



Statistic-algebraic Tuning Optimization for Regularized Iterative Reconstruction (PML and EMS) in Emission Tomography

Fayçal Ben Bouallègue, Jean-Francois Crouzet, A Dubois, Denis Mariano-Goulart

► To cite this version:

Fayçal Ben Bouallègue, Jean-Francois Crouzet, A Dubois, Denis Mariano-Goulart. Statistic-algebraic Tuning Optimization for Regularized Iterative Reconstruction (PML and EMS) in Emission Tomography. IEEE Transactions on Nuclear Science, 2013, 60 (1), pp.144-157. <10.1109/TNS.2012.2226911>. <hal-00799467>

HAL Id: hal-00799467

<https://hal.science/hal-00799467v1>

Submitted on 20 Apr 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

STATISTIC-ALGEBRAIC TUNING OPTIMIZATION FOR REGULARIZED ITERATIVE RECONSTRUCTION (PML AND EMS) IN EMISSION TOMOGRAPHY

F Ben Bouallègue, JF Crouzet, A Dubois, and D Mariano-Goulart

Abstract

Most of the regularized iterative reconstruction schemes employed in emission tomography (such as penalized maximum-likelihood, PML) usually require the adjustment of a scalar parameter β that determines the strength of the a-priori information regarding the studied object. Empirical selection of β remains hazardous since its optimal value depends on the morphological structure of the reconstructed image and the data signal to noise ratio (SNR), which explains partly the scarce utilization of penalized reconstruction in clinical routine. In this paper, we derive a simple optimization criterion for β that relies on a statistical description of the noise propagation when iteratively updating the image estimate and on a surrogate algebraic formulation holding for both PML and expectation-maximization-smooth (EMS) iterative reconstruction. When incorporated into each iteration step, the statistic-algebraic tuning optimization (SATO) yields two new optimized regularized iterative methods: SATO-PML and SATO-EMS. These methods are compared with classical MLEM reconstruction followed by optimal Gaussian post-filtering (ML-opt) through Monte-Carlo experiments involving the Hoffman brain phantom and the Shepp-Logan phantom. It is shown that, whatever the studied object and the count rate, SATO-PML and SATO-EMS are convergent in terms of β and produce images with bias, variance and contrast properties that are at least as good as those of the ML-opt images. The two proposed algorithms are also evaluated using realistic PET data from a Hoffman phantom produced using the GATE platform in order to demonstrate the feasibility of our SATO scheme with actual data.

Keywords

Emission tomography, iterative reconstruction, regularization, penalized maximum-likelihood, expectation-maximization-smooth.

INTRODUCTION

Emission tomography seeks to extract qualitative and quantitative information on radioactive tracer distribution from its projection record. Numerous image reconstruction methods have been proposed in the last few decades to perform this task, either based on analytical or iterative approaches. Since the reference publications of Shepp & Vardi [1] and Lange & Carson [2], iterative algorithms based on likelihood maximization have increasingly gained in popularity, in both the research field and routine clinical practice. Besides their ability to efficiently model the photon emission and detection phenomena inside the reconstruction process, most iterative techniques can account for the statistical properties of the recorded data [3]. However, a considerable number of papers on the maximum-likelihood expectation-maximization (MLEM) algorithm have clearly shown that, despite its asymptotically unbiased character, unregularized MLEM inevitably leads to an undesirable increase in the estimate variance while the iterations proceed, thereby leading to degradation in the produced image (in terms of the signal to noise ratio among others) [4-8]. The noise propagation remains complex from a theoretical

point of view [6-7], whereas its effects on the reconstructed images degrade qualitative interpretation in a clinical context. In order to propose smooth images suitable for clinical purpose, some regularization has to be introduced into the ML approach. The most natural regularization method consists of either arbitrarily stopping the iterative process at an early stage or a linear low-pass filtering of the object estimate [9] — a post-filtering option using a convolution kernel with adjustable size is therefore available on most of the commercial SPECT or PET reconstruction software suites. These two methods, albeit commonly employed, remain broadly empirical. Some statistical stopping criteria for MLEM reconstruction relying on feasibility or cross-validation have been proposed in the early nineties [10-14]. The expectation-maximization-smooth (EMS) approach performs the low-pass filtering in between each iteration step [15-17]. Although there is no proof of convergence, this algorithm always appears to converge in practice, and the image characteristics are very similar to those of post-filtered ML images (provided that the optimal filters are selected in both cases) [16]. Known as penalized maximum-likelihood (PML), the second main type of regularization method relies on Bayesian reconstruction using maximum a-posteriori estimators [18-21]. A simplified approach allowing an iterative implementation of PML estimation was developed by Green [22], and various energy functions have been proposed to model the a-priori information regarding the reconstructed image [23-28]. However, although the resolution and noise properties of the PML algorithm have been widely studied and well established [29-32], its exploitation in routine practice remains uncommon in part due to the lack of consensus on how to optimally fix the tuning parameter (usually noted β) determining the strength of the roughness penalty [23]. The same drawback holds for the EMS technique, in which the optimal size of the filter depends on the studied object and the data statistics. A decade ago, several papers have been written about computing β such that a known and approximately uniform resolution is obtained in the reconstructed image [33-34]. In the present article, which is based on a statistical description of the noise propagation at each iteration, we propose a simple criterion allowing to adjust the tuning parameter (β , filter size) of a regularized algorithm with particular applications to PML and EMS. Tuning optimization is achieved at each iteration throughout the reconstruction process using a surrogate algebraic formulation holding for the two studied algorithms. Our methodology is evaluated through Monte-Carlo simulations using the Hoffman brain phantom and the Shepp-Logan phantom and two levels of counting statistics. The performances of the two new regularized algorithms with tuning optimization are compared with those of classical MLEM reconstruction followed by optimal Gaussian post-filtering in terms of bias, variance, root-mean-square error, and contrast¹. Last, reconstructed slices of a Hoffman phantom PET acquisition generated using the GATE platform are displayed in order to provide insight into the potential efficiency of our algorithms when working with actual data.

MATERIALS AND METHODS

A. Classical regularized iterative reconstruction

Let \mathbf{p} be the projections recorded by the acquisition device (\mathbf{p}_i stands for the number of photons registered in bin i) and \mathbf{x} the activity distribution to be estimated (\mathbf{x}_j stands for the activity inside image pixel j). We denote with \mathbf{A} the system matrix whose component \mathbf{A}_{ij} accounts for the probability for a photon emitted from pixel j to be recorded into bin i . Ideally the system matrix is intended to perfectly model the emission/detection process, including side-effects such as auto-attenuation, Compton

¹ Obviously, the problem of regularization optimization is highly complex and should ideally be defined in terms of the clinical task that is to be achieved. For a first and global evaluation, the studied figures of merit have been chosen for their convenience and universality.

scattering, and detector response function. For a given reconstruction task, the objective function $O(\mathbf{x}; \mathbf{p})$ of an image \mathbf{x} is defined as its conditional probability given the record \mathbf{p} . Using Bayes rule, a linear Poisson model, and the monotonicity of the logarithm, finding the image that maximizes the objective function is equivalent to finding the maximal argument of:

$$\text{Log}O(\mathbf{x}; \mathbf{p}) = \sum_i [\mathbf{p}_i \log(\mathbf{q}_i) - \mathbf{q}_i - \log(\mathbf{p}_i!)] + \log(P(\mathbf{x})) - \log(P(\mathbf{p})) \quad (1)$$

where $\mathbf{q} = \mathbf{A}\mathbf{x}$ and $P(\mathbf{x})$ is the intrinsic probability of image \mathbf{x} and called the prior. The first term in the right-hand side is $\log(P(\mathbf{p}|\mathbf{x}))$ and called the log-likelihood. It ensures the adequacy between the recorded data and the reconstructed image. The second term can be viewed as a regularization term to reinforce the injectivity of the reconstruction problem $\mathbf{p} = \mathbf{A}\mathbf{x}$, which is known to be ill-conditioned. The third term is a constant for a given reconstruction problem, it has no impact on the maximization process and is thus omitted in all the subsequent computations. When no a priori assumption is made regarding the activity distribution, the prior is set to constant and the produced estimate $\bar{\mathbf{x}}$ corresponds to the unregularized maximum-likelihood (ML) estimate:

$$\bar{\mathbf{x}} = \underset{\mathbf{x} \in \mathbb{R}^N}{\operatorname{argmax}} \{ \sum_i [\mathbf{p}_i \log(\mathbf{q}_i) - \mathbf{q}_i] \} \quad (2)$$

The ML estimate is classically built using the iterative expectation-maximization (EM) algorithm whose n^{th} iterate can be written as (where f_{ML} stands for the ML updating function):

$$\mathbf{x}_j^n = [f_{ML}(\mathbf{x}^{n-1})]_j = \mathbf{x}_j^{n-1} \frac{1}{\sum_i \mathbf{A}_{ij}} \sum_i \mathbf{A}_{ij} \mathbf{p}_i / \mathbf{q}_i^{n-1} \quad (3)$$

The MLEM algorithm is known to provide an asymptotically unbiased estimator of the true activity distribution. However, achieving the low-bias estimate requires a large number of iteration steps, which inevitably increases the estimate variance and results in highly noisy images. In practice, no simple and effective bias-variance tradeoff can be globally established since the convergence properties of MLEM strongly depend on the studied object and the count level. Besides, stopping the iterative process at an early stage remains unsafe because the spatial resolution can be very position-dependent at low iteration numbers. In order to reduce noise propagation, regularized methods have been proposed that consist of adding some a priori knowledge regarding the smoothness of the studied object. In the PML algorithm, the prior² is usually expressed using a Gibbs energy function penalizing the likelihood [18-20]:

$$P(\mathbf{x}) = \frac{1}{Z} e^{-\beta U(\mathbf{x})} \quad (4)$$

where Z is a normalization constant, β a tuning parameter, and U the energy function. One choice for U which is also the most commonly employed is based on the quadratic penalty:

$$U(\mathbf{x}) = \frac{1}{4} \sum_j \sum_{k \in N_j} w_k (\mathbf{x}_j - \mathbf{x}_k)^2 \quad (5)$$

with N_j denoting the neighborhood of pixel j and w a normalized weighting function. In this paper, we only consider clique sizes of 8 with $w_k \propto 1$ for “direct” nearest neighbors and $w_k \propto 1/\sqrt{2}$ for “diagonal” next-nearest neighbors. The PML estimate $\bar{\mathbf{x}}$ is defined as:

$$\bar{\mathbf{x}} = \underset{\mathbf{x} \in \mathbb{R}^N}{\operatorname{argmax}} \{ \sum_i [\mathbf{p}_i \log(\mathbf{q}_i) - \mathbf{q}_i - \beta U(\mathbf{x})] \} \quad (6)$$

² Throughout this manuscript, the term “prior” will be used in the sense of “penalty” in the context of PML reconstruction, albeit the employed priors imply heuristic considerations rather than objective a-priori knowledge.

Using an EM scheme and the one-step-late (OSL) approach [22], the MAP estimate is built iteratively following:

$$\mathbf{x}_j^n = \left[f_{MAP}^\beta(\mathbf{x}^{n-1}) \right]_j = \frac{\mathbf{x}_j^{n-1}}{\sum_i \mathbf{A}_{ij} + \beta \partial_j U(\mathbf{x}^{n-1})} \sum_i \mathbf{A}_{ij} \frac{\mathbf{p}_i}{\mathbf{q}_i^{n-1}} = \frac{1}{1 + \beta \frac{\sum_{k \in N_j} w_k (\mathbf{x}_j^{n-1} - \mathbf{x}_k^{n-1})}{\sum_i \mathbf{A}_{ij}}} [f_{ML}(\mathbf{x}^{n-1})]_j \quad (7)$$

where f_{MAP}^β stands for the PML updating function with tuning parameter β . Another classical regularization technique involves a low-pass filtering of the estimator, either after the iterative process or in between each iteration step. In the EMS technique, the ML estimator is convolved after each iteration using a Gaussian kernel [15-17]. The EMS estimate $\bar{\mathbf{x}}$ is generated iteratively according to:

$$\mathbf{x}_j^n = \left[f_{EMS}^\beta(\mathbf{x}^{n-1}) \right]_j = [f_{ML}(\mathbf{x}^{n-1}) * G_\beta]_j \quad (8)$$

where $*$ stands for the convolution operator and G_β is a centered Gaussian kernel with a full width at half maximum (FWHM) of β pixels. f_{EMS}^β stands for the EMS updating function with tuning parameter β^3 . The ability of the two aforementioned methods (PML and EMS) to significantly reduce noise in the reconstructed slices while preserving sufficient resolution depends on the selection of the tuning parameter β which controls the strength of the a priori information concerning the reconstructed object. With regard to PML restoration, in [23] Herbert and Leahy observed that the optimal value of β varies with both the object shape and structure and the signal to noise ratio (SNR) in the recorded data. Contrariwise, when studying the same image with the same SNR, the optimal β value remains the same whatever the realization of the noise. These considerations also hold for EMS reconstruction. Based on the statistical properties of the iterative reconstruction process and a surrogate algebraic formulation, we propose in the following section a simple criterion for optimizing the tuning parameter value throughout the iterations.

B. Statistic-algebraic tuning optimization

Our goal here is to develop a procedure to optimize the choice of the tuning parameter β at each iteration. The optimal β value (denoted β^{opt} in the sequel) would naturally be the value which allows an optimal noise control while allowing efficient reconstruction of the useful signal. For a given regularized algorithm REG (REG = PML or EMS) with tuning parameter β and a given iteration index n , we define the correction $\delta^{\beta,n}$ as the difference between the updated image obtained using the regularized algorithm and the updated image obtained using the ML updating function:

$$\delta_j^{\beta,n} = \left[f_{REG}^\beta(\mathbf{x}^{n-1}) - f_{ML}(\mathbf{x}^{n-1}) \right]_j = \left[f_{REG}^\beta(\mathbf{x}^{n-1}) \right]_j - \mathbf{x}_j^{n-1} \frac{1}{\sum_i \mathbf{A}_{ij}} \sum_i \mathbf{A}_{ij} \mathbf{p}_i / \mathbf{q}_i^{n-1} \quad (9)$$

Let us denote with h the exact ML updating function:

$$[h(\mathbf{x}^{n-1})]_j = \mathbf{x}_j^{n-1} \frac{1}{\sum_i \mathbf{A}_{ij}} \sum_i \mathbf{A}_{ij} \hat{\mathbf{p}}_i / \mathbf{q}_i^{n-1} \quad (10)$$

where $\hat{\mathbf{p}}$ are the exact (noise-free) projections of the object. We define the ideal correction \mathbf{d}^n at iteration n as:

³ Although the tuning parameters of MAP and EMS have different interpretations and dimensions, the authors chose to name them with the same letter β in order to simplify the notations throughout the article.

$$\mathbf{d}_j^n = [h(\mathbf{x}^{n-1}) - f_{ML}(\mathbf{x}^{n-1})]_j = \mathbf{x}_j^{n-1} \frac{1}{\sum_i \mathbf{A}_{ij}} \sum_i \mathbf{A}_{ij} (\hat{\mathbf{p}}_i - \mathbf{p}_i) / \mathbf{q}_i^{n-1} \quad (11)$$

The aim of this section is to describe a statistic-algebraic approach in which, at each iteration, the optimal value of β is defined as that providing an adequacy between the correction $\boldsymbol{\delta}^{\beta,n}$ and the ideal correction \mathbf{d}^n . From a Bayesian point of view, $\boldsymbol{\delta}^{\beta,n}$ is a constant since the recorded projections \mathbf{p} are constant for a given reconstruction task. Besides, \mathbf{d}^n is a random variable whose statistical properties follow from the uncertainty about the exact projections $\hat{\mathbf{p}}$. Since $E(\hat{\mathbf{p}} - \mathbf{p}) = 0$ and $Var(\hat{\mathbf{p}} - \mathbf{p}) \approx Diag(\hat{\mathbf{p}}) \approx Diag(\mathbf{p})$, the ideal correction has a null mean and a variance:

$$\mathbf{V}^n \triangleq Cov(\mathbf{d}^n) ; \quad \boldsymbol{\sigma}_j^n \triangleq \mathbf{V}_{jj}^n = Var(\mathbf{d}_j^n) \approx (\mathbf{x}_j^{n-1})^2 \frac{1}{(\sum_i \mathbf{A}_{ij})^2} \sum_i \mathbf{A}_{ij}^2 \mathbf{p}_i / \mathbf{q}_i^{n-1}^2 \quad (12)$$

We define the objective function $J^n(\beta)$ as the mean square error between the correction $\boldsymbol{\delta}^{\beta,n}$ and the ideal correction \mathbf{d}^n at iteration n :

$$J^n(\beta) = E \{ \|\boldsymbol{\delta}^{\beta,n} - \mathbf{d}^n\|^2 \} \quad (13)$$

A purely statistical optimization of the tuning parameter would consist in finding at each iteration β^{opt} such that:

$$\beta^{opt} = \underset{\beta \in \mathbb{R}^+}{argmin} \{ J^n(\beta) \} \quad (14)$$

However, solving (14) using (12) leads trivially to:

$$\beta^{opt} = \underset{\beta \in \mathbb{R}^+}{argmin} \{ \|\boldsymbol{\delta}^{\beta,n}\|^2 + Tr(\mathbf{V}^n) \} = \underset{\beta \in \mathbb{R}^+}{argmin} \{ \|\boldsymbol{\delta}^{\beta,n}\|^2 \} = 0 \quad (15)$$

since $\boldsymbol{\delta}^{0,n} = 0$. In order to avoid this drawback that arises from a purely statistical formulation of the optimization problem, we propose to resort to a statistic-algebraic formulation based on the definition of a surrogate of the ideal correction \mathbf{d} denoted $\check{\mathbf{d}}$ associated with a surrogate objective function \check{J} . The surrogate $\check{\mathbf{d}}$ is a constant vector built using the statistical properties of \mathbf{d} . Various definitions of $\check{\mathbf{d}}$ may be considered and the authors have tested some of them, which led them to select empirically the following. The surrogate ideal correction $\check{\mathbf{d}}$ is built such that for each of its components:

$$\forall j, \check{\mathbf{d}}_j^{n^2} = E(\mathbf{d}_j^{n^2}) = Var(\mathbf{d}_j^n) = \mathbf{V}_{jj}^n ; \quad i.e., \quad \check{\mathbf{d}}_j \triangleq \pm \boldsymbol{\sigma}_j^n \quad (16)$$

The associated surrogate objective function is then:

$$\check{J}^n(\beta) = \|\boldsymbol{\delta}^{\beta,n} - \check{\mathbf{d}}^n\|^2 \quad (17)$$

The optimal value of the regularization parameter β^{opt} is now defined as:

$$\beta^{opt} = \underset{\beta \in \mathbb{R}^+}{argmin} \{ \check{J}^n(\beta) \} \quad (18)$$

It follows from (16) that:

$$\beta^{opt} = \underset{\beta \in \mathbb{R}^+}{argmin} \{ \|\boldsymbol{\delta}^{\beta,n}\|^2 + Tr(\mathbf{V}^n) - 2 \check{\mathbf{d}}^{n^T} \boldsymbol{\delta}^{\beta,n} \} \quad (19)$$

where T denotes the matrix transpose. Equation (19) is very close to equation (15) due to the definition of $\check{\mathbf{d}}^n$. However, $\check{J}^n(\beta)$ includes a third term that will allow for a non-trivial solution of the optimization problem. As the goal is to minimize $\check{J}^n(\beta)$, the sign of the components of $\check{\mathbf{d}}^n$ in equation (16) will be chosen so that, before the optimization process, this third term (and hence $\check{J}^n(\beta)$) is minimal which implies that:

$$\forall j, \check{\mathbf{d}}_j \triangleq \text{sign}(\delta_j^{\beta,n}) \sigma_j^n \quad (20)$$

where σ_j^n is estimated using equation (12). The surrogate $\check{\mathbf{d}}^n$ is a vector having the same order of magnitude as the ideal correction \mathbf{d}^n for every dimension of the image hyperspace. The sign of each component of $\check{\mathbf{d}}^n$ is arbitrarily chosen so as to fit with the sign of the corresponding component of the current correction $\delta^{\beta,n}$, which roughly implies that the correction brought by the regularized algorithm is implicitly considered to be a “good” correction. The situation at iteration n is illustrated on Figure 1. Let us denote $\gamma^{\beta,n}$ the projection of $\check{\mathbf{d}}^n$ on $\delta^{\beta,n}$:

$$\gamma^{\beta,n} = \delta^{\beta,n} \check{\mathbf{d}}^{nT} \delta^{\beta,n} / \|\delta^{\beta,n}\|^2 \quad (21)$$

Resolving equation (18) could be achieved using any classical minimization method such as the steepest descent or conjugate gradient. However, the solution becomes straightforward if we consider a particular class of corrections. We define a collinear correction as a correction such that $\delta^{\beta_1,n}$ and $\delta^{\beta_2,n}$ are collinear whatever $(\beta_1, \beta_2) \in \mathbb{R}^{*+2}$, i.e., there exists some λ such that $\delta^{\beta_1,n} = \lambda \delta^{\beta_2,n}$. As PML and EMS produce fairly collinear corrections in first approximation (see Appendix A), we will rely on this property to solve equation (18). Based on basic geometric considerations, for a collinear correction, the optimal value β^{opt} is such that:

$$\gamma^{\beta^{opt},n} = \delta^{\beta^{opt},n} \quad (22)$$

Defining the scaling factor $\kappa^{\beta,n}$ as:

$$\kappa^{\beta,n} = \|\gamma^{\beta,n}\| / \|\delta^{\beta,n}\| = \check{\mathbf{d}}^{nT} \delta^{\beta,n} / \|\delta^{\beta,n}\|^2 \quad (23)$$

it results that:

$$\kappa^{\beta^{opt},n} = 1 \quad (24)$$

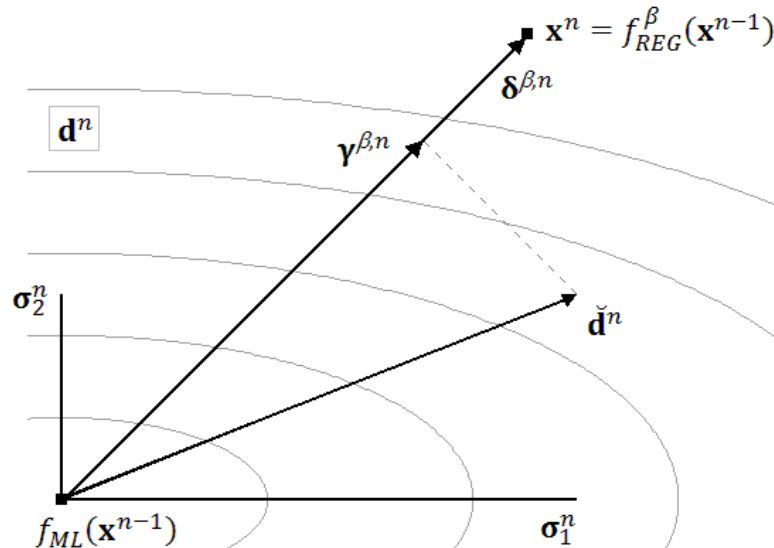


Figure 1. Situation at iteration n , displaying the distribution of the ideal correction \mathbf{d}^n (gray level curves and standard deviation σ^n), its surrogate $\check{\mathbf{d}}^n$, the current correction $\delta^{\beta,n}$, and $\gamma^{\beta,n}$.

From a qualitative point of view, equation (24) means that if $\kappa^{\beta,n} > 1$ the regularization is too weak and the correction has to be increased by increasing β . Contrariwise, if $\kappa^{\beta,n} < 1$, β has to be diminished. Equation (24) will be referred to as the statistic-algebraic tuning optimization (SATO) criterion. This criterion can be used to design SATO-PML and SATO-EMS algorithms which are conceived to progressively fulfill the SATO criterion throughout the iterations. In practice, the general recipe for the implementation of these two algorithms is the following:

- Initialize arbitrarily the image estimate \mathbf{x}^0 , for instance, using a non-filtered back-projection.
- Initialize β arbitrarily.
- Start the iterative process. At each iteration $n > 1$, do the following:
 - Compute $\check{\mathbf{d}}^n$ using (12) and (20).
 - Compute \mathbf{x}^n using (7) or (8), $\delta^{\beta,n}$ using (9), and $\kappa^{\beta,n}$ using (21) and (23).
 - Modify β in order to force $\kappa^{\beta,n}$ toward 1 using (see Appendix B for details):

$$\text{SATO-PML: } \beta^{new} = \kappa^{\beta,n} \beta^{old} \quad (25)$$

$$\text{SATO-EMS: } \beta^{new} = \beta^{old} / \sqrt{1 - \frac{1}{4}(\beta^{old})^2 \ln(\kappa^{\beta,n})/\ln(2)} \quad (26)$$

- Proceed to the next iteration.

C. Experimental assessment

The experimental assessment of our SATO methodology relied on numerical simulations and physical phantom acquisitions. In order to evaluate to what extent our technique overcomes the aforementioned drawbacks of classical regularized iterative reconstruction, we tested it using two numerical phantoms and two levels of counting statistics. The two studied phantoms were the Hoffman brain phantom and the Shepp-Logan phantom (Figure 2). In these two phantoms, a hot spot was included standing for a hyperactive tumor. Table I summarizes the quantitative specifications of the two phantoms. The performances of SATO-PML and SATO-EMS were compared with those provided by classical MLEM reconstruction followed by Gaussian post-filtering (denoted ML-filter). The objects were defined and reconstructed over a 128×128 grid. The projections were simulated using 64 projection angles over 180° (128 bins per projection angle). The system matrix used to simulate and reconstruct the data was based on a uniformly distributed pixel activity model and did not model the attenuation, scatter, and detector response. The two levels of counting statistics considered were 100,000 events on average (low count rate) and 1 million events on average (high count rate). The Poisson noise was added to the noisy replicates of the projections using Knuth's algorithm [35].

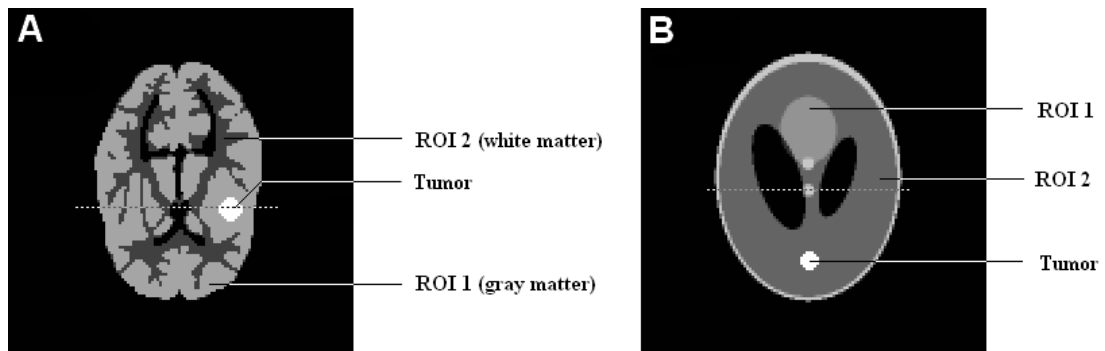


Figure 2. The two studied numerical phantoms (A: Hoffman. B: Shepp-Logan).

TABLE I. QUANTITATIVE SPECIFICATIONS OF THE TWO STUDIED PHANTOMS

	Tumor size (pixels)	Tumor/neighborhood contrast	ROI 1/ROI 2 contrast
Hoffman brain phantom	57	50 %	150 %
Shepp-Logan phantom	37	150 %	50 %

For each pair {phantom, count level}, the modulus operandi was the following:

- Simulate 50 noisy realizations of the phantom projections.
- Reconstruct each realization with SATO-PML, SATO-EMS and MLEM. The total number of iterations was 150 for the low count rate sinograms and 300 for the high count rate sinograms. For the SATO algorithms, the first image estimate \mathbf{x}^0 was initialized using an unfiltered back-projection, and β was randomly initialized inside the interval $[10^{-5} \ 10^{-1}]$ for SATO-PML and inside the interval $[0.5 \ 2.5]$ for SATO-EMS.
- Post-filter the MLEM estimate with Gaussian kernels of FWHM ranging from 0.5 to 5 pixels by steps of 0.05. Based on the computation of the root mean square (RMS) error between the filtered estimate and the exact object, define the optimal ML-filter estimate as that with the lower RMS error. This optimal estimate is denoted ML-opt.
- For each of the reconstructed objects, compute the following characteristics:
 - RMS error between the reconstructed image and the exact object.
 - Image mean coefficient of variation (CV), expressed in percent:

$$CV = \sqrt{\sum_j Var(\bar{\mathbf{x}}_j) / \sum_j E(\bar{\mathbf{x}}_j)^2} \quad (27)$$

- Tumor to neighborhood contrast, expressed in percent of the exact value given in Table I:

$$\text{Tumor contrast} = \frac{(\text{mean tumor activity} - \text{mean neighborhood activity}) / \text{mean neighborhood activity}}{\text{exact contrast value}} \quad (28)$$

- ROI 1 to ROI 2 contrast (gray to white matter contrast for the Hoffman phantom), expressed in percent of the exact value given in Table I:

$$\text{ROI contrast} = \frac{(\text{mean ROI 1 activity} - \text{mean ROI 2 activity}) / \text{mean ROI 2 activity}}{\text{exact contrast value}} \quad (29)$$

The three described methods were finally used to reconstruct two slices of a Hoffman brain phantom produced using the GATE (Geant4 Application in Emission Tomography) platform [36]. The numerical Hoffman phantom employed was an axially invariant 3D phantom whose 2D slice was similar to that described above (except that it contained three hot spots) 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) [37]. Positron range and non-collinearity were not modeled. Two 3D data sets were produced: a first with a total activity of 40 MBq (^{18}F FDG) and an acquisition time of 3 minutes, and a second with a total activity of 80 MBq (^{18}F FDG) and an acquisition time of 12 minutes. The scatter and random annihilations were corrected using the scatter and random estimations provided by the software [38-39]. For each of the two 3D data sets, the 25 (among 29) central transaxial sinograms were summed in order to produce two high-statistics 2D sinograms gathering about 1.3 million and 10 million prompts

(total net trues). The system matrix employed for the reconstructions was built using a uniformly distributed pixel activity model and corrected for attenuation and normalization 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. The total number of iterations was 200 for the SATO algorithms and 400 for MLEM. The MLEM estimate was post-convoluted with Gaussian kernels of FWHM 2.5, 3 and 3.5 pixels for the 1.3 million count sinogram and 2, 2.5 and 3 pixels for the 10 million count sinogram. The SATO-PML algorithm was implemented using two priors: the quadratic prior described in equations (4-5), and a median prior defined by the energy function [26]:

$$U(\mathbf{x}) = \frac{1}{2} \sum_j \sum_{k \in N_j} w_k |\mathbf{x}_j - \mathbf{m}_k| \quad (30)$$

where \mathbf{m}_k stands for the median pixel value in the neighborhood N_k of pixel k , and with w the same weighting function as that used in equation (5).

RESULTS

Figure 3 shows the results of the numerical experiment corresponding to the Hoffman brain phantom with low count rate data. Figures 4, 5 and 6 are designed similarly and are respectively related to the Hoffman phantom with high count rate, the Shepp-Logan phantom with low count rate and the Shepp-Logan phantom with high count rate. Each of these four figures is composed of four subplots. Subplot A illustrates the convergence properties of SATO-PML (top row) and SATO-EMS (bottom row). From left to right are the change with iteration of scaling factor κ , the tuning parameter β , and the RMS error between the estimates and the exact object. Each figure of subplot A shows the 95% confidence interval of the corresponding variable, defined as the mean value ± 1.96 times the standard deviation over the 50 realizations. Subplot B displays the tradeoff curves for ML-filter, SATO-PML and SATO-EMS (the tradeoff curves for ML-filter were obtained by varying the size of the post-smoothing kernel as explained above). The bias-variance tradeoff (where the bias is defined as the RMS error of the mean estimate), the RMS error-variance tradeoff, and the two contrast-variance tradeoffs (tumor then ROI contrasts) are displayed from top left to bottom right. Subplot C proposes reconstructed images. The results for ML-opt, SATO-PML and SATO-EMS are shown from left to right. The top row shows a representative reconstruction for each method. The middle and last rows correspond respectively to the mean and standard deviation of the pixel values. Last, subplot D displays profiles through the dotted lines shown on Figure 2. The mean profile of the pixel values is shown on top and below is the profile of the standard deviation of the pixel values. Table II (resp. Table III) summarizes the relative performances of SATO-PML (resp. SATO-EMS) compared with ML-opt in terms of bias, variance (via the CV), RMS error, and contrast (tumor and ROI), for the four numerical experiments. Significant improvement or worsening (over 5% in absolute value) is highlighted using respectively light or dark gray cell shading.

Figures 7 shows the results for the GATE Hoffman phantom study (A: 1.3 million count sinogram and B: 10 million count sinogram). The top images are the reconstructions obtained using MLEM followed by Gaussian post-filtering of FWHM 2.5, 3 and 3.5 pixels (A) and 2, 2.5 and 3 pixels (B). The bottom images are the reconstructions produced by (from left to right) SATO-PML with the quadratic prior, SATO-PML with the median prior, and SATO-EMS. The curves show the convergence of the tuning parameter β for (left to right) SATO-PML with the quadratic prior, SATO-PML with the median prior, and SATO-EMS.

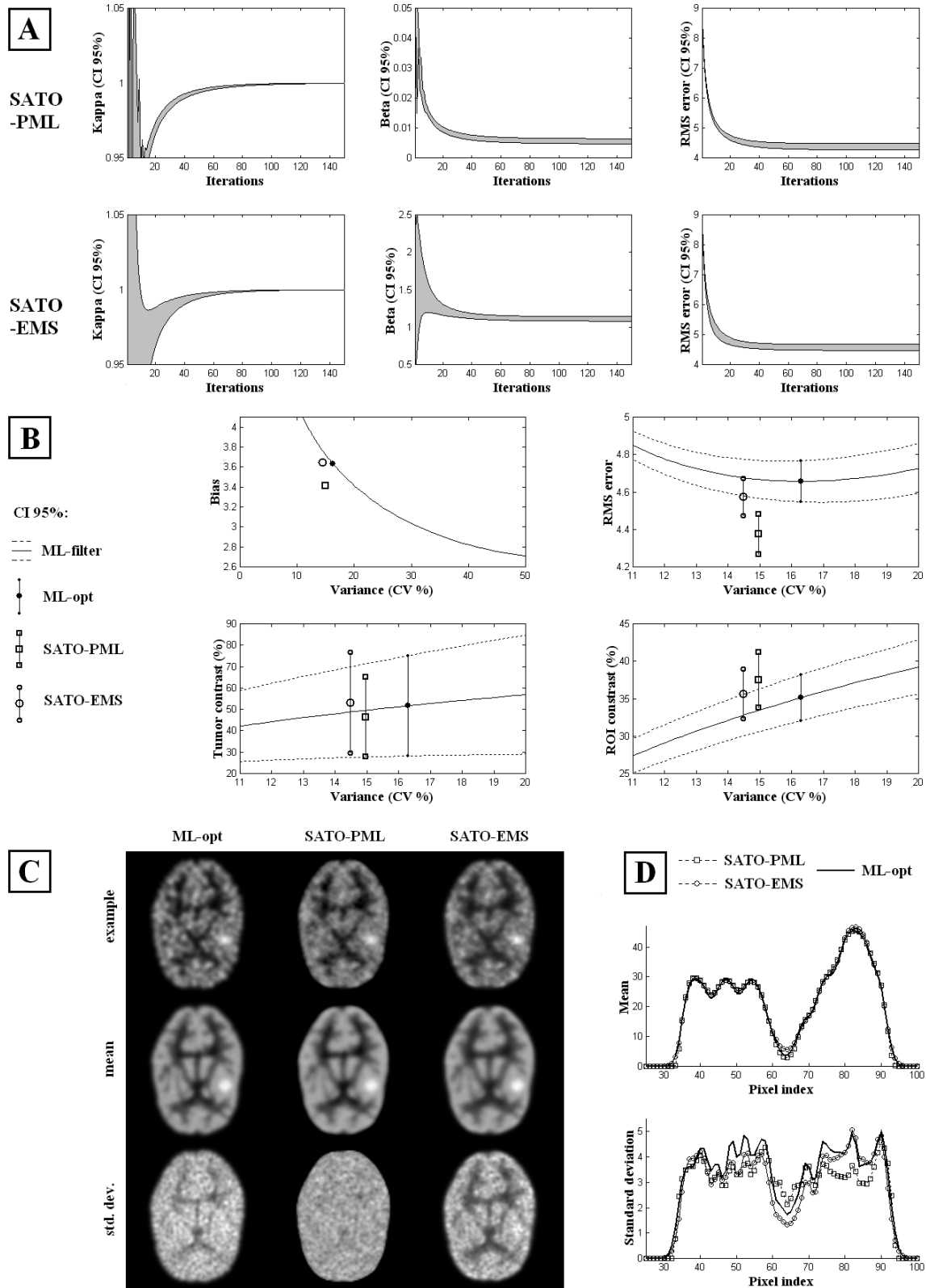


Figure 3. Results for the Hoffman phantom at low count rate. **A:** convergence. Evolution with the iteration of (from left to right): scaling factor κ , tuning parameter β , and RMS error between the estimate and the exact object (top row: SATO-PML; bottom row: SATO-EMS). **B:** tradeoff curves. Top left to bottom right: bias-variance, RMS error-variance, and contrast-variance (tumor then ROI). **C:** reconstructed images. Left to right: ML-opt, SATO-PML and SATO-EMS. Top to bottom: example, mean, and standard deviation. **D:** profiles through the lines displayed on Figure 2. Left: mean. Right: standard deviation.

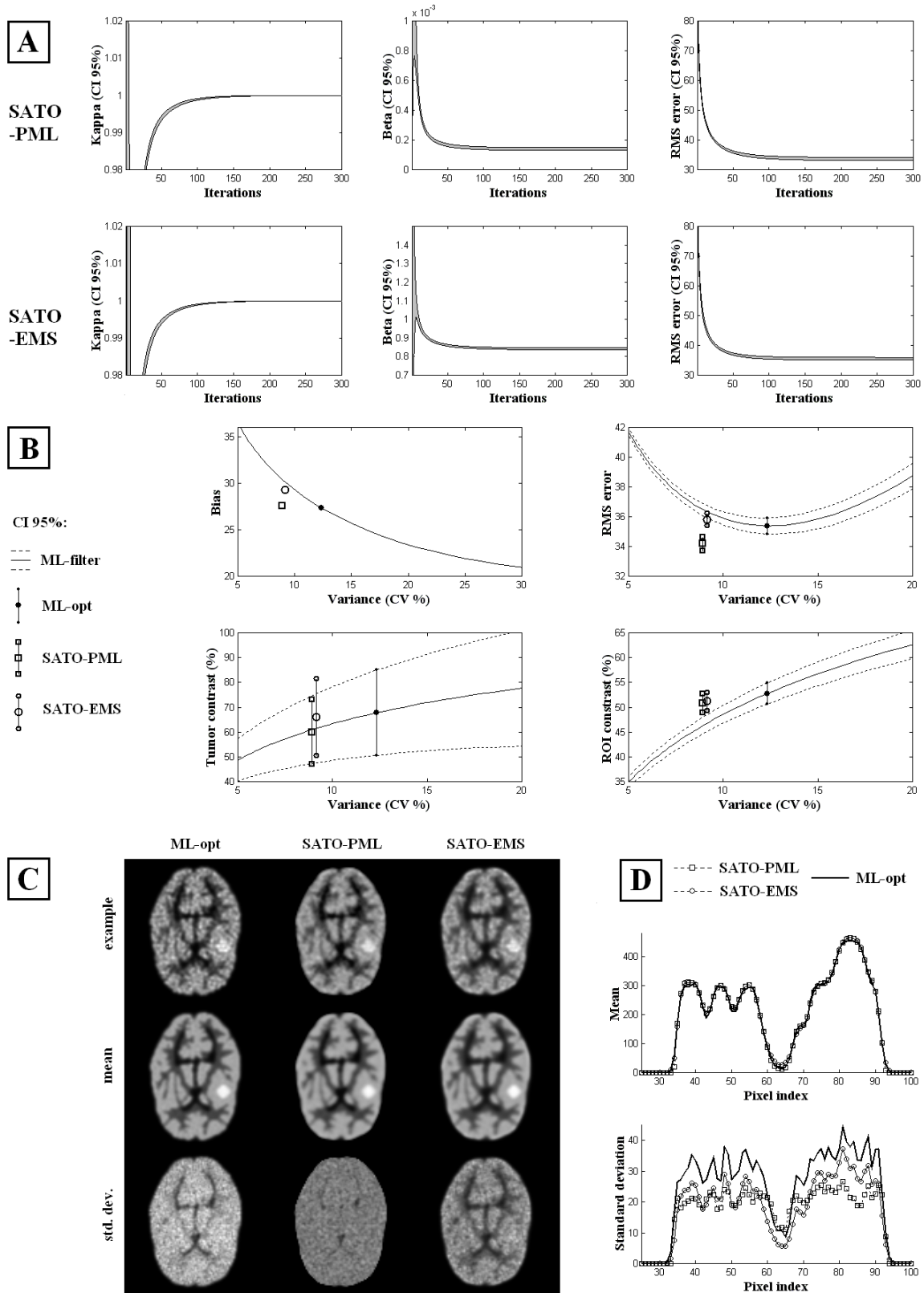


Figure 4. Results for the Hoffman phantom at high count rate. **A**: convergence. Evolution with the iteration of (from left to right): scaling factor κ , tuning parameter β , and RMS error between the estimate and the exact object (top row: SATO-PML; bottom row: SATO-EMS). **B**: tradeoff curves. Top left to bottom right: bias-variance, RMS error-variance, and contrast-variance (tumor then ROI). **C**: reconstructed images. Left to right: ML-opt, SATO-PML and SATO-EMS. Top to bottom: example, mean, and standard deviation. **D**: profiles through the lines displayed on Figure 2. Left: mean. Right: standard deviation.

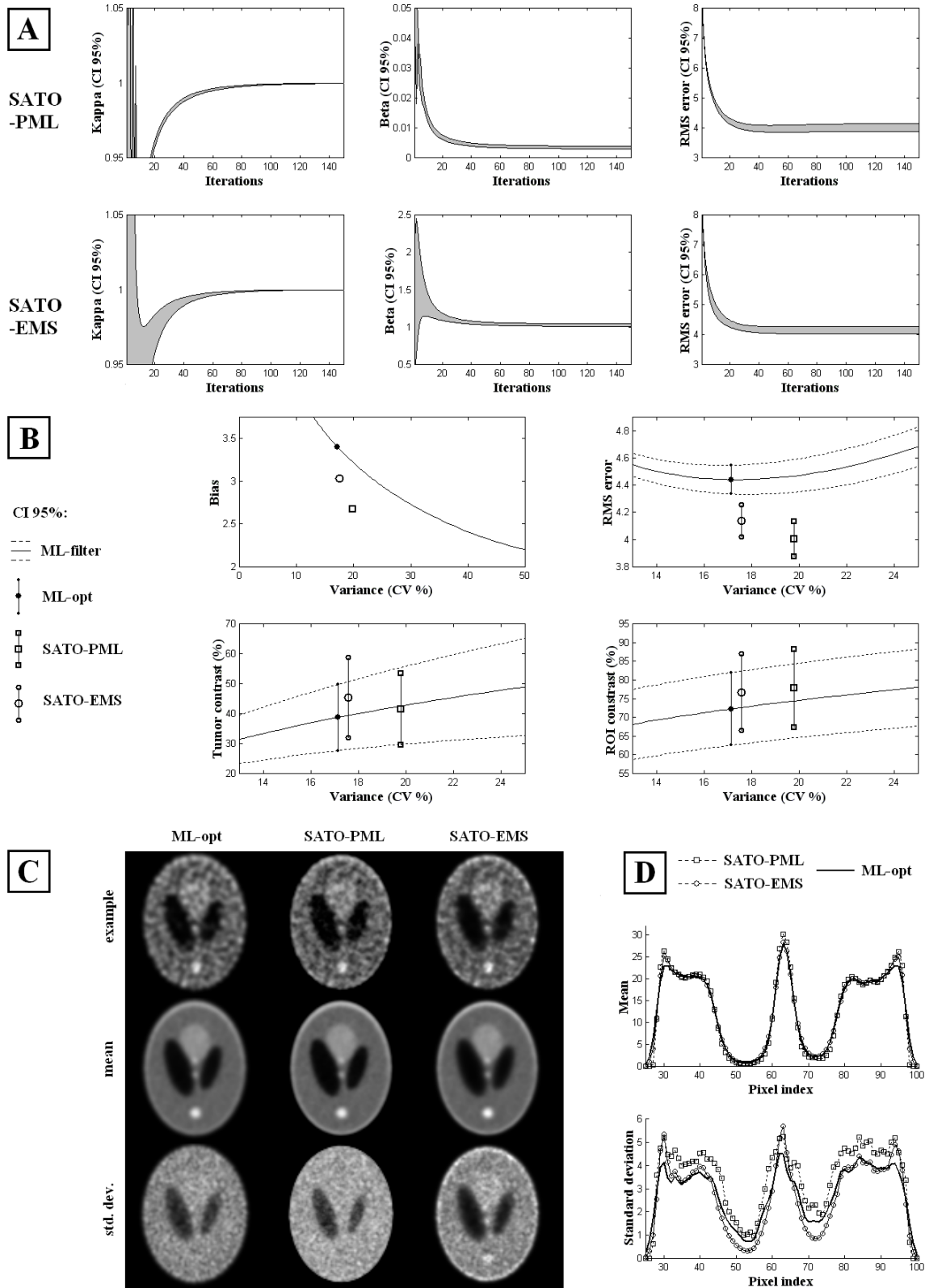


Figure 5. Results for the Shepp-Logan phantom at low count rate. **A:** convergence. Evolution with the iteration of (from left to right): scaling factor κ , tuning parameter β , and RMS error between the estimate and the exact object (top row: SATO-PML; bottom row: SATO-EMS). **B:** tradeoff curves. Top left to bottom right: bias-variance, RMS error-variance, and contrast-variance (tumor then ROI). **C:** reconstructed images. Left to right: ML-opt, SATO-PML and SATO-EMS. Top to bottom: example, mean, and standard deviation. **D:** profiles through the lines displayed on Figure 2. Left: mean. Right: standard deviation.

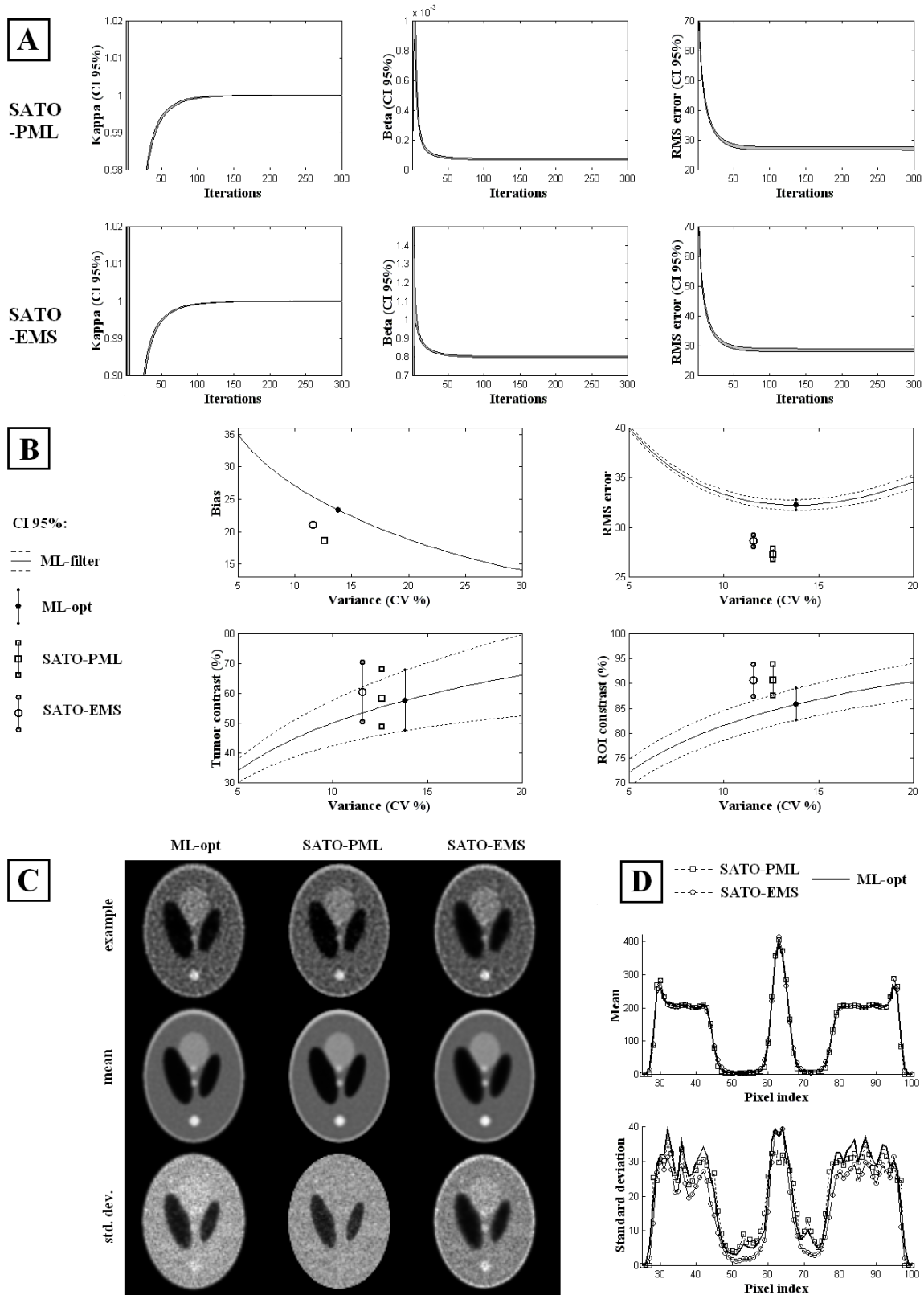


Figure 6. Results for the Shepp-Logan phantom at high count rate. **A**: convergence. Evolution with the iteration of (from left to right): scaling factor κ , tuning parameter β , and RMS error between the estimate and the exact object (top row: SATO-PML; bottom row: SATO-EMS). **B**: tradeoff curves. Top left to bottom right: bias-variance, RMS error-variance, and contrast-variance (tumor then ROI). **C**: reconstructed images. Left to right: ML-opt, SATO-PML and SATO-EMS. Top to bottom: example, mean, and standard deviation. **D**: profiles through the lines displayed on Figure 2. Left: mean. Right: standard deviation.

TABLE II. RELATIVE PERFORMANCES OF SATO-PML WITH RESPECT TO ML-OPT. CELLS COLORED IN LIGHT GRAY AND DARK GRAY RESPECTIVELY REFER TO A SIGNIFICANT IMPROVEMENT AND WORSENING (SIGNIFICANCE THRESHOLD: 5%).

	Low count rate		High count rate	
	Hoffman	Shepp-Logan	Hoffman	Shepp-Logan
Bias	- 6 %	- 21.5 %	+ 0.5 %	- 20 %
Variance (CV)	- 8 %	+ 15.5 %	- 27.5 %	- 8.5 %
Mean RMS error	- 6 %	- 9.5 %	- 3.5 %	- 15.5 %
Mean tumor contrast	- 10 %	+ 7 %	- 11.5 %	+ 1 %
Mean ROI contrast	+ 7 %	+ 7.5 %	- 4 %	+ 5.5 %

TABLE III. RELATIVE PERFORMANCES OF SATO-EMS WITH RESPECT TO ML-OPT. CELLS COLORED IN LIGHT GRAY AND DARK GRAY RESPECTIVELY REFER TO A SIGNIFICANT IMPROVEMENT AND WORSENING (SIGNIFICANCE THRESHOLD: 5%).

	Low count rate		High count rate	
	Hoffman	Shepp-Logan	Hoffman	Shepp-Logan
Bias	+ 0.5 %	- 11 %	+ 7 %	- 9.5 %
Variance (CV)	- 11 %	+ 3 %	- 25 %	- 16 %
Mean RMS error	- 2 %	- 7 %	+ 1 %	- 11 %
Mean tumor contrast	+ 3 %	+ 17.5 %	- 2.5 %	+ 5 %
Mean ROI contrast	+ 1.5 %	+ 6 %	- 3 %	+ 5.5 %

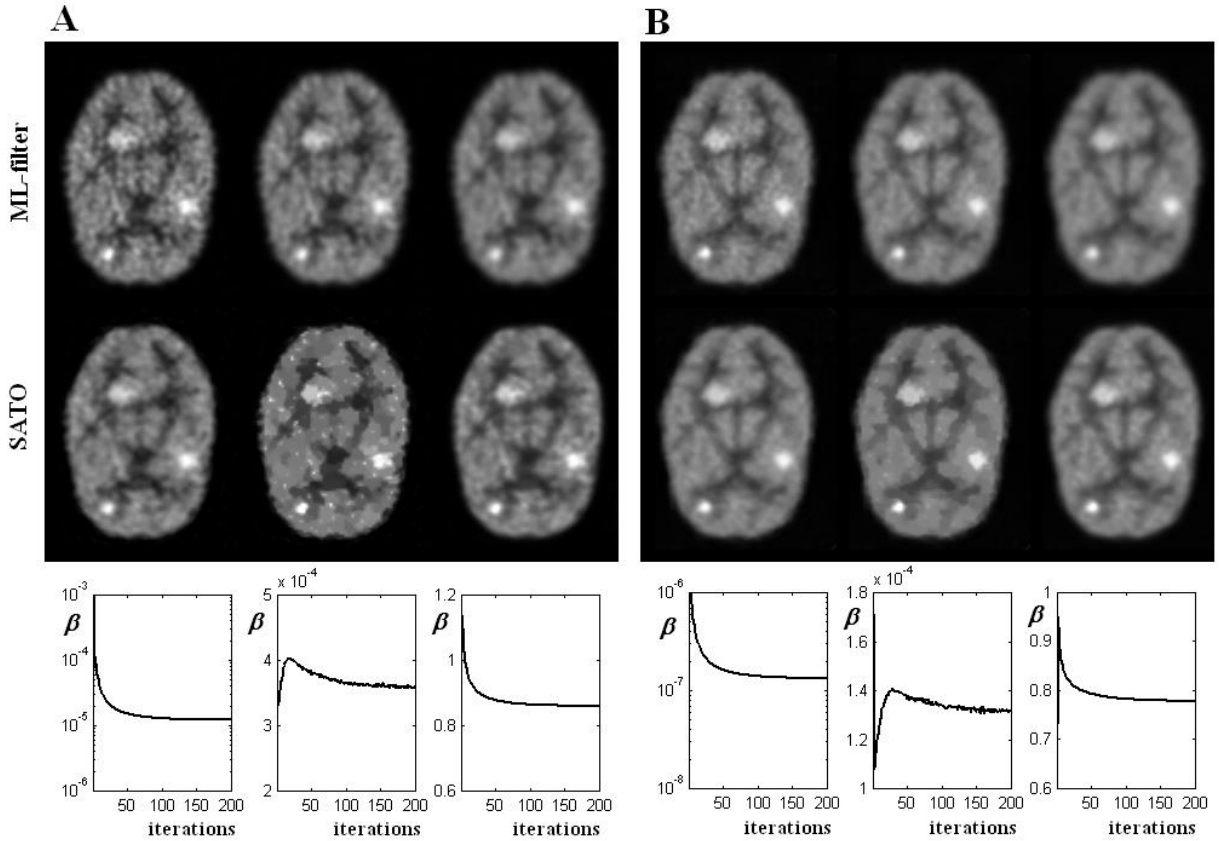


Figure 7. GATE Hoffman phantom study. A: 1.3 million count sinogram. B: 10 million count sinogram. *Top images*: reconstruction using MLEM with Gaussian post-filtering of FWHM 2.5, 3, and 3.5 pixels (A) and 2, 2.5, and 3 pixels (B). *Bottom images*, from left to right: reconstruction using SATO-PML with the quadratic prior, SATO-PML with the median prior, and SATO-EMS. *Curves*, from left to right: convergence of the tuning parameter for SATO-PML with the quadratic prior, SATO-PML with the median prior, and SATO-EMS.

DISCUSSION

The following discussion regarding the numerical experiments proceeds from the observation of Figures 3 to 6, as well as Tables II and III. It appears from Figures 3 to 6 (subplots A) that our two methods are globally convergent since the scaling factor κ reaches the optimal value of 1 (corresponding to the SATO criterion) during the iterative process. Whatever the random initialization of β , the plateau is reached after 100 to 150 iterations depending on the data statistics, the convergence rate being rather independent of the studied object and the initial value of β . The convergence of κ toward 1 implies the stabilization of the tuning parameter at its optimal value, which is weakly dependent on the noise realization according to the relatively small amplitude of the 95% confidence interval for β at convergence. Furthermore, our results confirm that the convergence of β ensures the stabilization of the RMS error which reaches a plateau (and does not further explode as in MLEM reconstruction). The stabilization of the RMS error and its small dispersion at convergence clearly demonstrate the regularized and robust nature of the two SATO algorithms. The RMS error at convergence is quite similar for the two algorithms, SATO-PML providing slightly lower convergence errors than SATO-EMS whatever the experiment. The summary provided in Tables II and III shows that from a global point of view SATO-PML and SATO-EMS allow a significant improvement in the image quality compared with ML-opt, be it in terms of bias, variance, RMS error, or contrast. The significant worsening appears to be scarce and isolated (no significant worsening exceeds 10 to 15 %) and is clearly balanced by significant improvements regarding the other figures of merit. The bias-variance tradeoff curves displayed on Figures 3 to 6 (subplot B) show that the points corresponding to SATO-PML and SATO-EMS are always located below the tradeoff curve of ML-filter, which means that the two proposed methods allow a reduction in either the bias or the variance compared with ML-filter in general and ML-opt in particular. The study of the variance-RMS error and variance-contrast plots reveals that the results of SATO-PML and SATO-EMS in terms of RMS error and contrast are always equivalent or significantly better than those of ML-filter for a similar variance level. The dispersion of the figures of merit (characterized using their 95% confidence interval) is highly similar for the three methods, thus demonstrating an equivalent robustness.

When the images are examined in detail (Figures 3 to 6, subplots C), it appears that SATO-PML provides images with a slightly noisier texture but moderately more accurate compared with SATO-EMS. This is also apparent on the mean line profiles (Figures 3 to 6, subplots D) where SATO-PML provides a better recovery of the object silhouette, especially at low count rates. According to the standard deviation images and profiles, the local variance of the SATO-PML estimate is almost uniform and globally similar to that of the ML-opt estimate. As for the SATO-EMS images, they prove to be significantly less noisy than the ML-opt images while preserving equivalent resolution characteristics as appreciable on the mean profiles. Their local variance is dependent on the local mean (the high activity regions having higher variance than the low activity ones), while remaining globally similar or lower than that of the ML-opt images.

The properties of the SATO-PML and SATO-EMS estimates discussed above on the basis of the Monte-Carlo experiments are fairly well reflected through the GATE phantom study. The three ML-filter estimates proposed in Figure 7 for each count level give an insight into which would probably be the optimal one (ML-opt, top middle image), roughly corresponding to a kernel FWHM of 3 pixels (A) and 2.5 pixel (B). As expected, the SATO-PML reconstructions (bottom left images) and the SATO-EMS reconstructions (bottom right images) show resolution properties qualitatively similar to those of ML-opt. The SATO-PML estimates based on a median prior (bottom middle images) are there to attest that our statistic-algebraic optimization criterion can virtually be applied to any kind of regularized iterative scheme using a tuning parameter to modulate the strength of the penalization. Indeed, the SATO-PML estimates obtained using a median prior exhibit the classical image texture of a reconstruction using an

edge-preserving feature, their noise and resolution characteristics remaining very close to those of the ML-opt estimates. The convergence curves displayed at the bottom show that the tuning parameter effectively converges when working with actual data. The middle curves corresponding to SATO-PML with a median prior exhibit a slight instability of β during its convergence toward its optimal value. This behavior is likely due to the form of the prior (equation 24) whose first derivative is non-continuous, hence leading to abrupt modifications in the direction of the correction $\delta^{\beta,n}$ while the iterations proceed. The experiment using the GATE phantom was also intended to demonstrate the feasibility of our optimization technique when working with actual data associated with an imperfectly modeled system matrix, unlike the Gibbs parameter selection method proposed by Hebert and Leahy in [23] and based on an adequacy criterion between the recorded projections and the estimate projections. Indeed, as the authors point out in their discussion, significant modeling errors in the system matrix are susceptible to alter the validity of their criterion and induce the convergence of the PML estimate toward the noisy ML estimate (see [23] section VI for details).

Before concluding, one word has to be said regarding the differential performances of the two SATO methods. SATO-PML performs better in terms of bias reduction, leading to noisier but more accurate images, while SATO-EMS performs better in terms of variance reduction, leading to smoother images. As further discussed in the appendix, this is likely due to the manner in which the two algorithms evaluate their correction $\delta^{\beta,n}$ (equations (31) and (35)). Equation (35) shows that the EMS correction is to some extent rather homogenous over the image and proportional to the pixel non-uniformity. This explains the tendency of SATO-EMS to smooth the image estimates in a homogenous way. Regarding PML, due to the quadratic form of the prior, the correction detailed in equation (31) shows a clear non uniformity as it is additionally proportional to the pixel value, hence leading to a stronger correction in the high activity regions. The SATO-PML image estimates consequently exhibit better edge restoration at the cost of a slightly lower noise reduction.

CONCLUSIONS

In this paper, we have proposed a simple statistical description of the noise propagation at a given iteration when exploiting ML-related iterative reconstruction in emission tomography. The two penalized methods examined (PML and EMS) proved to bring corrections for which an approximate algebraic criterion can naturally be exploited in order to iteratively optimize their tuning parameter β . We have proposed a general recipe for the practical implementation of our statistic-algebraic tuning optimization (SATO), with a particular application for PML (with a quadratic penalty) and EMS, yielding the SATO-PML and SATO-EMS algorithms. We have evaluated these two algorithms using Monte-Carlo simulations involving two numerical phantoms (Hoffman and Shepp-Logan) and two noise levels. The proposed methods were compared with classical MLEM reconstruction followed by optimal Gaussian post-filtering (ML-opt). Our two methods proved to be convergent in terms of β and to enable the stabilization of the RMS error. Their robustness was attested by the low dispersion of the tuning parameter and RMS error at convergence. They produced estimates with bias, variance and contrast properties at least as good as those of the ML-opt estimate. From a qualitative point of view, the generated images were fairly similar, the SATO-PML images showing moderately better resolution at the price of a tiny noise increment, and the SATO-EMS images being globally smoother without significant loss of resolution. Finally, the proposed methods and the reference one were used for the reconstruction of a Hoffman phantom whose projection data were produced by the GATE platform using the specifications of the GEMINI GXL PET scanner and two count levels. The images obtained by SATO-PML and SATO-EMS were visually very close to the “qualitatively” determined ML-opt image.

Reconstructions using SATO-PML and a median prior have also been proposed in order to provide insight into the potential scope of our optimization technique.

APPENDIX

A. Collinearity

Concerning PML, the correction for a given pixel j is:

$$\delta_j^{\beta,n} = \frac{1}{1 + \beta \Delta_j^{n-1}} [f_{ML}(\mathbf{x}^{n-1})]_j - [f_{ML}(\mathbf{x}^{n-1})]_j = -\beta \Delta_j^{n-1} \frac{1}{1 + \beta \Delta_j^{n-1}} [f_{ML}(\mathbf{x}^{n-1})]_j \quad (31)$$

where:

$$\Delta_j^{n-1} = \frac{\sum_{k \in N_j} w_k (\mathbf{x}_j^{n-1} - \mathbf{x}_k^{n-1})}{\sum_i A_{ij}} \quad (32)$$

is the (normalized) difference between the value of pixel j and the weighted mean value of the neighboring pixels at iteration $n-1$ (due to the OSL approach). Noting that $1 + \beta \Delta_j^{n-1} > 0 \forall j$ in order to avoid the cancellation of the denominator in the PML updating function, we have:

$$\beta \ll \min_k \left\{ \frac{1}{|\Delta_k^{n-1}|}, \Delta_k^{n-1} < 0 \right\} \quad (33)$$

The sign of Δ_j^n being randomly distributed, it follows that for any pixel j :

$$\beta |\Delta_j^{n-1}| \ll 1 \quad (34)$$

The PML correction thus writes as:

$$\delta_j^{\beta,n} = -\beta \Delta_j^{n-1} u_{\pm}(\beta |\Delta_j^{n-1}|) [f_{ML}(\mathbf{x}^{n-1})]_j \quad (35)$$

with:

$$u_{\pm}(x) = \frac{1}{1 \pm x} \quad (36)$$

where u_+ is chosen if $\Delta_j^{n-1} > 0$ and u_- is chosen if $\Delta_j^{n-1} < 0$. As shown on Figure 8, the two functions $u_{\pm}(x)$ remain almost constant in the practical range of variation of β (i.e., for $x \ll 1$) and are fairly well approximated as $u_{\pm}(x) \approx 1$. Consequently, the PML correction can be approximated as:

$$\delta_j^{\beta,n} \approx -\beta \Delta_j^{n-1} [f_{ML}(\mathbf{x}^{n-1})]_j \quad (37)$$

hence demonstrating its approximate collinear character.

Regarding EMS, the correction writes for a given pixel j :

$$\delta_j^{\beta,n} = [f_{ML}(\mathbf{x}^{n-1}) * G_{\beta} - f_{ML}(\mathbf{x}^{n-1})]_j = [f_{ML}(\mathbf{x}^{n-1}) * (G_{\beta} - I)]_j \quad (38)$$

with I the identity kernel function. Denoting G_β^* the normalized kernel such that $G_\beta^*(0,0) = 0$ and $G_\beta^*(m,n) \propto G_\beta(m,n)$ for $(m,n) \neq (0,0)$ yields:

$$G_\beta = (1 - \omega^\beta) I + \omega^\beta G_\beta^* \quad (39)$$

where $\omega^\beta = 1 - G_\beta(0,0)$ is monotonic in β . Basic manipulation leads to:

$$G_\beta = I - \omega^\beta (I - G_\beta^*) \quad (40)$$

and hence:

$$\delta_j^{\beta,n} = -\omega^\beta [f_{ML}(\mathbf{x}_j^{n-1}) * (I - G_\beta^*)]_j = -\omega^\beta \Delta_j^{\beta,n} \quad (41)$$

where $\Delta_j^{\beta,n} = [f_{ML}(\mathbf{x}_j^{n-1}) * (I - G_\beta^*)]_j$ has the same interpretation as Δ_j^{n-1} in the PML case and can be seen as the difference between the value of pixel j and the weighted mean value of the neighboring pixels. In contrast to equation (32), $\Delta_j^{\beta,n}$ is here computed at iteration n (previously at iteration $n-1$) on the basis of the ML update $f_{ML}(\mathbf{x}^{n-1})$ (previously on the basis of \mathbf{x}^{n-1}). Provided that $\Delta_j^{\beta,n}$ is sufficiently stable when β varies (which is the case when using a progressive Gaussian filtering), equation (41) ensures that the EMS correction is approximately collinear since for any (β_1, β_2) :

$$\delta_j^{\beta_1,n} \approx \frac{\omega^{\beta_1}}{\omega^{\beta_2}} \delta_j^{\beta_2,n} \quad (42)$$

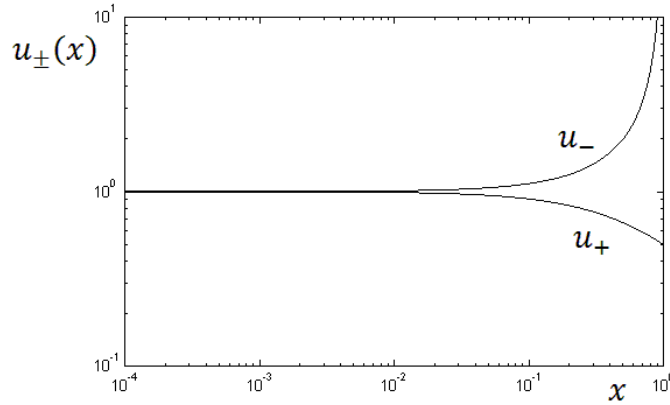


Figure 8. Plot of the functions $u_\pm(x)$ for $x \in [10^{-4} ; 1]$. Note the log-log scale.

B. Update formulas for β

Regarding SATO-PML, as the correction $\delta^{\beta,n}$ is in first approximation proportional to β (see eq. 31), it is natural to update β by simply multiplying it by the scaling factor $\kappa^{\beta,n}$, hence equation (25). As for SATO-EMS, equation (41) tells us that $\delta^{\beta,n}$ is in first approximation proportional to ω^β . Noting σ the standard deviation of the Gaussian kernel, one has:

$$\omega^\beta = 1 - G_\beta(0,0) = 1 - \text{erf}^2\left(\frac{1}{\sigma\sqrt{2}}\right) \quad (43)$$

One can easily check that the function $1 - \operatorname{erf}^2\left(\frac{1}{\sigma\sqrt{2}}\right)$ is very well approximated using $\exp\left(-\frac{1}{2\sigma^2}\right)$. Introducing $\beta = 2\sqrt{2\ln(2)}\sigma$ it comes that:

$$\delta^{\beta,n} \propto \exp(-\ln(2)/\beta^2) \quad (44)$$

Then solving

$$\exp\left(-\frac{4\ln(2)}{\beta_{new}^2}\right) = \kappa^{\beta,n} \exp\left(-\frac{4\ln(2)}{\beta_{old}^2}\right) \quad (45)$$

yields equation (26).

REFERENCES

- [1] L. A. Shepp and Y. Vardi, "Maximum likelihood reconstruction for emission tomography". *IEEE Trans. Med. Imag.*, vol. 1, pp. 113-122, 1982.
- [2] K. Lange and R. Carson, "EM reconstruction algorithms for emission and transmission tomography". *J. Comput. Assist. Tomogr.*, vol. 2, pp. 306-316, 1984.
- [3] M. I. Miller, D. L. Snyder, and T. R. Miller, "Maximum likelihood reconstruction for single-photon emission computed tomography". *IEEE Trans. Nucl. Sci.*, vol. 32, pp. 769-778, 1985.
- [4] D. W. Wilson and B. W. Tsui, "Noise properties of filtered backprojection and ML-EM reconstructed emission tomography images". *IEEE Trans. Nucl. Sc.*, vol. 40-4, pp. 1198-1203, 1993.
- [5] S. C. Liew, B. H. Hasegawat, J. K. Brown and T. F. Lang, "Noise propagation in SPECT images reconstructed using an iterative maximum-likelihood algorithm". *Phys. Med. Biol.*, vol. 38, pp. 1713-1726, 1993.
- [6] H. H. Barrett, D. W. Wilson and B. M. Tsui, "Noise properties of the EM algorithm: I. theory". *Phys. Med. Biol.*, vol. 39, pp. 833-846, 1994.
- [7] H. H. Barrett, D. W. Wilson and B. M. Tsui, "Noise properties of the EM algorithm: II. Monte-Carlo Simulations". *Phys. Med. Biol.*, vol. 39, pp. 847-872, 1994.
- [8] J. Nuyts, "On estimating the variance of smoothed MLEM images". *IEEE Trans. Nucl. Sc.*, vol. 40-3, pp. 714-721, 2002.
- [9] F. J. Beekman, E. T. Slijpen, and W. J. Niessen, "Selection of task-dependent diffusion filters for the post-processing of SPECT images". *Phys. Med. Biol.*, vol. 43, pp. 1713-1730, 1998.
- [10] E. Veklerov and J. Llacer, "Stopping rule for the MLE algorithm based on statistical hypothesis testing". *IEEE Trans. Med. Imag.*, vol. 6, pp. 313-19, 1987.
- [11] J. Llacer and E. Veklerov, "Feasible images and practical stopping rules for iterative algorithms in emission tomography". *IEEE Trans. Med. Imag.*, vol. 8, pp. 186-93, 1989.

- [12] T. J. Hebert, "Statistical stopping criteria for iterative maximum likelihood reconstruction of emission images" *Phys. Med. Biol.*, vol 35, pp. 1221-32, 1990.
- [13] K. J. Coakley and J. Llaser, "Use of cross-validation as a stopping rule in emission tomography image reconstruction". *Proc. SPIE Vol. 1443*, p. 226-233. Schneider Ed, 1991.
- [14] K.J. Coakley, "A cross-validation procedure for stopping the EM algorithm and deconvolution of neutron depth profile spectra". *IEEE Trans. Nucl. Sci.*, vol. 38, pp. 9-15, 1991.
- [15] B. W. Silverman et al, "A smoothed EM approach to indirect estimation problems, with particular reference to stereology and emission tomography". *Journal of the Royal Statistical Society, Series B*, vol 52-2, pp. 271-324, 1990.
- [16] E. T. P. Slijpen and F. J. Beekman, "Comparison of post-filtering and filtering between iterations for SPECT reconstruction". *IEEE Trans. Nucl. Sci.*, vol. 46, pp. 2233-2238, 1999.
- [17] S. Mustafovic, K. Thielemans, D. Hogg *et al.*, "Object dependency of resolution and convergence rate in OSEM with filtering" *IEEE Nuclear Science Symposium Conference Record*, vol. 3, pp 1786-1790, 2001.
- [18] S. Geman and D. E. McClure, "Bayesian image analysis: an application to single photon emission tomography," *Proc. Statist. Comput. Sect., Amer. Statist. Soc.*, Washington, DC, 1985, pp 12-18.
- [19] E. Levitan and G. T. Herman, "A maximum a posteriori probability expectation maximization algorithm for image reconstruction in emission tomography". *IEEE Trans. Med. Imag.*, vol. MI-6, pp. 185-192, 1987.
- [20] T. Hebert and R. Leahy, "A generalized EM algorithm for 3-D Bayesian reconstruction from Poisson data using Gibbs priors". *IEEE Trans. Med. Imag.*, vol. 8-2, pp. 194-202, 1989.
- [21] Z. Liang, R. Jaszczak and K. Greer, "On Bayesian image reconstruction from projections: uniform and non uniform a priori source information". *IEEE Trans. Med. Imag.*, vol. 8, pp. 227-235, 1989.
- [22] P. J. Green, "Bayesian reconstructions from emission tomography data using a modified EM algorithm". *IEEE Trans Med Imag.*, vol. 9, pp. 84-93, 1990.
- [23] T. J. Hebert and R. Leahy, "Statistic-based PML image reconstruction from Poisson data using Gibbs priors". *IEEE Trans. Signal. Proc.*, vol. 40-9, pp. 2290-2302, 1992.
- [24] E. U. Mumcuoglu, R. M. Leahy, and S. R. Cherry, "Bayesian reconstruction of PET images: methodology and performance analysis". *Phys. Med. Biol.*, vol. 41, pp. 1777-1807, 1996.
- [25] A. R. de Pierro, "On the convergence of an EM-type algorithm for penalized likelihood estimation in emission tomography". *IEEE Trans. Med. Imag.*, vol. 14, pp. 762-765, 1995.

- [26] I. T. Hsiao, A. Rangarajan, and G. Gindi, "A new convex edge-preserving median prior with applications to tomography". *IEEE Trans. Med. Imag.*, vol. 22-5, pp. 580-585, 2003.
- [27] J. Nuyts, D. Beque, P. Dupont, and L. Mortelmans, "A concave prior penalizing relative differences for maximum-a-posteriori reconstruction in emission tomography". *IEEE Trans. Nucl. Sci.*, vol. 49, pp. 56-60, 2002.
- [28] S. Ahn and J. A. Fessler, "Globally convergent image reconstruction for emission tomography using relaxed ordered subsets algorithms". *IEEE Trans. Med. Imag.* Vol. 22, pp. 613-626, 2003.
- [29] K. Lange, "Convergence of EM image reconstruction algorithms with Gibbs smoothing," *IEEE Trans. Med. Imag.*, vol. 9-4, pp. 439-446, 1990.
- [30] J. A. Fessler, "Mean and variance of implicitly defined biased estimators (such as penalized maximum likelihood): applications to tomography". *IEEE Trans. Image. Proc.*, vol. 5-3, pp. 493-506, 1996.
- [31] J. A. Fessler and W. L. Rogers, "Spatial resolution properties of penalized likelihood image reconstruction: space-invariant tomographs". *IEEE Trans. Imag. Proc.*, vol. 5, pp. 1346-1358, 1996.
- [32] W. Wang and G. Gindi, "Noise analysis of MAP-EM algorithms for emission tomography". *Phys. Med. Biol.*, vol. 42-11, pp. 2215-2232, 1997.
- [33] J. W. Stayman and J. A. Fessler, "Regularization for uniform spatial resolution properties in penalized-likelihood image reconstruction". *IEEE Trans Med Imag.*, vol. 19(6), pp. 601-15, 2000.
- [34] J. Qi and R. Leahy, "Resolution and noise properties of MAP reconstructions in fully 3D PET". *IEEE Trans Med Imag.*, vol. 19(5), pp. 493-506, 2000
- [35] D. E. Knuth, "Seminumerical algorithms. The art of computer programming, Volume 2". *Addison Wesley*, 1969.
- [36] S. Jan *et al* (for the OpenGATE collaboration), "GATE: a simulation toolkit for PET and SPECT". *Phys. Med. Biol.*, vol. 49, pp. 4543-61, 2004.
- [37] F. Lamare, A. Turzo, Y. Bizais, C. Cheze Le Rest, and D. Visvikis, "Validation of a Monte Carlo simulation of the Philips Allegro/GEMINI PET systems using GATE". *Phys. Med. Biol.*, vol. 51, pp. 943-62, 2006.
- [38] D. L. Bailey and S. R. Meikle, "A convolution-subtraction scatter correction method for 3D PET", *Phys. Med. Biol.*, vol. 39, pp. 411-24, 1994.
- [39] D. J. Kadrmas, "LOR-OSEM: statistical PET reconstruction from raw line-of-response histograms", *Phys. Med. Biol.*, vol. 49, pp. 4731-44, 2004.