,000 total projection counts" for iteration numbers 10, 30, 50, 100 and 200 running through "linear Maximum-a-Posteriori-Expectation Maximization (MAP-EM) algorithm"using $\alpha = 16, q = 15$, "β = value of the phantom" corresponding with their respective profiles. Here we checked the "Monte-Carlo (MC)" simulations with 80,000 counts that is the mean are dependent at each iteration 10, 30, 50, 100 and 200. Figure 1 show the excellent agreement of profiles of mean images through the lesion center at 80,000 projection counts at different iterations 10, 30, 50, 100 and 200 respectively for MC and theory. However, it is to be noted that for a linear reconstruction algorithm, the noise-free reconstruction and ensemble mean reconstruction are exactly equal. This equality is very well defined through the fact that is well approximated by the nonlinear MAP-EM is gratifying.



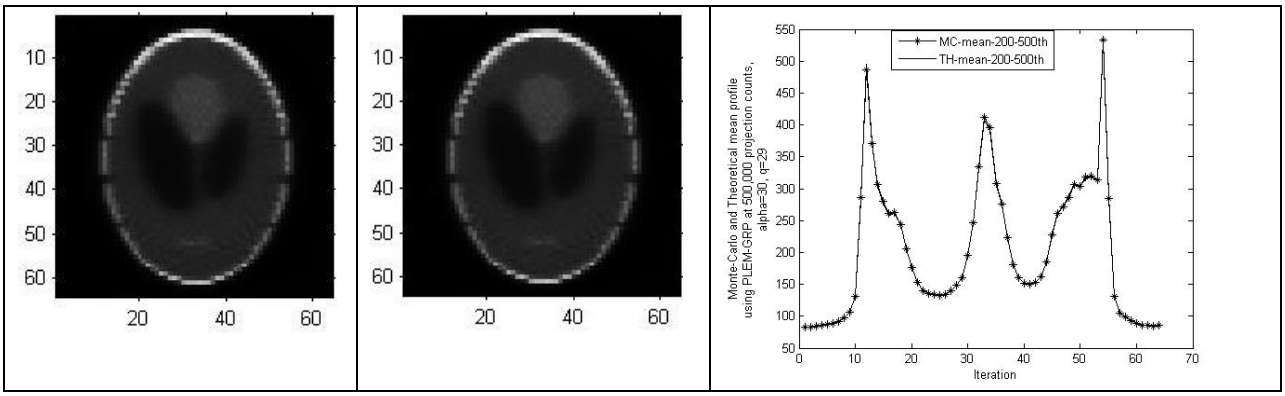
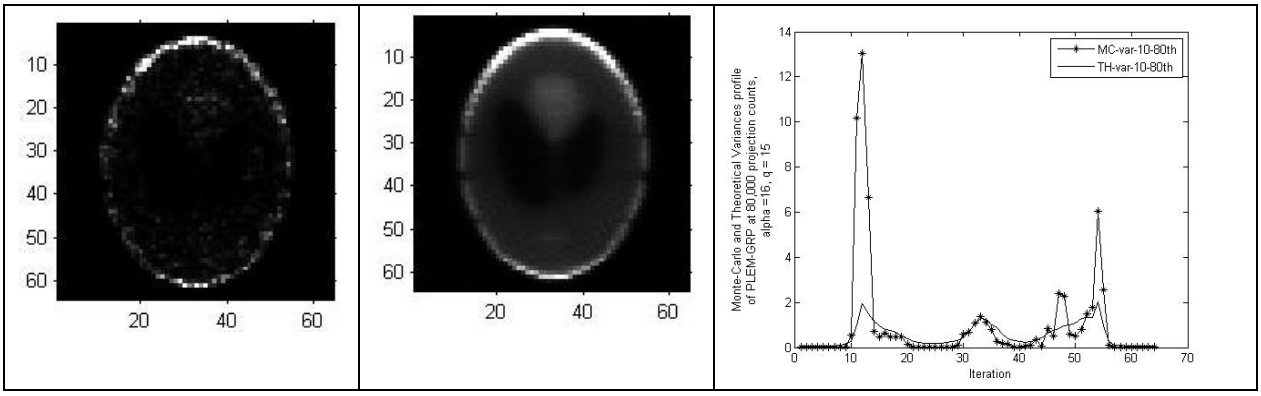


Figure 2: Average and theoretical mean corresponding to their profiles at each iteration (10, 30, 50, 100, and 200) using nonlinear PLEM and linear MAP-EM algorithm through GRP at 500,000 projection counts, α=30, q =29 and β = value of the phantom.

The Monte Carlo and theoretical mean images of total 500 noisy reconstructions at iterations 10,30, 50, 100 and 200 are shown in figure (2) whereas, projection counts are used as 500,000 counts. The top left corner of the figure (2) shows the noisy mean images of 500 noisy images at 500,000 total projection counts reconstructed through nonlinear PLEM (Penalized Maximum-Likelihood Expectation Maximization) algorithm using Gamma Root Prior (GRP) for iterations equal to 10, 30, 50, 100 and 200 respectively, whereas, the middle portion presents the theoretical mean images at 500,000 total projection counts for iteration numbers 10, 30, 50, 100 and 200 running through linear MAP-EM algorithm using $\alpha = 16, q = 15$, β = value of the phantom corresponding with their respective profiles. Here we checked the Monte-Carlo (MC) validation with 500,000 noise realizations, however, for each experiment, the results of mean are very consistent at iterations 10, 30, 50, 100 and 200 for both MC simulation and theoretical analysis. Figure 2 shows the excellent agreement of profiles of mean images through the lesion center at 500,000 projection counts at different iterations 10, 30, 50, 100 and 200 respectively for MC and theory.



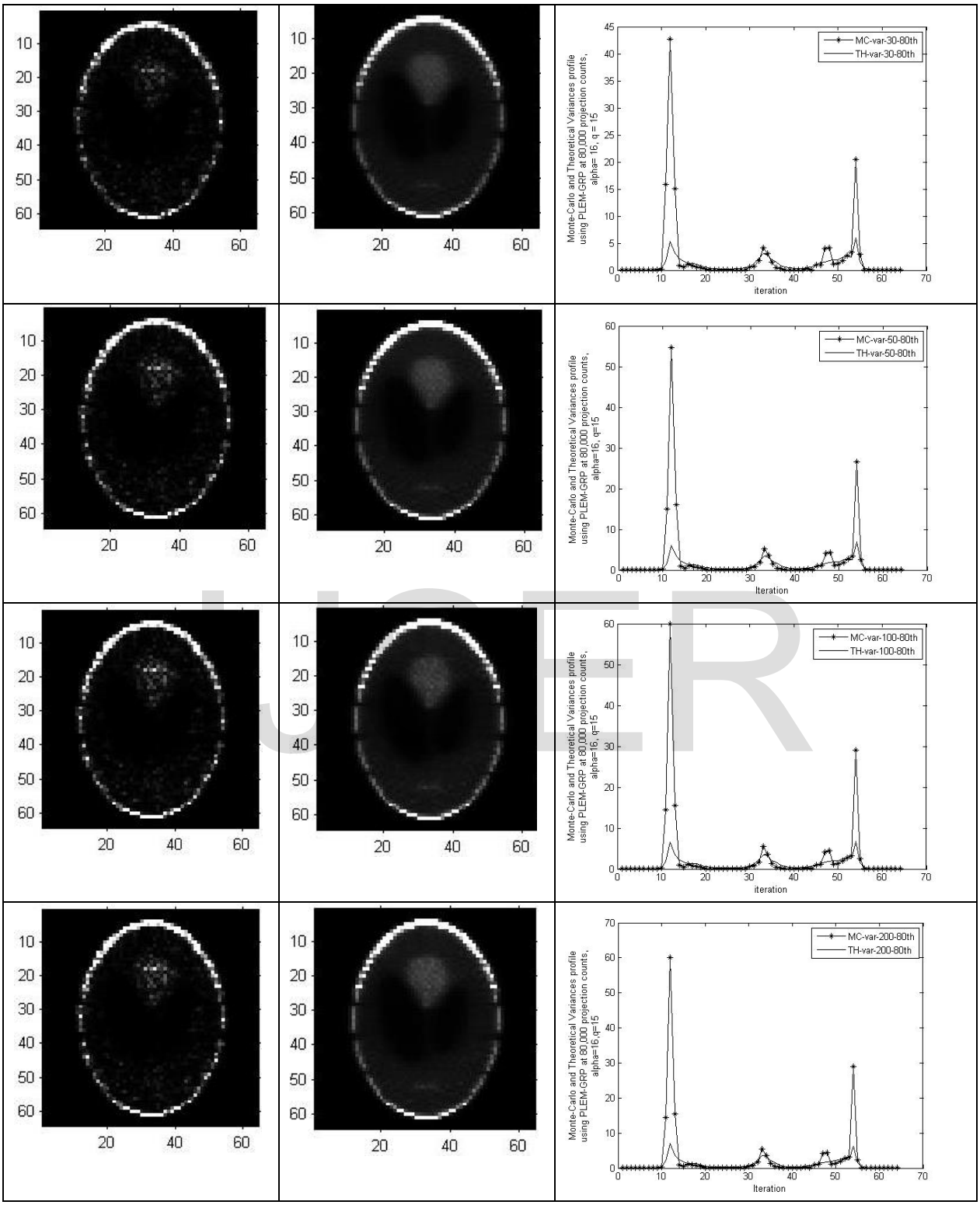
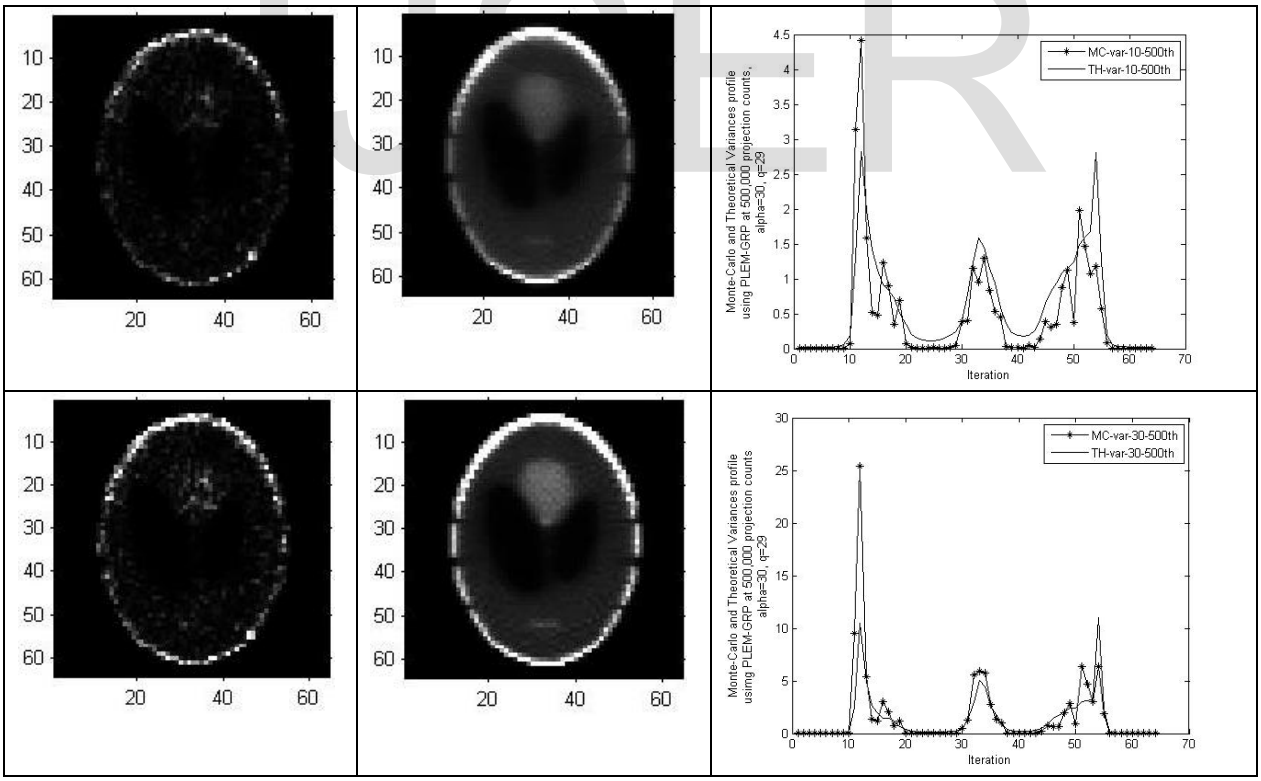


Figure 3: Average and theoretical variance corresponding to their profiles at each iteration (10, 30, 50, 100, and 200) using nonlinear PLEM and linear MAP-EM algorithm using GRP at 80,000 projection counts, α=16, q=15 and β =value of the phantom.

The "Monte Carlo" variance images of total 500 noisy images for iteration number equal to 10,30, 50, 100 and 200 are shown in figure (3) whereas, the total projection counts for this case

were 80,000 counts. The top left corner of the figure (3) shows the noisy mean images of 500 noisy images at 80,000 total projection counts reconstructed through" nonlinear PLEM" algorithm using "GRP" for iterations equal to 10, 30, 50, 100 and 200 respectively, whereas, the middle portion presents the theoretical variance images at 80,000 total projection counts for iteration numbers 10, 30, 50, 100 and 200 running through "linear MAP-EM" algorithm for "Gamma Root Prior (GRP)"using $\alpha = 16, q = 15$, β = value of the phantom corresponding with their respective profiles at through the image center. However, the variation of noise at the edges as shown in figure 3 shows the difference in non-linear PLEM and linear MAP-EM algorithms, as "PLEM algorithm" is nonlinear; hence the variation at the edges is very large as compared to the linear "MAP-EM" algorithm. Moreover, at lower iteration, the variance resembles like the phantom but as there will be the variation in iteration, there will be the increment in the noise behavior shown in figure. Through variance, the edge preserving results are explained gracefully through iterative methods, however, variance plots for this case as shown in figure 3 for 80,000 total projection counts.



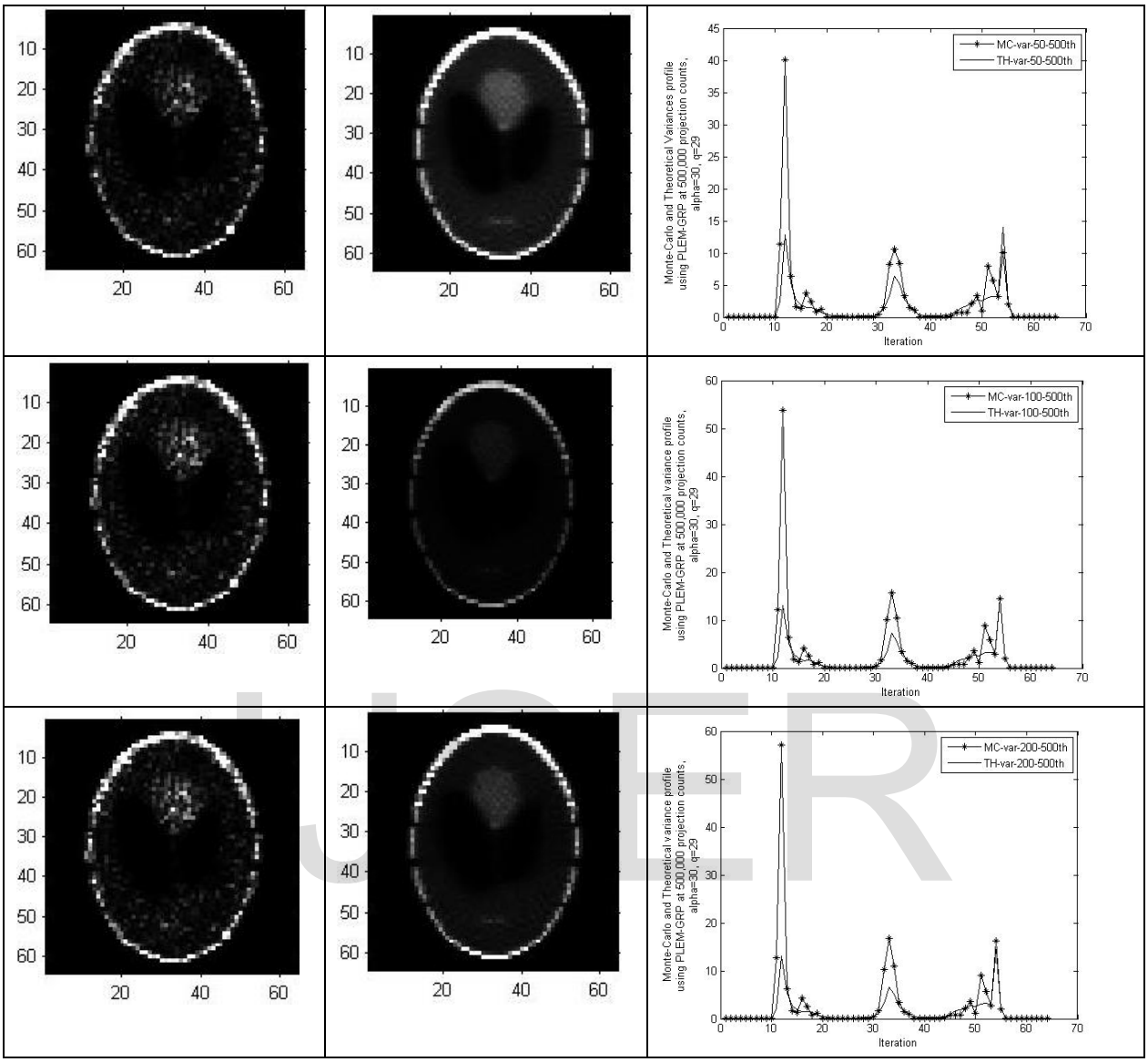


Figure 4: Average and theoretical variance corresponding to their profiles at each iteration (10, 30, 50, 100, and 200) using nonlinear PLEM and linear MAP-EM algorithm using GRP at 500,000 projection counts, α=30, q=29 and β=value of the phantom.

The "Monte Carlo" variance images of total 500 noisy images at iterations 10,30, 50, 100 and 200 are shown in figure (4) whereas, the projection counts are 500,000 counts. The top left corner of the figure (4) shows the mean images of 500 noisy images at 500,000 total projection counts reconstructed through "PLEMGRP" at each iteration, whereas, the middle portion presents the theoretical variance images at 500,000 projection counts for each iteration through MAP-EM-GRP using $\alpha = 30, q = 29$, β indicates the pixel weight corresponding with their respective profiles at through the image center. However, the variation of noise at the edges as shown in figure 4 shows the difference in "PLEM" and " MAP-EM" algorithms, as "PLEM algorithm is nonlinear; hence the variation at the edges is very large as compared to the linear

MAP-EM algorithm". Moreover, the lower iteration, the variance behaves like the phantom, whereas, variance in iteration comparatively increase of noise in the reconstruction. The variance image is smeared leading to the noiseless images as shown in the middle portion of figure 4. The increment in neighborhoods will smoothed the noise variations increase and finally smoothed out the large shown in figure 4. The variance is explained better through the linear algorithm rather than non-linear. The comparison of variance plots for this case as shown in figure 5 for 500,000 projection counts

Table 1: Monte-Carlo and Theoretical estimates of mean and variance at each iteration (10, 30, 50, 100, 200) using PLEM-GRP algorithm, α = 16, q=15, β = value of the phantom, at 80,000 projection counts.

iter	MC .mean	Th.mean	MC.variance	Th.variance
10	147.3827	147.5703	1.1506	0.9909
30	148.3736	148.5211	2.2866	2.0035
50	148.1938	149.0270	2.4968	2.1605
100	148.0590	148.7470	2.5577	2.1553
200	148.0353	149.8157	2.5596	2.2066

Table 2: Monte-Carlo and Theoretical estimates of mean and variance at each iteration (10, 30, 50, 100, 200) using PLEM-GRP algorithm, α = 30, q=29, β = value of the phantom, at 500,000 projection counts.

iter	MC .mean	Th.mean	MC.variance	Th.variance
10	296.4945	297.1547	0.6680	1.1481
30	314.1501	313.3849	3.3005	2.8158
50	316.6665	316.4774	5.1778	3.2299
100	316.9292	319.0601	6.9149	3.4288
200	316.7447	314.6408	7.2430	3.2421

The above Table 1 and Table 2 show typical results for our experiment done for "nonlinear PLEM and linear MAP-EM algorithms using Gamma Root Prior (GRP)" through 80,000 and 500,000 projection counts with different parameters respectively. Table 1, as shown above, gives the analysis of MC-theory experiment for low projection counts i.e. 80,000 total projection counts using "PLEM and MAP-EM algorithms using GRP" with parameters $\alpha = 16\ and\ \beta = value\ of\ the\ phantom$. However, as we can see from the tables, the "(MAP-EM)-GRP" performs well for both 80,000 (low projection) counts and 500,000 (high projection) counts. Table 2, as shown above, gives the analysis of MC-theory experiment for high projection counts i.e. 500,000 total projection counts using "MAP-EM" algorithm using GRP with parameters $\alpha = 30\ and\ \beta = value\ of\ the\ phantom$. It is seen from the above tables that on the average, the higher counts that exhibit high signal-noise ratio perform better and gives us the good results than low projection counts with low signal to noise ratio. However, the results that are shown in

Table 1 and Table 2 for MC-theory mean and variance are dependent at each iteration for "Monte-Carlo (MC)" simulations.

Conclusions and Discussion

The numerical assessment of theoretical statistical attributes of the noise in an expectation-maximization reconstruction is evaluated through theoretical formula. The theoretical average image was elaborated through the ideal image at the *kth* iteration. The iteration number increases linearly through the centroid of the elaboration. The large wavering might be vitiated that expansion instead the final reconstructed maximum likelihood image. When the iteration are greater enough, there will be low projection counts that would occur the errors.

However, the noise is largest, the projections of the reconstructed image look ideally but just for few iterations. The algebraic form of the equations is considered to be the simplification of the estimated object. The analysis gives the results for the mean and variance as a function of the target, iteration number and system matrix. The square of theoretical mean is comparable to the variance at the point that makes the noise more local.

The low noise is relative to the low intensity images. The variance dependency on mean is not quadratic. The trends of variance and covariance of noise and iteration are observed linearly and recapitulated through the theory. The theoretical noise analysis through "MAP-EM" recapitulate the application of two different noise levels on the "Shepp-Logan phantom". However, to compute the "covariance $K_{F|f}^{k}$", an important evaluation must be done through the theory.

The derivation of the theoretical formula is validated to apply the study of noise propagation through iterations in this thesis. The quadratic prior's plays an important role in reducing the noise in topical neighborhoods. The mean image that runs algorithm on the noise free data is traced by the theory. The imprecise solution is calculated through simulations. The forms for $A^k B^k \; and \; C^k$ without approximation would be disorder, U^k is logical, however, the unrealistic bias to be introduced for the breakdown.

References

1. Silverman B. W., Jennison C., Stander J., Brown T. C. The specification of edge penalties for regular and irregular pixel images. IEEE Trans PattAnaly Mach Intel. 1990; 12(10):1017-24.

2. Tang J., Zhang L., Chen Z., Xing Y., Cheng J. A BPF-type algorithm for CT with a curved PI detector.
3. Kearfott K. J., Votaw J. R. The Basics of Positron Emission Tomographic (PET) Imaging. In: Emran A., editor. Chemists Views of Imaging Centers.1995 Springer US.
4. Wang W and Gindi 1996a Noise analysis of regularized EM algorithms for SPECT Technical Report, Depts. Of Radiology and Electrical Engineering, State University of New York at Stony Brook MIPL-96-1.
5. W. Wang and G. Gindi, Noise analysis of MAP-EM algorithms for emission tomography: Validation and Task Performance Application to Quantitation, Technical Report MIPL-96-3, Depts. Of Radiology and Electrical Engineering, State University of New York at Stony Brook, Oct. 1996.
6. H. H. Barrett, D. W. Wilson, and B. M. W. Tsui, Noise Properties of the EM Algorithm: 1. Theory, Phys. Med. Biol., 39, pp.833-846, 1994.
7. Vardi, Y. Shepp, and L. A., (1982), "Maximum Likelihood Reconstruction for Emission Tomography," IEEE Transactions on Medical Imaging, 1, 113-122.
8. Shepp, L. A., Vardi Y, Kaufman L.," A statistical model for positron emission tomography. J Am Stat Assoc 1985; 80 (389): 8-20.
9. Mark Christiaens, Bjorn De Sutter, Koen De Bosschere, Jan Van Campenhout and Ignace Lemahieu (1999),"A fast, cache-aware algorithm for the calculation of radiological paths exploiting subword parallelism," Journal of Systems Architecture 45, 781-790.
10. Fessler J. A., Rogers W. L. Spatial resolution properties of penalized-likelihood image reconstruction methods: Space-invariant tomographs. IEEE Trans. Img. Proc. 1996;5(9):1346-58.
11. Fessler J. A., Rogers L. Uniform quadratic penalties cause non-uniform image resolution. IEEE Nucl. Sci. Symp. Med. Img. Conf.; 1994.: p. 1915-9. Biomed. Engg. Appl. Basis Comm. 2003;15:143-9.
12. Jinhua Sheng, Lei Ying (2005), "A fast image reconstruction algorithm based on penalized-likelihood estimate," Medical Engineering & Physics 27, 679-686.

13. Meltzer C. C., Kinahan P. E., Green P. J., icholas T. E., omtat C., antwell M. N., in M. P., rice J. C. Correction of PET data for partial volume effects in human cerebral cortex by MR imaging. J Comp. Assist. Tomo 1999;40(12):2053-6.