Author's personal copy

Ann Nucl Med
DOI 10.1007/s12149-012-0657-5

TECHNICAL NOTE

A heuristic statistical stopping rule for iterative reconstruction in emission tomography

F. Ben Bouallègue · J. F. Crouzet · D. Mariano-Goulart

Received: 2 May 2011 / Accepted: 19 September 2012
© The Japanese Society of Nuclear Medicine 2012

Abstract
Objective We propose a statistical stopping criterion for iterative reconstruction in emission tomography based on a heuristic statistical description of the reconstruction process.
Methods The method was assessed for MLEM reconstruction. Based on Monte-Carlo numerical simulations and using a perfectly modeled system matrix, our method was compared with classical iterative reconstruction followed by low-pass filtering in terms of Euclidian distance to the exact object, noise, and resolution. The stopping criterion was then evaluated with realistic PET data of a Hoffman brain phantom produced using the GATE platform for different count levels.
Results The numerical experiments showed that compared with the classical method, our technique yielded significant improvement of the noise-resolution tradeoff for a wide range of counting statistics compatible with routine clinical settings. When working with realistic data, the stopping rule allowed a qualitatively and quantitatively efficient determination of the optimal image.
Conclusions Our method appears to give a reliable estimation of the optimal stopping point for iterative reconstruction. It should thus be of practical interest as it produces images with similar or better quality than classical post-filtered iterative reconstruction with a mastered computation time.

F. Ben Bouallègue (✉) · D. Mariano-Goulart
Nuclear Medicine Department, Lapeyronie University Hospital, 371 avenue du Doyen Gaston Giraud, 34295 Montpellier Cedex 5, France
e-mail: faybenb@hotmail.com

J. F. Crouzet
Department of Mathematics, Montpellier Science University, 34095 Montpellier, France

Keywords Emission tomography · Iterative reconstruction · MLEM · Stopping criterion

Introduction

In the medical field, emission tomography is an imaging technique that seeks to estimate the distribution of a radioactive tracer confined inside the human body based on a record of the electromagnetic waves emitted by the tracer in a given set of directions. Tomographic reconstruction is an inverse problem that is known to have analytical solution under certain conditions. However, due to their ability to incorporate various corrections (geometrical system response, Compton scattering, and attenuation), it has now become common practice to exploit algebraic iterative methods for clinical applications. Many original iterative algorithms have been proposed since the 1970s in the frame of emission tomography, among which the algebraic reconstruction technique ART [1, 2], the iterative least-square technique ILST [3], the simultaneous iterative reconstruction technique SIRT [4], the steepest descent method STP [5], the iterative filtered back-projection [5], and the simultaneous multiplicative algebraic reconstruction technique SMART [6, 7]. In 1977, Dempster et al. [8] proposed to rely on the statistical properties of the tomographic record when defining the functional that has to be minimized throughout the reconstruction process. Based on the Poisson nature of the record [9], they derived the maximum likelihood expectation maximization (MLEM) technique. MLEM is known to converge toward the image that is the most likely to have produced the record [10]. Rescaled block-iterative (RBI) methods have been designed in order to fasten the convergence of MLEM [11, 12], among which the ordered subset expectation maximization OSEM [13] and the

Published online: 04 October 2012

Author's personal copy

row-action maximum likelihood algorithm RAMLA [14]. The power of statistical iterative methods as well as their advantage over other reconstruction methods has been well established [15–17]. However, due to the ill-posedness of the tomographic problem and to the inconsistency of the record, the noise in the reconstructed image usually increases as iterations of the algorithm proceed. This noise is composed of a statistical high-frequency noise, which depends on the noise in the record, and a deterministic inaccuracy, which depends on convergence properties of the algorithm, truncation, round-off errors, and imperfection of the model [18, 19]. Consequently, in clinical routine, most users impose the termination of the reconstruction process after only a few iterations [20] in order to limit the noise propagation, which remains a speculative method that leads to uncertain results [21, 22]. In order to prevent noise propagation, regularized methods have been proposed that consist in adding some a priori knowledge regarding the smoothness of the studied object. In maximum a posteriori (MAP) algorithms, the prior is usually expressed using a Gibbs energy function penalizing the likelihood [23–25]. Another classical regularization technique involves a low-pass filtering of the estimated object, either after the iterative process or in between each iteration step [26]. Regularized methods efficiently stem the noise spreading but inevitably lead to substantial loss of resolution. Moreover, they generally require the empirical tuning of some scalar parameters adjusting the desired amount of smoothing. In this paper, we rely on a heuristic description of the iterative reconstruction process to define a stopping criterion based on the statistical properties linking the exact object with the record. We first show through numerical Monte-Carlo experiments that our method provides resulting estimates that are almost as good as the better estimate reached during the reconstruction in terms of Euclidian distance. We then assess our method versus a classical compromise consisting in applying a huge number of iterations followed by a low-pass filtering of the resulting image. The comparison is performed using numerical simulations and involves root mean square (RMS) error and noise-resolution tradeoff assessment. We finally evaluate our method with realistic PET data of a Hoffman brain phantom produced using the GATE platform.

Materials and methods

Context and notations

Emission tomography consists in studying an object $\overline{\boldsymbol{\theta}}$ corresponding to a radioactive distribution confined into the field of view of a tomographic device. Numerically, this object can be considered as a function defined over a Cartesian grid composed of N pixels, i.e. $\overline{\boldsymbol{\theta}} \in \mathrm{IR}^N$. As the direct point-wise measurement of the activity is impossible, one relies on the tomographic process to produce an estimate $\boldsymbol{\theta}$ of the exact object $\overline{\boldsymbol{\theta}}$. This process can be decomposed into two steps. First, the tomographic recording consists in counting the number of emitted photons for a given set of detection bins. The record is denoted $\mathbf{p} \in \mathrm{IR}^M$ with M, the total number of bins. Let us name $\overline{\mathbf{p}} \in \mathrm{IR}^M$ the integrals of the exact activity over the lines of response linked with the detection bins. The integrals $\overline{\mathbf{p}}$ can be expressed as $\overline{\mathbf{p}} = \Lambda\overline{\boldsymbol{\theta}}$ with $\Lambda \in \mathrm{IR}^{M\times N}$ the Radon operator. Due to the physical properties of the radioactive decay, we have that $\mathbf{p} = \phi(\overline{\mathbf{p}})$ where ϕ stands for the Poisson noise operator. The Poisson noise is known to have a dispersion characterized by $E\left[(\mathbf{p}-\overline{\mathbf{p}})^2\right] = \overline{\mathbf{p}}$. Second, the tomographic reconstruction consists in producing the estimate $\boldsymbol{\theta}$ based on the record $\mathbf{p}$. The iterative reconstruction techniques use a step by step error correction procedure in order to find a solution $\boldsymbol{\theta}$ such that $\boldsymbol{\theta}$ is positive and $\Lambda\boldsymbol{\theta} = \mathbf{p}$. Some additional regularity constraints may be imposed on $\boldsymbol{\theta}$ based on a priori information concerning the studied object $\overline{\boldsymbol{\theta}}$. The MLEM algorithm seeks the estimate $\boldsymbol{\theta}$ as the most likely to produce the record $\mathbf{p}$ through the following procedure (where n denotes the iteration index):

$$\boldsymbol{\theta}_j^0 = cst\ \forall j; \quad \boldsymbol{\theta}_j^{n+1} = \boldsymbol{\theta}_j^n \frac{\sum_i \Lambda_{ij} \frac{\mathbf{p}_i}{(\Lambda\boldsymbol{\theta})_i}}{\sum_i \Lambda_{ij}} \tag{1}$$

Heuristic approach—definition of the stopping rule

As the Poisson noise affecting the records is known to be additive, one can decompose $\mathbf{p}$ into two terms:

$$\mathbf{p} = \overline{\mathbf{p}} + \boldsymbol{\beta} \tag{2}$$

where $\overline{\mathbf{p}}$ can be seen as the useful signal and where $\boldsymbol{\beta}$ is the noise term which order of magnitude is $O(\sqrt{\overline{\mathbf{p}}})$. During the reconstruction process, the algorithm reconstructs both terms and we hence can consider that, each iteration, the resulting image is also composed of a useful image and a noise image as:

$$\boldsymbol{\theta}^n = \boldsymbol{\mu}^n + \boldsymbol{\varepsilon}^n \tag{3}$$

with $\boldsymbol{\mu}$ the useful image and $\boldsymbol{\varepsilon}$ the noise image. The MLEM algorithm is designed to converge toward a solution denoted $\widetilde{\boldsymbol{\theta}}$ such that:

$$\boldsymbol{\theta}^n \to \widetilde{\boldsymbol{\theta}};\ \Lambda\widetilde{\boldsymbol{\theta}} = \mathbf{p} \tag{4}$$

This solution may as well be decomposed as:

$$\widetilde{\boldsymbol{\theta}} = \widetilde{\boldsymbol{\mu}} + \widetilde{\boldsymbol{\varepsilon}};\ \Lambda\widetilde{\boldsymbol{\mu}} = \overline{\mathbf{p}};\ \Lambda\widetilde{\boldsymbol{\varepsilon}} = \boldsymbol{\beta} \tag{5}$$

Author's personal copy

where one can legitimately choose $\widetilde{\boldsymbol{\mu}} = \overline{\boldsymbol{\theta}}$. At the first iteration, as the estimate is initialized as a constant, no noise has yet been reconstructed and one can write:

$$\boldsymbol{\theta}^0 = \boldsymbol{\mu}^0;\ \boldsymbol{\varepsilon}^0 = 0 \tag{6}$$

We now define two indicators that will help us interpret the evolution of the reconstruction process throughout the iterations. The first one is an objective indicator which is unknown in practice:

$$\mathrm{RMS}^n = \sqrt{E\left(\left[\boldsymbol{\theta}^n - \overline{\boldsymbol{\theta}}\right]^2\right)} \tag{7}$$

i.e. the root mean square error between the reconstructed image and the exact object. The behavior of this first indicator has already been studied. RMS^n is known to decrease until it reaches its minimum and then to increase [18, 19]. The reconstructed image for which RMS^n reaches its minimum is denoted $\widehat{\boldsymbol{\theta}}$ and called ideal estimation. The second indicator is empirical but may be computed in practice each iteration:

$$J^n = \frac{\sum_i \left(\mathbf{p}_i - \Lambda\boldsymbol{\theta}_i^n\right)^2}{\sum_i \Lambda\boldsymbol{\theta}_i^n} \tag{8}$$

(where $\Lambda\boldsymbol{\theta}_i = \sum_j \Lambda_{ij}\boldsymbol{\theta}_j$), i.e. the mean of the quadratic error between the record and the image integrals normalized with the mean of the image integrals. When computed using the exact object, this indicator yields due to the dispersion properties of the Poisson noise:

$$\overline{J} = \frac{\sum_i (\mathbf{p}_i - \overline{\mathbf{p}}_i)^2}{\sum_i \overline{\mathbf{p}}_i} \approx 1 \tag{9}$$

At the first iteration, as the constant initial image may be chosen as small as one wants, we have that $\boldsymbol{\theta}^0 = \boldsymbol{\mu}^0 \ll 1$ so that $\Lambda\boldsymbol{\theta}^0 \ll \mathbf{p}$ and thus $J^0 > 1$. When $n \to \infty$, the reconstruction algorithm seeks to build an object whose projections fit with the recorded data. The indicator J^n will hence decrease monotonically and ideally tend toward 0 (if an image $\widetilde{\boldsymbol{\theta}}$ such that $\Lambda\widetilde{\boldsymbol{\theta}} = \mathbf{p}$ is found), i.e.:

$$J^n \to \tilde{J} = \frac{\sum_i \left(\mathbf{p}_i - \Lambda\widetilde{\boldsymbol{\theta}}_i^n\right)^2}{\sum_i \Lambda\boldsymbol{\theta}_i^n} = 0 \tag{10}$$

Consequently, during the reconstruction, J^n starts from $J^0 > 1$ and decreases toward 0 hence having at one moment a value of 1. The reconstructed image for which $J^n = 1$ is denoted $\breve{\boldsymbol{\theta}}$. By virtue of Poisson's law describing the statistical uncertainty in the tomographic record, we will call (physically) realistic an image for which the indicator J is close to 1. One can also consider that the more J is close to 1, the more the related image is realistic. In that sense, $\breve{\boldsymbol{\theta}}$ is the most realistic image reached by the algorithm during the reconstruction. The ensemble of the realistic images gathers the images for which the record obeys a Poisson statistics around the image integrals. This ensemble obviously contains the exact object $\overline{\boldsymbol{\theta}}$. Figure 1 represents schematically the reconstruction process. Figure 1a displays in two dimensions the evolution of the estimated image with the useful image in abscissa and the noise image in ordinate. Each one-dimensional axis stands for a hyperspace of dimension N. The first iterations contribute to the reconstruction of the useful image without reconstruction of noise (zone A) until the estimation reaches a realistic image for which $J = 1$. From there, the algorithm starts to reconstruct the noise (zone B) and passed a certain point the algorithm only reconstructs the noise (zone C). Figure 1b shows the evolution of the (RMS, J) doublet along with the iterations. The key images of the reconstruction are displayed as well as the exact object. Zones A and B are delimited by the line $J = 1$. The border between zone B and zone C is indicated arbitrarily ($J = \alpha$). We denote $\widehat{J}$ the value of J when RMS reaches its minimum, i.e. the value of J associated with the image $\widehat{\boldsymbol{\theta}}$ that is the closest to $\overline{\boldsymbol{\theta}}$ in terms of Euclidian distance. As the ideal estimation $\widehat{\boldsymbol{\theta}}$ is reached inside zone B, we know that $\widehat{J} \in [\alpha 1]$ with α an unknown strictly positive parameter. In the next section, we show experimentally that, when working with a perfectly defined system matrix, the value of $\widehat{J}$ is highly stable and very close to 1 whatever the studied object and the count rate. If we denote J_{opt} the mean value of $\widehat{J}$, the condition $J \leq J_{\mathrm{opt}} \approx 1$ can thus be seen as a natural stopping rule for iterative reconstruction.

Assessment of the method through Monte-Carlo simulations

The numerically simulated tomographic records were constituted of 64 projection angles over 180° and 64 projection bin per angle. The projections were parallel. The images were reconstructed over a 64 × 64 Cartesian grid. When simulating the records and reconstructing, we did not take into account the attenuation or the scatter, and we assumed a perfect detector response. We considered a uniform pixel model to build up the system matrix Λ. The Poisson noise was simulated using Knuth's algorithm [27].

Validation of the stopping rule

We generated 500 exact objects constituted of a central uniform disk of radius 25 pixels and random activity between 0 and 2, plus several (1–5) additional disks of random radius (2–10 pixels) activity (0–10) and position (Fig. 2 shows a possible configuration). For each object, a tomographic record

Author's personal copy

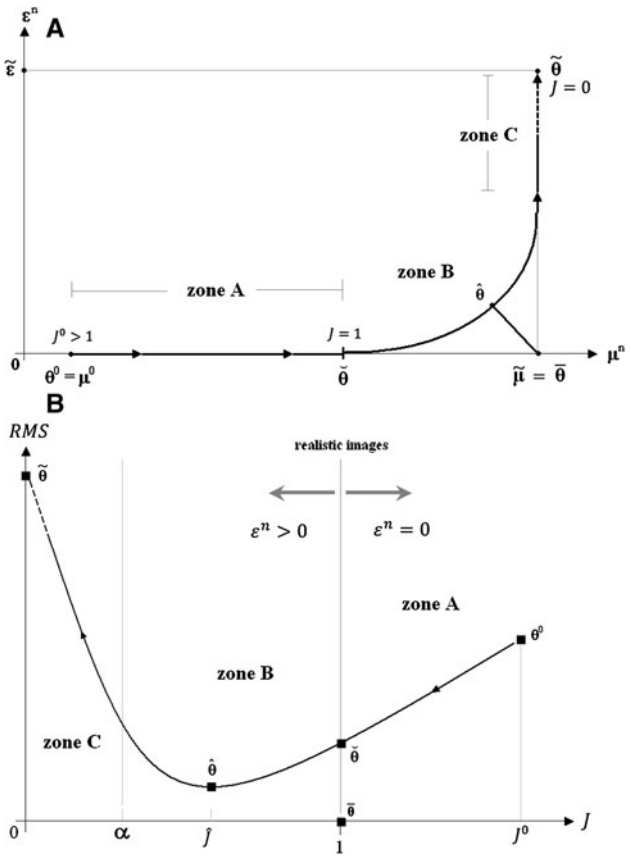

Fig. 1 Schematic representation of the reconstruction process. **a** Evolution of the reconstructed image in terms of useful and noise images. **b** Evolution of the two defined indicators *J* and *RMS*

was simulated with a random counting statistics comprised roughly between 5,000 and 140,000 counts. This record was reconstructed using MLEM and, each iteration, RMS^n and J^n were computed. After 100 iterations (which proved to be always sufficient), $\widehat{J}$ was identified as the value of J^n corresponding to the minimum of RMS^n.

Comparison with a classical methodology in terms of RMS

We then assessed our method (MLEM stopped when $J \leq J_{opt} \approx 1$) called in the sequel MLEM-STOP with comparison to the standard method consisting in applying MLEM with a high number of iterations (here 100) followed by a convolution of the resulting estimate with a Gaussian kernel of full-width at half-maximum (FWHM) the resolution of the recording device (1 pixel in our configuration), called in the sequel MLEM-CONV. We reissued the experiment carried out in the last section: 500 records corresponding to various exact objects were generated with counting statistics ranging from 5,000 to 140,000 total counts. Each record was reconstructed and three indicators were computed:

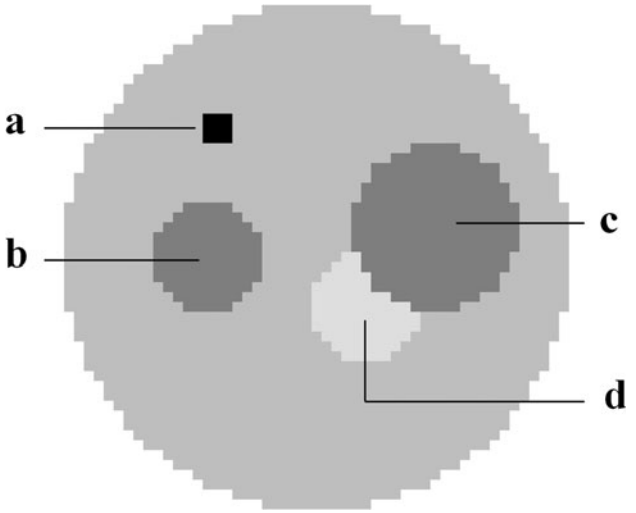

Fig. 2 One possible configuration of the first studied object. *Background* radius 25 pixels and relative uptake 1. *Area a* radius 1.5 pixels and relative uptake 4. *Area b* radius 5.5 pixels and relative uptake 2. *Area c* radius 8.5 pixels and relative uptake 2. *Area d* radius 5.5 pixels and relative uptake 0.5

- $RMS_{MLEM\text{-}STOP}$ the root mean square error between the estimate produced by MLEM-STOP and the exact object.
- $RMS_{MLEM\text{-}CONV}$ the root mean square error between the estimate produced by MLEM-CONV and the exact object.
- RMS_{MIN} the minimum root mean square error between the reconstructed image and the exact object reached within 100 iterations of MLEM.

Comparison with a classical methodology in terms of resolution and noise

The study focused on an object constituted of a uniform background disk (Area 0) of radius 24 pixels and activity 1 (Fig. 3). Inside were disposed three uniform small disks (Areas 1, 2 and 3) of radius 4, 2 and 1 pixels having a relative uptake of 10 with respect to Area 0. We defined around each small disk an annular neighborhood (belonging to Area 0) of radius the double of that of the corresponding area. We generated tomographic records of this object for 15 levels of counting statistics ranging from 10,000 to 300,000 counts (this range of count rates is in agreement with routine clinical settings: in our department of nuclear medicine, the count rates in routine practice range from about 10,000 events per slice for gated myocardial SPECT to about 150,000 events per slice for MIBG SPECT. For thorax–abdomen–pelvis PET, the count rate per 2D slice ranges roughly from 200,000 to 500,000). For each of the 15 statistic levels, 50 records were generated

Author's personal copy

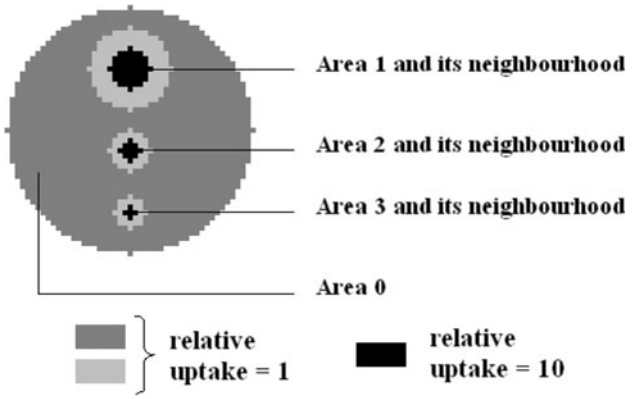

Fig. 3 Composition of the second studied object

and reconstructed using MLEM-STOP and MLEM-CONV. We studied two figures of merit:

- Noise: relative standard deviation of the reconstructed pixel values inside Area 0 (in %).
- Resolution: relative mean uptake of Areas 1, 2, and 3 with respect to the mean uptake of Area 0 and with respect to the mean uptake of their own neighborhood. This relative mean uptake was expected to be as close as possible to its exact value of 10 but was usually lower due to partial volume effects affecting the reconstructed images.

Evaluation of the method using realistic data

In order to evaluate the qualities of our stopping criterion with realistic data, we studied a Hoffman brain phantom containing three hot spots. Figure 4 shows the 2D activity and density maps and Table 1 details the relative uptakes and attenuation coefficients of the different tissues. The projection data were produced using the GATE (Geant4 Application in Emission Tomography) platform [28]. The numerical Hoffman phantom employed was an axially invariant 3D phantom whose 2D activity and density maps were similar to those displayed on Fig. 4 and defined on a 128 × 128 grid with a pixel size of 2 mm. The data were simulated using the specifications of the GEMINI GXL PET scanner (Philips Healthcare) [29]. Positron range and non-collinearity were not modeled. The scatter and random annihilations were corrected using the scatter and random estimations provided by the software [30, 31]. The system matrix employed for the reconstructions was built using a uniformly distributed pixel activity model and corrected for attenuation and detector sensitivity using the appropriate correction factors provided by the software. The images were reconstructed on a 128 × 128 grid with a pixel size of 2 mm. Four 3D date sets were produced with respective total activities of 30, 50, 100, and 150 MBq (^{18}FDG) and an acquisition time of 60 s. The

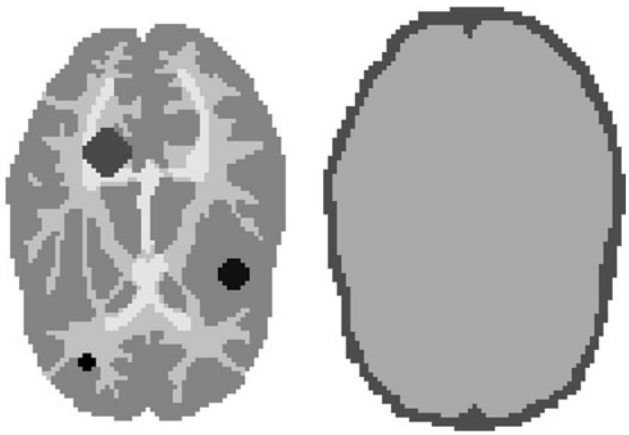

Fig. 4 Hoffman brain phantom used for the GATE PET simulations. *Left* activity map and *right* density map

Table 1 Hoffman phantom composition

Tissue	Relative uptake	Attenuation (cm^{-1})
Bone	0	0.15
Gray matter	1	0.1
White matter	0.5	0.1
Cerebrospinal fluid	0.2	0.1
Large tumor	1.5	0.1
Medium tumor	2	0.1
Small tumor	3	0.1

mean count levels in the 2D sinograms were, respectively, of 300,000, 500,000, 1 million, and 1.5 million. For each of the four data sets, the 25 (among 29) central transaxial 2D sinograms were reconstructed using MLEM and stopped using the described criterion. Furthermore, at each iteration step, three figures of merit were computed over the 25 reconstructed images: the bias, the standard deviation, and the RMS error. These three figures of merit were evaluated over the entire image and normalized (i.e. expressed in percent of the total image activity). The computation of the RMS error allowed to identify the value of $\widehat{J}$ corresponding to the minimal RMS error (RMS_{MIN}). The relative RMS error increase was computed as the relative difference between the RMS error of the image produced by the stopping rule and RMS_{MIN}.

Results

Assessment of the method through Monte-Carlo simulations

Figure 5 shows the histogram and the dispersion with the count rate of $\widehat{J}$ for the 500 studied objects. The distribution

Author's personal copy

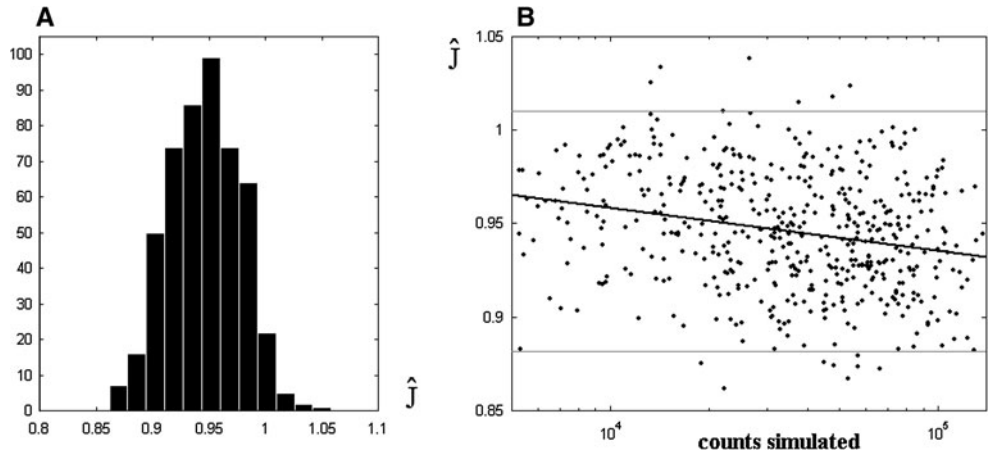

Fig. 5 **a** Histogram of the values of $\hat{J}$ for the 500 studied objects. **b** Dispersion of the values of $\hat{J}$ with the count rate (*black line* linear regression, *gray lines* 95 % confidence interval)

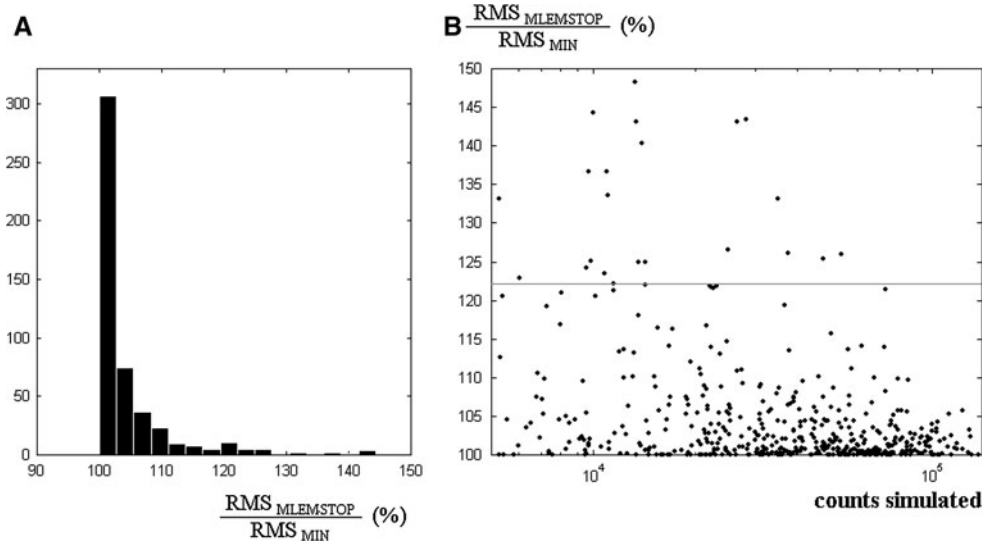

Fig. 6 Comparison between $RMS_{MLEM\text{-}STOP}$ and RMS_{MIN}. **a** Histogram of the ratio. **b** Evolution of the ratio with the count rate (*gray line* 95 % confidence interval)

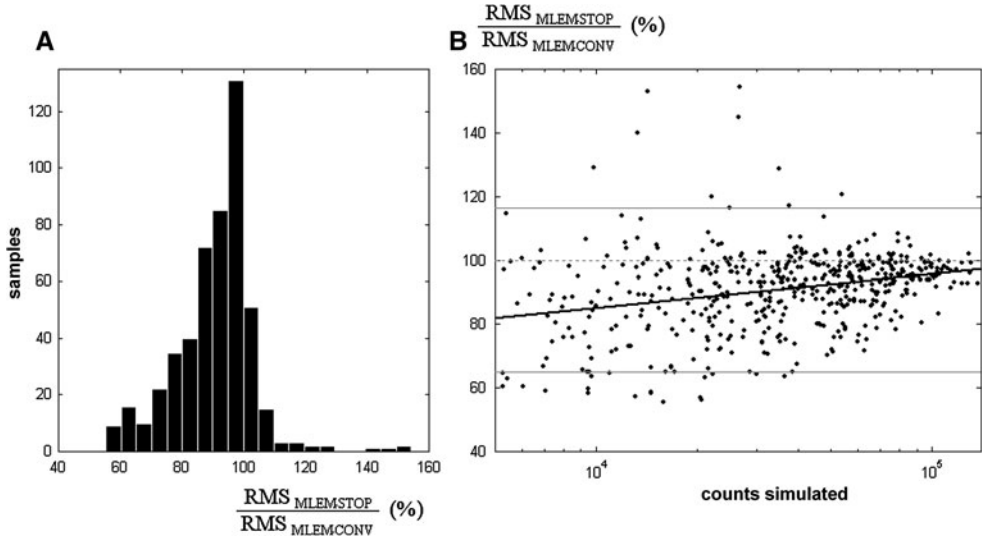

Fig. 7 Comparison between $RMS_{MLEM\text{-}STOP}$ and $RMS_{MLEM\text{-}CONV}$. **a** Histogram of the ratio. **b** Evolution of the ratio with the count rate (*black line* linear regression, *gray lines* 95 % confidence interval)

Author's personal copy

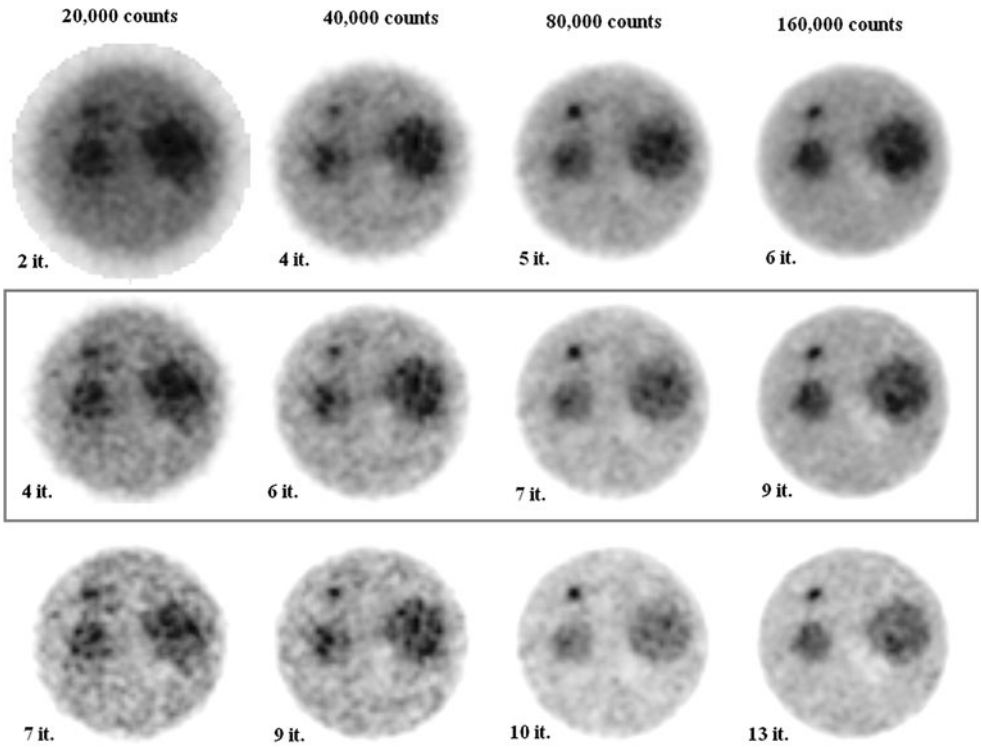

Fig. 8 Representative reconstructions provided by MLEM-STOP for the object described on Fig. 2 and for four count rates. *From left to right* 20,000, 40,000, 80,000 and 140,000 counts simulated. *Middle row* images obtained using MLEM-STOP. *Top and bottom rows* images obtained for a few less (*top*) or more (*bottom*) iterations during the reconstruction process. The number of iterations is mentioned under each slice

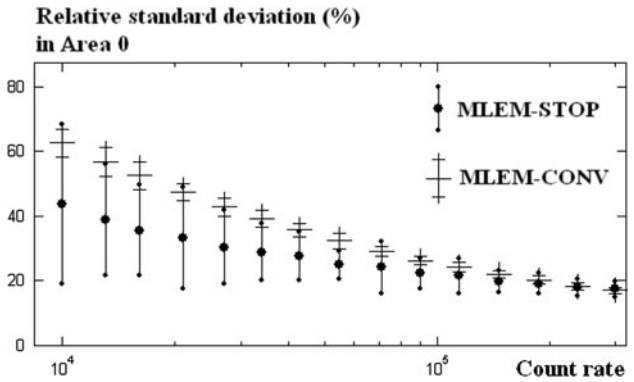

Fig. 9 Evolution of the noise figure of merit (mean and standard deviation) along with the counting statistics for MLEM-STOP and MLEM-CONV

of $\widehat{J}$ has a mean value of 0.946 and a standard deviation of 0.032, these parameters appearing to stay stable whatever the counting statistics. The 95 % confidence interval over the whole statistics range is [0.88 1.01]. The mean value being quite close to 1 and the relative standard deviation being very low (3.4 %), one can reasonably use the condition $J \leq 1$ as a stopping rule for MLEM reconstruction.

Figure 6 shows the histogram and the dispersion with the count rate of the ratio (in %) between $RMS_{MLEM\text{-}STOP}$ and RMS_{MIN}. The mean of the ratio is 105 % and the 95 % confidence interval is [100 % 122 %]. Figure 7 shows the histogram and the dispersion with the count rate of the ratio (in %) between $RMS_{MLEM\text{-}STOP}$ and $RMS_{MLEM\text{-}CONV}$. The mean of the ratio is 90.7 % and its standard deviation 12.8 %. Figure 8 (middle row) shows representative reconstruction provided by MLEM-STOP for the object described in Fig. 2 and for four count rates (20,000, 40,000, 80,000, and 140,000 counts). For comparison, the top and bottom rows provide the images obtained for a few less (top) or more (bottom) iterations.

Figure 9 displays the evolution of the noise figure of merit along with the statistics level. For each level and for each of the two tested methods, we show the mean and standard deviation of the figure of merit. Figure 10 displays the evolution of the resolution figure of merit along with the statistics level for each of the three studied areas. Here again are shown the mean and standard deviation of the figure of merit. Figure 11 displays representative reconstructions obtained with MLEM-STOP (top row) and MLEM-CONV (bottom row) for six count rates (10,000, 20,000, 35,000, 75,000, 150,000, and 300,000 counts).

Evaluation of the method using realistic data

Figure 12 shows the evolution with the iterations of the mean J value, the bias, the standard deviation, and the RMS error. On each subplot, the abscissa corresponds to the J values which decrease with the iterations, and the ordinate corresponds to one of the three figures of merit (bias, standard deviation and RMS error). Each round marker stands for one iteration step. For the four simulated count rates, $\widehat{J}$ was very stable and close to 1. The mean

Author's personal copy

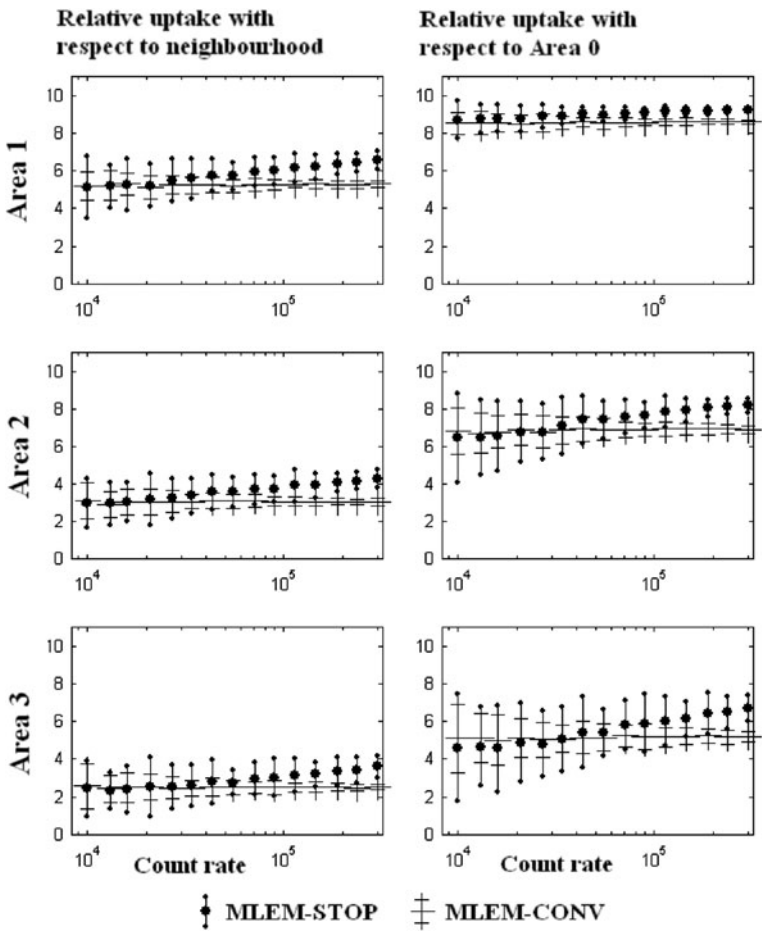

Fig. 10 Evolution of the resolution figure of merit (mean and standard deviation) along with the counting statistics for MLEM-STOP and MLEM-CONV

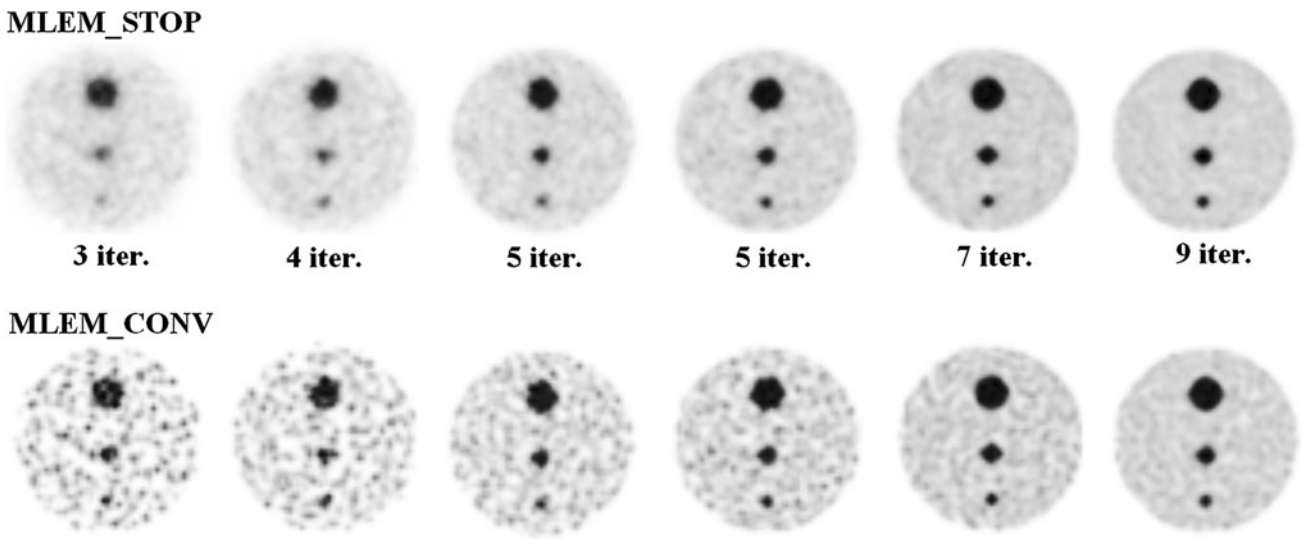

Fig. 11 Representative reconstructions obtained with MLEM-STOP (*top*) and MLEM-CONV (*bottom*). *From left to right* 10,000, 20,000, 35,000, 75,000, 150,000, and 300,000 counts simulated. For MLEM_STOP, the number of iterations after which the stopping criterion was reached is mentioned under each slice

value of $\widehat{J}$ over the 25 sinograms was 0.971 ($CI_{95\ \%} = [0.948\ 0.995]$) for the 300,000 count sinograms, 0.977 ($CI_{95\ \%} = [0.955\ 0.999]$) for the 500,000 count sinograms, 0.998 ($CI_{95\ \%} = [0.989\ 1.007]$) for the 1 million count sinograms, and 1.001 ($CI_{95\ \%} = [1.000\ 1.002]$) for the 1.5 million count sinograms. The relative RMS

Author's personal copy

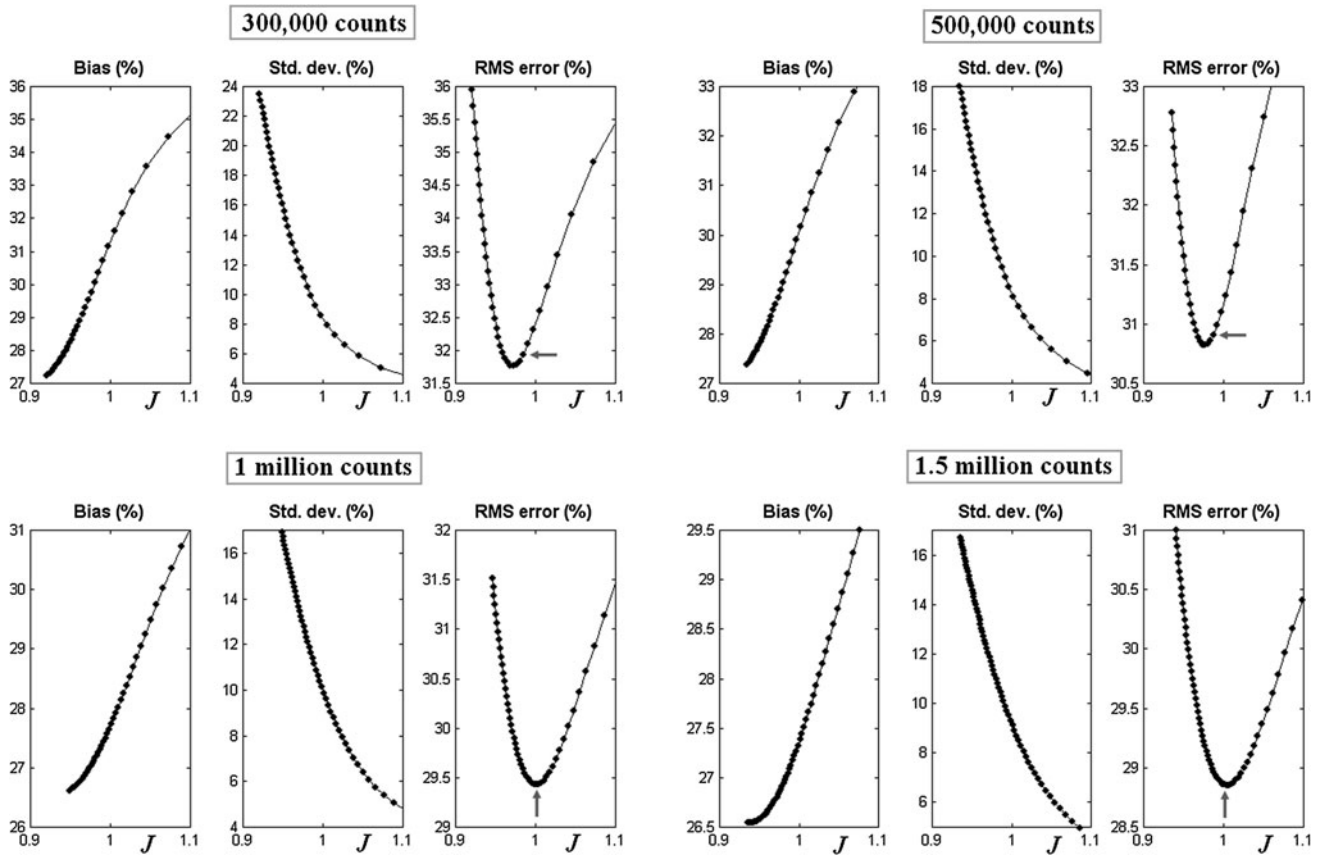

Fig. 12 GATE Hoffman phantom study. *From top left to bottom right* are shown the results corresponding to 300,000 counts, 500,000 counts, 1 million counts, and 1.5 million counts. For each count rate are displayed the evolution with J of (from *left to right*) the bias, the standard deviation, and the RMS error. Each *round marker* corresponds to an iteration step. The *gray arrows* indicate the iteration step for which the stopping rule is fulfilled

error increase was always low. Its {mean; maximum} values were, respectively, {1.9 %; 4.5 %} for the 300,000 count sinograms, {1.1 %; 3.2 %} for the 500,000 count sinograms, {0.06 %; 0.3 %} for the 1 million count sinograms, and {0.1 %; 0.2 %} for the 1.5 million count sinograms. Figure 13 shows the reconstructed slices of the GATE Hoffman phantom for the four studied count levels (one sinogram chosen arbitrarily over the 25 available for each count level). The middle framed column corresponds to the optimal image as defined by our stopping rule. For comparison, the other columns display the images obtained for a few less or more iterations.

Discussion

Assessment of the method through Monte-Carlo simulations

We have first validated the value of the indicator J_{opt} that conditions our stopping rule for MLEM. The results shown on Fig. 5 attest that the values of $\widehat{J}$ are highly localized, the majority being comprised inside the interval [0.9 1] with a mean of 0.946 and a standard deviation of 0.032. This localization is weakly dependent on the data statistics as appreciable from the regression line on Fig. 5b: the linear regression remains confined between 0.935 and 0.965 while the counting statistics vary from 5,000 to 140,000 counts. It is therefore well founded to rely on the condition $J \leq J_{\text{opt}} = E\left(\widehat{J}\right) \approx 1$ as a stopping rule for MLEM reconstruction whatever the studied object and the total counts in the record.

We have then assessed our method MLEM-STOP in terms of RMS, noise and resolution versus a classical regularization approach. Figure 6 shows that terminating the MLEM algorithm using our stopping rule leads to a resulting RMS error in the reconstructed image that is in average 5 % higher than the minimal RMS error reached during the reconstruction process. This supplementary RMS error is below 1.7 % for 50 % of the tested objects and below 22 % for 95 % of them. We also compared in terms of RMS error the results of MLEM-STOP with those of MLEM-CONV (100 iterations of MLEM followed by suitable convolution). The numerical experiment showed that MLEM-STOP yields an RMS error in the reconstructed

Author's personal copy

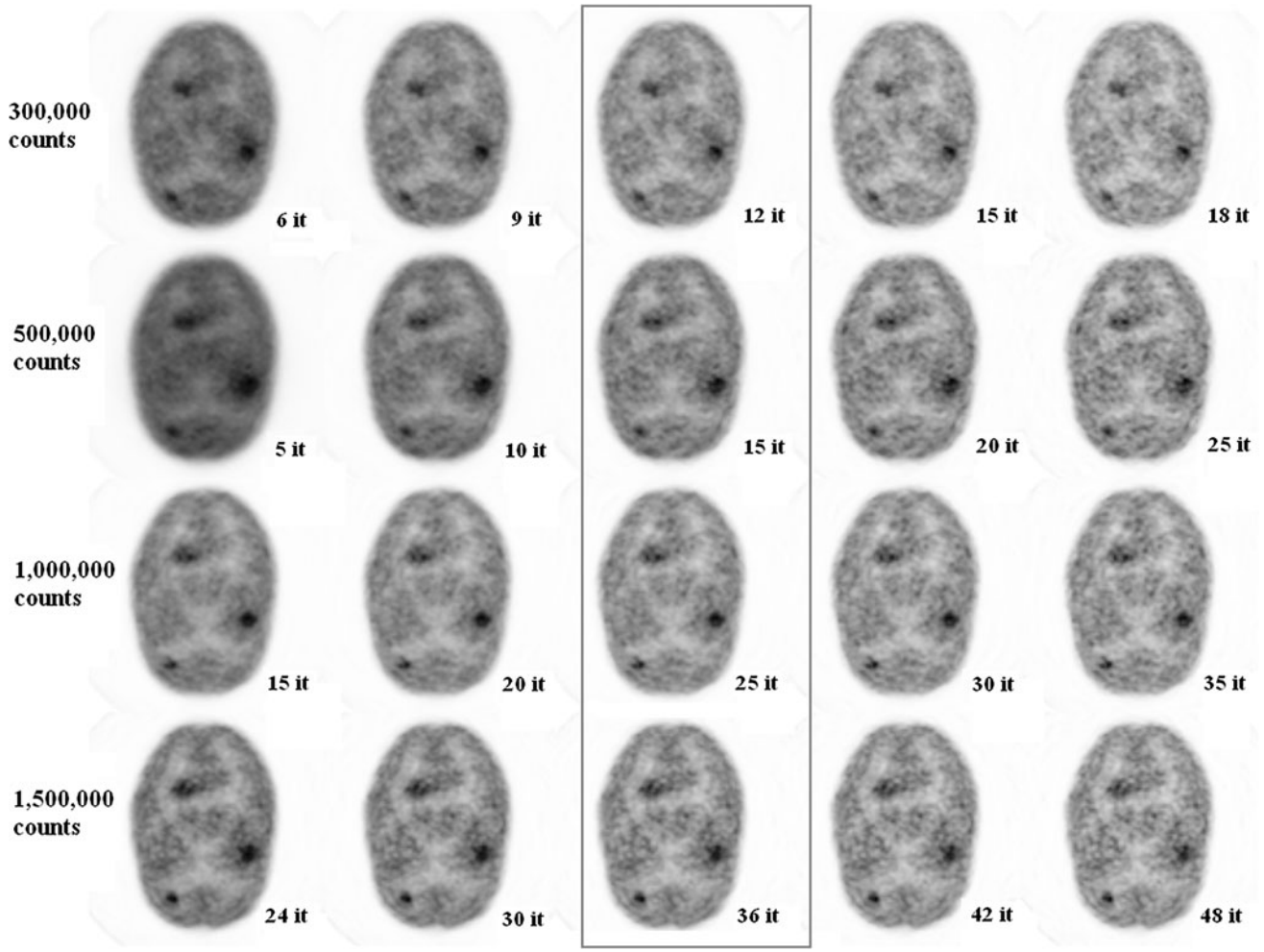

Fig. 13 Reconstructions of the GATE Hoffman phantom. *From top to bottom* mean count rate of 300,000, 500,000, 1 million, and 1.5 million counts. *Middle column* images obtained using our stopping rule. *Left and right columns* images obtained for a few less (*left*) or more (*right*) iterations during the reconstruction process. The number of iterations is mentioned under each slice

Table 2 Qualitative appreciation of the advantage of MLEM-STOP over MLEM-CONV depending on the counting statistics range, in terms of noise, resolution and noise-resolution tradeoff (0 stands for an equality and + for an advantage)

Count rate	10,000	40,000	100,000	300,000
Noise	++	+		0
Resolution	0	+		++
Tradeoff	++	++		++

images that is in average 9 % less than that obtained with MLEM-CONV (Fig. 7). For 95 % of the tested objects, $RMS_{MLEM\text{-}STOP}$ is comprised between 65 and 116 % of $RMS_{MLEM\text{-}CONV}$. The linear regression displayed on Fig. 7b attests that whatever the statistics (from 5,000 to 140,000) the image reconstructed with MLEM-STOP has an average RMS error lower than the average RMS error of the images reconstructed using MLEM-CONV (82 % for 5,000 counts to 97 % for 140,000 counts).

We finally compared MLEM-STOP with MLEM-CONV in terms of noise and resolution through numerical simulations using an object constituted of a uniform background plus three small hot disks. Figure 9 shows that MLEM-STOP produces images whose noise is in average

Author's personal copy

always lower than the noise in the images resulting from MLEM-CONV. However, as appreciable on the figure, due to the dispersion of the noise figure of merit values, the advantage of MLEM-STOP over MLEM-CONV becomes clear starting from a 5,000 count rate. Beyond 200,000 counts, the two methods tend to provide results that are no longer significantly different. Concerning the resolution, Fig. 10 shows that, as for the noise, the dispersion of the figure of merit is always higher using MLEM-STOP. However, if one focuses on the mean of the resolution figure of merit, it appears that MLEM-STOP produces images that are at least as accurate as those provided by MLEM-CONV for counting statistics higher than 10,000. Beyond 40,000 counts, the advantage of MLEM-STOP over MLEM-CONV becomes clear whatever the size of the studied disk. These latest results allow distinguishing roughly three ranges of counting statistics summarized in Table 2:

- From 10,000 to 40,000 counts: sharp advantage for MLEM-STOP in terms of noise and fairly similar results for the two methods in terms of resolution.
- From 40,000 to 100,000 counts: significant benefit in terms of noise and resolution when using MLEM-STOP.
- Over 100,000 counts: MLEM-STOP produces images significantly better in terms of resolution with noise properties that tend to be equivalent to those of the images obtained using MLEM-CONV.

It appears therefore that, inside the interval [40,000 100,000] counts, MLEM-STOP permits a significant benefit compared with MLEM-CONV regarding both noise and resolution properties of the reconstructed images. Outside this interval, MLEM-STOP always allow a significant improvement compared with MLEM-CONV, either concerning noise behavior (for low statistics from 10,000 to 40,000 counts) or resolution performances (for high statistics over 100,000 counts), hence leading to a better noise-resolution tradeoff.

Evaluation of the method using realistic data

In order to judge the feasibility of our stopping rule when working with real data, we reconstructed 2D sinograms of a Hoffman phantom simulated using the GATE platform for four different count rates ranging from 300,000 to 1.5 million. The sinograms were corrected for scatter and random and the system matrix used for reconstruction took into account the attenuation and detector sensitivity. The study shows (Fig. 12) that during the iterative process the bias decreases monotonically whereas the standard deviation increases monotonically. Initially, the resulting RMS error follows the trend of the bias and decreases. At one point, the positive slope of the variance becomes larger than the negative slope of the bias. The RMS reaches a minimum when the two slopes are equal and then increases following the trend of the variance. It appears that the optimal stopping value $\widehat{J}$ (corresponding to the RMS error minimum) is always very close to the theoretical value of 1, whatever the count rate. This confirms that our stopping rule allows to efficiently target the optimal bias-variance tradeoff. The consecutive relative RMS error increase resulting from an abortion of the iterative process once $J \leq 1$ is always mild and never exceeds 5 %. The results exposed in Fig. 13 confirm from a qualitative point of view that whatever the count rate the images obtained with our stopping criterion exhibit a good noise-resolution tradeoff with comparison to the images obtained for a few less or more iterations during the reconstruction process.

Conclusion

In this paper, starting from a heuristic description of the iterative reconstruction process in emission tomography and based on the statistical properties of the tomographic record, we proposed a simple stopping criterion which is theoretically relevant for any iterative reconstruction algorithm. Using Monte-Carlo simulations, we showed that our criterion allows stopping the reconstruction process for an estimated image that is very close to the optimal estimate in terms of Euclidian distance. In comparison with post-filtered MLEM, we proved that our stopping rule behaves significantly better in terms of RMS error and noise-resolution tradeoff for a wide range of counting statistics (10,000 to 300,000 counts) compatible with usual clinical settings. We then tested our stopping rule on realistic data produced with the GATE platform for a Hoffman brain phantom. Here again, the produced images are very close to the optimal ones in terms of RMS error and hence efficiently target the optimal bias-variance tradeoff during the iterative process.

Acknowledgments The authors would like to thank I. Buvat and A. Dubois for their technical support regarding the GATE simulations.

References

1. Gordon R, Bender R, Herman GT. Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and X-ray photography. J Theor Biol. 1970;29:471–81.
2. Brooks RA, Di Chiro G. Principles of computer assisted tomography (CAT) in radiographic and radioisotopic imaging. Phys Med Biol. 1976;21:689–732.
3. Goitein M. Three-dimensional density reconstruction from a series of twodimensional projections. Nucl Instrum Methods. 1972;101:509–18.
4. Gilbert PFC. Iterative methods for the three-dimensional reconstruction of an object from projections. J Theor Biol. 1972;36:105–17.

Author's personal copy

5. Xu XL, Liow JS, Strother SC. Iterative algebraic reconstruction algorithms for emission computed tomography: a unified framework and its application to positron emission tomography. Med Phys. 1993;20:1675–84.
6. Schmidlin P. Iterative separation of sections in tomographic scintigrams. Nucl Med. 1972;11(1):1–16.
7. Darroch JN, Ratcliff D. Generalized iterative scaling for log-linear models. Ann Math Stat. 1972;43:1470–80.
8. Dempster A, Laird N, Rubin D. Maximum likelihood from incomplete data via the EM algorithm. J R Stat Soc. 1977;39: 1–38.
9. Barrett HH, Swindell W. Radiological imaging—the theory of image formation, detection, and processing. New York: Academic Press; 1981.
10. Shepp VA, Vardi Y. Maximum likelihood reconstruction for emission tomography. IEEE Trans Med Imaging. 1982;1:113–22.
11. Byrne CL. Block-iterative methods for image reconstruction from projections. IEEE Trans Image Proc. 1996;5:792–4.
12. Byrne CL. Accelerating the EMML algorithm and related iterative algorithms by rescaled block-iterative methods. IEEE Trans Image Proc. 1998;7:100–9.
13. Hudson HM, Larkin RS. Accelerated image reconstruction using ordered subsets of projection data. IEEE Trans Med Imaging. 1994;13:601–9.
14. Browne J, de Pierro AB. A row-action alternative to the EM algorithm for maximizing likelihood in emission tomography. IEEE Trans Med Imaging. 1996;15:687–99.
15. Mesina CT, Boellaard R, Jongbloed G, Van der Vaart AW, Lammertsma AA. Experimental evaluation of iterative reconstruction versus filtered backprojection for 3D [15O]water PET activation studies using statistical parametric mapping analysis. Neuroimage. 2003;19:1170–9.
16. Lubberink M, Boellaard R, Van der Weerdt AP, Visser AC, Lammertsma AA. Quantitative comparison of analytic and iterative reconstruction methods in 2- and 3-dimensional dynamic cardiac 18F-FDG PET. J Nucl Med. 2004;45:2008–15.
17. Razifar P, Lubberink M, Schneider H, Lâgström B, Bengtsson E, Bergström M. Non-isotropic noise correlation in PET data reconstructed by FBP but not by OSEM demonstrated using autocorrelation function. BMC Med Imaging. 2005;5(1):3.
18. Liew SC, Hasegawa BH, Brown JK, Lang TF. Noise propagation in SPECT images reconstructed using an iterative maximum likelihood algorithm. Phys Med Biol. 1993;38:1713–27.
19. Mariano-Goulart D, Fourcade M, Bernon JL, Rossi M, Zanca M. Experimental study of stochastic noise propagation in SPECT images reconstructed using the conjugate gradient algorithm. Comput Med Imaging Graph. 2003;27:53–63.
20. Hebert TJ. Statistical stopping criteria for iterative maximum likelihood reconstruction of emission images. Phys Med Biol. 1990;35:1221–32.
21. Snyder DL, Miller MI, Thomas LJ, Politte DG. Noise and edge artifacts in maximum-likelihood reconstructions for emission tomography. IEEE Trans Med Imaging. 1987;6:228–38.
22. Falcon C, Juvells I, Pavia J, Ros D. Evaluation of a cross-validation stopping rule in MLE SPECT reconstruction. Phys Med Biol. 1998;43:1271–85.
23. Liang Z, Jaszczak R, Greer K. On Bayesian image reconstruction from projections: uniform and nonuniform a priori source information. IEEE Trans Med Imaging. 1989;8:227–35.
24. Fessler JA, Hero AO. Penalized maximum-likelihood image reconstruction using space-alternating generalized EM algorithms. IEEE Trans Image Proc. 1995;4:1417–29.
25. Tanaka E, Kudo H. Subset-dependent relaxation in block-iterative algorithms for image reconstruction in emission tomography. Phys Med Biol. 2003;48:1405–22.
26. Slijpen ETP, Beekman FJ. Comparison of post-filtering and filtering between iterations for SPECT reconstruction. IEEE Trans Nucl Sci. 1999;46:2233–8.
27. Knuth DE. Seminumerical algorithms. The art of computer programming, vol 2. Boston: Addison Wesley; 1969.
28. Jan S et al. for the OpenGATE collaboration, et al. GATE: a simulation toolkit for PET and SPECT. Phys Med Biol. 2004;49:4543–61.
29. Lamare F, Turzo A, Bizais Y, Le Rest CC, Visvikis D. Validation of a Monte Carlo simulation of the Philips Allegro/GEMINI PET systems using GATE. Phys Med Biol. 2006;51:943–62.
30. Bailey DL, Meikle SR. A convolution-subtraction scatter correction method for 3D PET. Phys Med Biol. 1994;39:411–24.
31. Kadrmas DJ. LOR-OSEM: statistical PET reconstruction from raw line-of-response histograms. Phys Med Biol. 2004;49:4731–44.