Least-squares dual characterization for ROI assessment in emission tomography

2013 Phys. Med. Biol. 58 4175

(http://iopscience.iop.org/0031-9155/58/12/4175)

IOP Publishing Physics in Medicine and Biology

Phys. Med. Biol. **58** (2013) 4175–4194 doi:10.1088/0031-9155/58/12/4175

Least-squares dual characterization for ROI assessment in emission tomography

F Ben Bouallègue[1,4], J F Crouzet[2], A Dubois[3], I Buvat[3] and D Mariano-Goulart[1]

[1] Nuclear Medicine Department, Lapeyronie University Hospital, 371 avenue du Doyen Gaston Giraud, F-34295 Montpellier Cedex 5, France
[2] Department of Mathematics, Montpellier Science University, F-34095 Montpellier, France
[3] IMNC—UMR 8165 CNRS, Paris 7 and Paris 11 Universities, 15 rue Georges Clémenceau, F-91406 Orsay Cedex, France

E-mail: faybenb@hotmail.com

Received 15 February 2013, in final form 24 March 2013
Published 28 May 2013
Online at stacks.iop.org/PMB/58/4175

Abstract

Our aim is to describe an original method for estimating the statistical properties of regions of interest (ROIs) in emission tomography. Drawn upon the works of Louis on the approximate inverse, we propose a dual formulation of the ROI estimation problem to derive the ROI activity and variance directly from the measured data without any image reconstruction. The method requires the definition of an ROI characteristic function that can be extracted from a co-registered morphological image. This characteristic function can be smoothed to optimize the resolution-variance tradeoff. An iterative procedure is detailed for the solution of the dual problem in the least-squares sense (least-squares dual (LSD) characterization), and a linear extrapolation scheme is described to compensate for sampling partial volume effect and reduce the estimation bias (LSD-ex). LSD and LSD-ex are compared with classical ROI estimation using pixel summation after image reconstruction and with Huesman's method. For this comparison, we used Monte Carlo simulations (GATE simulation tool) of 2D PET data of a Hoffman brain phantom containing three small uniform high-contrast ROIs and a large non-uniform low-contrast ROI. Our results show that the performances of LSD characterization are at least as good as those of the classical methods in terms of root mean square (RMS) error. For the three small tumor regions, LSD-ex allows a reduction in the estimation bias by up to 14%, resulting in a reduction in the RMS error of up to 8.5%, compared with the optimal classical estimation. For the large non-specific region, LSD using appropriate smoothing could intuitively and efficiently handle the resolution-variance tradeoff.

[4] Author to whom any correspondence should be addressed.

1. Introduction

Emission tomography addresses the problem of reconstructing the activity map of a radiotracer from a series of projections, each projection corresponding to the number of photons received by the tomograph in a given direction. These projections can be seen as a set of estimates of the line integrals of the activity through the field of view of the imaging device, so that the theoretical background and numerical tools of tomographic reconstruction can be used to produce an estimate of the studied object. However, due to the random nature of radioactive decay and to the detection process, the measured projections are affected by Poisson noise which propagates into the reconstructed object. When a quantitative analysis is necessary, it is useful to estimate the statistical uncertainty characterizing the reconstructed activity map. This is especially true when one wants to compare the total activity inside one or several regions of interest (ROIs) with other measurements performed in different regions or at a different time. The problem of estimating image variance in emission tomography has been extensively studied. Analytical and numerical approximations have been proposed, many of them focusing on a specific reconstruction algorithm. Both filtered back-projection and maximum-likelihood algorithms have been employed since the late 1970s in the field of emission tomography and many formulations have been proposed to describe the propagation of the uncertainty from the measured projections to the reconstructed images (Huesman 1977, Budinger *et al* 1977, Alpert *et al* 1982, Moore *et al* 1988, Gillen 1992, Wilson and Tsui 1993, Liew *et al* 1993, Kim *et al* 1993a, 1993b, Barrett *et al* 1994a, 1994b, Pan and Metz 1995, Wang and Gindi 1997, Nuyts 2002, Mariano-Goulart *et al* 2003, Nickerson *et al* 2003, Fessler 1996). Original methods based on bootstrapping techniques (Haynor and Woods 1989, Buvat 2002, Lartizien *et al* 2010) or intervalist prediction (Rico *et al* 2009, Strauss *et al* 2009) have recently been proposed. The computation of the whole reconstructed image covariance matrix however, remains a computational intensive task due to the huge dimensions of the system matrix in typical acquisition settings. The difficult handling of the subsequent numerical errors and heavy computational burden make these calculations unsuitable for clinical routine. In addition, most image interpretation tasks require region of interest assessment rather than pixel-wise variance estimation. Various attempts have been made to design ROI-based quantification methods in emission tomography (Huesman 1984, Carson 1986, Formiconi 1993, Carson *et al* 1993). However, the proposed techniques either rely on the analytical formulation of the tomographic inverse problem (i.e., when using the filtered back-projection algorithm) that cannot support an accurate modeling of the emission–detection process, or they assume that the studied ROIs have uniform activity. A recent work relying on a local segmentation technique yielded promising results with μSPECT data (Moore *et al* 2012), but the method also assumes that the activity inside each segmented tissue is uniform. Elsewhere, the importance of the ROI definition has already been pointed out (Wells 1999).

Drawn upon the works of Louis on the approximate inverse (Louis and Maass 1990, Louis 1996, Louis *et al* 2008), the present paper describes an original method for estimating the statistical properties of an ROI. The method starts with a dual formulation of the ROI estimation problem to derive the ROI activity and variance directly from the measured data without any image reconstruction. It requires the definition of the ROI characteristic function, which can be obtained from a co-registered morphological image. This characteristic function can be smoothed to adjust the resolution-variance tradeoff. An iterative procedure is proposed to solve the dual problem in the least-squares sense and a linear extrapolation scheme is described to reduce the estimation bias caused by sampling partial volume effect. Dual characterization is compared with classical ROI estimation using pixel summation and with

Huesman's reference method (Huesman 1984) through Monte Carlo simulations involving 2D PET data of a Hoffman brain phantom.

2. Materials and methods

2.1. Direct characterization

Let $\mathbf{f} \in \mathbb{R}^{\mathrm{N}}$ denote an original activity map defined on a 2D Cartesian grid whose pixels are indexed with the letter $j (j = 1, \ldots, N)$ and R a subset of that grid called ROI. The characteristic function of R is noted $\mathbf{k} \in \mathbb{R}^{\mathrm{N}}$ and is such that $\mathbf{k}_j = 1$ if $j \in R$ and $\mathbf{k}_j = 0$ otherwise. The total activity inside ROI R is then $\mathrm{F} = \mathbf{k}^{\mathrm{T}}\mathbf{f}$ where T stands for the matrix transpose. In emission tomography, the original activity map is estimated using a series of projections noted $\mathbf{p} \in \mathbb{R}^{\mathrm{M}}$, where M is the number of projection bins. This estimation requires the modeling of the emission–detection process through a system matrix $\mathbf{M} \in \mathbb{R}^{\mathrm{N}\times\mathrm{M}}$ whose element $\mathbf{M}_{ij}$ models the probability for a photon (in SPECT) or photon pair (in PET) emitted from pixel j to be measured in bin i. An estimate $\boldsymbol{\varphi}$ of $\mathbf{f}$ is obtained by solving the inverse problem $\mathbf{M}\boldsymbol{\varphi} = \mathbf{p}$. Starting from this estimate, an estimate Φ of the ROI activity is deduced using $\Phi = \mathbf{k}^{\mathrm{T}}\boldsymbol{\varphi}$. The set of equations:

$$\mathbf{M}\boldsymbol{\varphi} = \mathbf{p} \tag{1a}$$

$$\Phi = \mathbf{k}^{\mathrm{T}}\boldsymbol{\varphi} \tag{1b}$$

will be referred to as the direct problem and the resulting solution as the direct characterization of the studied ROI. In practice, due to the physical properties of the radioactive decay and detection process, projection $\mathbf{p}$ is a random variable which can be written as:

$$\mathbf{p} = \bar{\mathbf{p}} + \boldsymbol{\nu} \tag{2}$$

where $\bar{\mathbf{p}}$ stands for the ideal noise-free projection such that $\bar{\mathbf{p}} = \mathbf{M}\mathbf{f}$, and $\boldsymbol{\nu}$ is a Poisson noise with $E(\boldsymbol{\nu}) = 0$ and $\mathrm{Var}(\boldsymbol{\nu}) = \mathrm{Diag}(\bar{\mathbf{p}})$. While $\mathbf{f}$ and $\bar{\mathbf{p}}$ remain unknown for a given reconstruction problem, it is commonly accepted that $\mathrm{Var}(\boldsymbol{\nu}) = \mathrm{Var}(\mathbf{p}) \approx \mathrm{Diag}(\mathbf{p})$ represents an acceptable approximation. An exact solution to the direct problem (1*a*) does not exist in general due to the inconsistency resulting from the Poisson noise. The estimated solution $\hat{\boldsymbol{\varphi}}$ can be written as:

$$\widehat{\boldsymbol{\varphi}} = \mathbf{H}\mathbf{f} + \mathbf{w} \tag{3}$$

where $\mathbf{H} \in \mathbb{R}^{\mathrm{N}\times\mathrm{N}}$ is a convolution operator that is intended to be as close as possible to the identity matrix $\mathbf{I}$ but departs from it due to the non-injectivity of the tomographic problem and the modeling errors affecting the system matrix, and $\mathbf{w} \in \mathbb{R}^{\mathrm{N}}$ is the image noise resulting from the propagation of the statistical noise from the projections to the reconstructed image. Using (1*b*), the corresponding estimation $\widehat{\Phi}$ of the ROI value is given by:

$$\widehat{\Phi} = \mathbf{k}^{\mathrm{T}}\widehat{\boldsymbol{\varphi}} = \mathbf{k}^{\mathrm{T}}\mathbf{H}\mathbf{f} + \mathbf{k}^{\mathrm{T}}\mathbf{w} \tag{4}$$

The first term of equation (4) determines the estimation bias through the resolution kernel $\mathbf{H}^{\mathrm{T}}\mathbf{k}$ ($\mathrm{bias} = |E(\widehat{\Phi}) - \mathrm{F}| = |\mathbf{k}^{\mathrm{T}}(\mathbf{H} - \mathbf{I})\mathbf{f}|$). The second term determines the variance of the estimate that can be computed as:

$$\mathrm{Var}(\widehat{\Phi}) = \mathbf{k}^{\mathrm{T}}\mathrm{Var}(\mathbf{w})\mathbf{k}. \tag{5}$$

As most modern techniques employed to solve (1*a*) are nonlinear (especially the algorithms based on likelihood maximization), simple and accurate expressions for $\mathbf{H}$ and $\mathbf{w}$ are generally not available and a statistical characterization of $\widehat{\Phi}$ (in terms of resolution/bias and variance) remains difficult to establish in a fast and effective way.

2.2. Huesman's method

In 1984, Huesman proposed a fast algorithm for the evaluation of ROIs and statistical uncertainty based on the analytical solution of the tomographic problem known as filtered back-projection (Huesman 1984). Huesman's ROI estimate can be written as:

$$\widehat{\Phi} = \mathbf{k}^{\mathrm{T}}\mathbf{BCAp} \tag{6}$$

where $\mathbf{A} \in \mathbb{R}^{\mathrm{M}\times\mathrm{M}}$ is a diagonal matrix accounting for the attenuation correction of the measured projections, $\mathbf{C} \in \mathbb{R}^{\mathrm{M}\times\mathrm{M}}$ is the convolution matrix mapping the unfiltered projections into the filtered ones using the ramp filter, and $\mathbf{B} \in \mathbb{R}^{\mathrm{N}\times\mathrm{M}}$ is the geometrical back-projection matrix. As the formulation is strictly linear, the statistical uncertainty affecting $\widehat{\Phi}$ can be approximated as:

$$\mathrm{Var}(\widehat{\Phi}) = \mathbf{k}^{\mathrm{T}}\mathbf{BCA}\mathrm{Var}(\mathbf{p})(\mathbf{BCA})^{\mathrm{T}}\mathbf{k} \approx \mathbf{k}^{\mathrm{T}}\mathbf{BCA}\,\mathrm{Diag}(\mathbf{p})(\mathbf{BCA})^{\mathrm{T}}\mathbf{k}. \tag{7}$$

Huesman's analytical method has two main drawbacks. First, the attenuation correction as expressed above is only available in PET and not in SPECT, leading to substantial inaccuracy in SPECT. Second, it suffers from the inability of filtered back-projection to force the non-negativity of the reconstructed image (unlike most of the algebraic reconstruction schemes which satisfy this constrain by using multiplicative iterative corrections). An extended formulation of Huesman's method has also been proposed (Muzic *et al* 1998) to compensate for partial volume effects.

2.3. Least-squares dual characterization

This sub-section is in part inspired from the works of Louis on the approximate inverse formulation for the solution of inverse problems (Louis and Maass 1990, Louis 1996, Louis *et al* 2008). The goal of the dual characterization is to express the ROI estimate as $\Phi = \boldsymbol{\delta}^{\mathrm{T}}\mathbf{p}$ where $\boldsymbol{\delta} \in \mathbb{R}^{\mathrm{M}}$ is called the dual characteristic function through which Φ can be computed in the projection space rather than in the image space as in (1*b*). From (1*a*) one deduces $\Phi = \boldsymbol{\delta}^{\mathrm{T}}\mathbf{p} = \boldsymbol{\delta}^{\mathrm{T}}\mathbf{M}\boldsymbol{\varphi}$, and from (1*b*) $\Phi = \boldsymbol{\delta}^{\mathrm{T}}\mathbf{p} = \mathbf{k}^{\mathrm{T}}\boldsymbol{\varphi}$. These two equations lead to $\mathbf{k}^{\mathrm{T}}\boldsymbol{\varphi} = \boldsymbol{\delta}^{\mathrm{T}}\mathbf{M}\boldsymbol{\varphi}$, which can be simplified into $\mathbf{k}^{\mathrm{T}} = \boldsymbol{\delta}^{\mathrm{T}}\mathbf{M}$, or equivalently $\mathbf{M}^{\mathrm{T}}\boldsymbol{\delta} = \mathbf{k}$. The set of equations:

$$\mathbf{M}^{\mathrm{T}}\boldsymbol{\delta} = \mathbf{k} \tag{8a}$$

$$\Phi = \boldsymbol{\delta}^{\mathrm{T}}\mathbf{p} \tag{8b}$$

will be referred to as the dual problem and the resulting solution as the dual characterization of the studied ROI. Here, equation (8*a*) is independent of the projection data hence $\boldsymbol{\delta}$ is a deterministic vector depending only on the system matrix and the studied ROI. The dual inverse problem (8*a*) is highly similar to the direct one (1*a*) since it involves the same system matrix $\mathbf{M}$, which is known to be ill-conditioned. As a result, an exact solution to the dual problem does not exist in general and the estimated solution will be denoted $\widehat{\boldsymbol{\delta}}$. The corresponding estimate $\widehat{\Phi}$ for the ROI value can be derived using (8*b*):

$$\widehat{\Phi} = \widehat{\boldsymbol{\delta}}^{\mathrm{T}}\mathbf{p} = \widehat{\boldsymbol{\delta}}^{\mathrm{T}}(\bar{\mathbf{p}} + \boldsymbol{\nu}) = \widehat{\boldsymbol{\delta}}^{\mathrm{T}}\mathbf{Mf} + \widehat{\boldsymbol{\delta}}^{\mathrm{T}}\boldsymbol{\nu} \tag{9}$$

The first term of equation (9) determines the estimation bias through the resolution kernel $\mathbf{M}^{\mathrm{T}}\widehat{\boldsymbol{\delta}}$ (bias $= |E(\widehat{\Phi}) - \mathrm{F}| = |(\widehat{\boldsymbol{\delta}}^{\mathrm{T}}\mathbf{M} - \mathbf{k}^{\mathrm{T}})\mathbf{f}|$). The second term determines the variance of the estimate:

$$\mathrm{Var}(\hat{\Phi}) = \widehat{\boldsymbol{\delta}}^{\mathrm{T}}\mathrm{Var}(\boldsymbol{\nu})\widehat{\boldsymbol{\delta}} \approx \widehat{\boldsymbol{\delta}}^{\mathrm{T}}\mathrm{Diag}(\mathbf{p})\widehat{\boldsymbol{\delta}} \tag{10}$$

Let us stress here the difference between direct and dual characterization. In both cases, the ROI estimates given by equations (4) and (9) can be written as:

$$\widehat{\Phi} = \mathbf{h}^{\mathrm{T}}\mathbf{f} + \omega \tag{11}$$

where $\mathbf{h} \in \mathbb{R}^{\mathrm{N}}$ is the resolution kernel and ω is the noise. Direct characterization of the ROI is performed using classical iterative nonlinear algorithms for which $\mathbf{h}$ and ω remain unknown (or at least not easily computable). On the contrary, the dual formulation makes it possible to estimate both $\mathbf{h} = \mathbf{M}^{\mathrm{T}}\widehat{\boldsymbol{\delta}}$ (within modeling errors affecting the system matrix) and $\omega = \widehat{\boldsymbol{\delta}}^{\mathrm{T}}\boldsymbol{\nu}$. The knowledge of the noise ω leads to a simple and robust expression for the ROI estimate variance as given in equation (10). The relative error e affecting the variance estimate:

$$e = \frac{\Delta\mathrm{Var}(\widehat{\Phi})}{\mathrm{Var}(\widehat{\Phi})} = \frac{\sqrt{\sum \widehat{\delta}_i^4 \bar{\mathbf{p}}_i}}{\sum \widehat{\delta}_i^2 \bar{\mathbf{p}}_i} \approx \frac{1}{\sqrt{qO(\bar{\mathbf{p}})}} \tag{12}$$

is expected to decrease with the count rate and with the number q of non-null components in $\widehat{\boldsymbol{\delta}}$ (i.e., the number of projections involved in the estimation of $\widehat{\Phi}$). For instance, basic computations yield $e = O(3)\%$[5] for $q = 100$ and $O(\bar{\mathbf{p}}) = 10$ (very tiny ROI at low count rate), and $e = O(0.3)\%$ for $q = 1000$ and $O(\bar{\mathbf{p}}) = 100$ (large ROI at high count rate). In addition, the knowledge of the resolution kernel $\mathbf{h} = \mathbf{M}^{\mathrm{T}}\widehat{\boldsymbol{\delta}}$ can help manage the bias-variance tradeoff as will be discussed thereafter. Since the solution of the dual problem may have negative components, (8*a*) cannot be solved like the direct problem using maximum-likelihood algorithms involving multiplicative corrections and will have to be solved with a descent method. Let us consider the following objective function:

$$J(\boldsymbol{\delta}) = \|\mathbf{M}^{\mathrm{T}}\boldsymbol{\delta} - \mathbf{k}\|^2 + \beta\|\boldsymbol{\delta}\|^2 \tag{13}$$

which minimal argument $\widehat{\boldsymbol{\delta}} = \underset{\boldsymbol{\delta}\in\mathbb{R}^{\mathrm{M}}}{\operatorname{argmin}}\{J(\boldsymbol{\delta})\}$ corresponds to the least-squares solution of (8*a*) with a Tikhonov regularization. The first term is a matching term between $\mathbf{M}^{\mathrm{T}}\widehat{\boldsymbol{\delta}}$ and $\mathbf{k}$ whose minimization results in the minimization of the estimation bias on $\widehat{\Phi}$. The second term is a penalty whose roughness is determined by the scalar parameter β and whose minimization results in the minimization of $\mathrm{Var}(\widehat{\Phi})$. Yet, these two terms are computed in different spaces and their relative order of magnitude is hard to evaluate. Although optimization techniques have been developed to tune the Tikhonov parameter β (Hansen and O'Leary 1993, Golub and von Matt 1997), a less time-consuming and more intuitive manner to handle the bias-variance tradeoff is to use the following objective function:

$$J(\boldsymbol{\delta}) = \left\|\mathbf{M}^{\mathrm{T}}\boldsymbol{\delta} - \breve{\mathbf{k}}_\alpha\right\|^2 \tag{14}$$

where $\breve{\mathbf{k}}_\alpha$ is a smoothed characteristic function of the ROI:

$$\breve{\mathbf{k}}_\alpha = \mathbf{k} * \mathbf{G}_\alpha \tag{15}$$

with $*$ standing for the 2D convolution operator, and $\mathbf{G}_\alpha$ for the 2D Gaussian kernel of full-width at half-maximum (FWHM) α pixels[6].When the regularization parameters (β and α) tend to zero, both equations (13) and (14) tend to the natural least-squares objective function $J(\boldsymbol{\delta}) = \|\mathbf{M}^{\mathrm{T}}\boldsymbol{\delta} - \mathbf{k}\|^2$. However, unlike the Tikhonov method, the regularization parameter α has a clear physical and numerical interpretation and can be used to tune the regularization roughness in terms of smoothing distance. The impact of the smoothing on the ROI characterization can be assessed in terms of resolution kernel. Indeed, the solution of the dual problem $\widehat{\boldsymbol{\delta}}$ is built so that $\mathbf{M}^{\mathrm{T}}\widehat{\boldsymbol{\delta}}$ is as close as possible to $\breve{\mathbf{k}}_\alpha$ and the smoothed

[5] Where O stands for 'order of magnitude'.

[6] Obviously, when working with 3D data, * stands for the 3D convolution operator and $\mathbf{G}_\alpha$ for a 3D Gaussian kernel.

characteristic function $\breve{\mathbf{k}}_\alpha$ may thus be viewed as the target resolution kernel for the ROI estimate. We propose to minimize the objective function using the following iterative procedure:

- Initialize $\widehat{\boldsymbol{\delta}}^0$ to:

$$\forall i, \begin{cases} \widehat{\delta}_i^0 = 1 \text{ if } \left\{\exists j \in R \mid \mathbf{M}_{ij} > 0\right\} \\ \widehat{\delta}_i^0 = 0 \text{ otherwise} \end{cases} \tag{16}$$

  then normalize it so that:

$$\sum_j (\mathbf{M}^{\mathrm{T}} \widehat{\boldsymbol{\delta}}^0)_j = \sum_j \breve{\mathbf{k}}_{\alpha,j}. \tag{17}$$

- At each iteration n, do:
- Compute the descent direction $\boldsymbol{\gamma}^n$ using an approximate Newton method:

$$\boldsymbol{\gamma}^n \approx -\nabla J(\boldsymbol{\delta}^n)(\nabla^2 J(\boldsymbol{\delta}^n))^{-1} \tag{18}$$

  As the inverse of the Hessian matrix $\nabla^2 J(\boldsymbol{\delta}^n)$ is hard to compute due to the huge dimensions of the system, the following approximation is used:

$$\forall i, \boldsymbol{\gamma}_i^n = \frac{-[\nabla J(\boldsymbol{\delta}^n)]_i}{[\nabla^2 J(\boldsymbol{\delta}^n)]_{ii}} \tag{19}$$

  i.e., only the diagonal elements of the Hessian matrix are taken into account. Basic computations show that this descent direction is equivalent to:

$$\forall i, \boldsymbol{\gamma}_i^n = \underset{\gamma \in \mathbb{R}}{\operatorname{argmin}} \left\{J\left(\widehat{\delta}_1^n, \ldots, \widehat{\delta}_{i-1}^n, \widehat{\delta}_i^n + \gamma, \widehat{\delta}_{i+1}^n, \ldots, \widehat{\delta}_M^n\right)\right\} \tag{20}$$

  The reason why this approximate Newton descent direction is employed instead of a classical gradient $\boldsymbol{\gamma}^n = -\nabla J(\boldsymbol{\delta}^n)$ is that the gradient results in a very slow convergence rate when the components of the system matrix $\mathbf{M}$ exhibit large amplitude variations, which is the case when attenuation correction is taken into account.
- Project $\boldsymbol{\gamma}^n$ onto the hyperplane $\sum_j (\mathbf{M}^{\mathrm{T}} \boldsymbol{\gamma}^n)_j = 0$ so as to preserve the normalization throughout the iterations.
- Compute $\breve{\varepsilon}$ as:

$$\breve{\varepsilon} = \underset{\varepsilon \in \mathbb{R}}{\operatorname{argmin}} \{J(\boldsymbol{\delta}^n + \varepsilon \boldsymbol{\gamma}^n)\} \tag{21}$$

  The minimization here is straightforward since J is quadratic in ε.
- Update $\widehat{\boldsymbol{\delta}}^n$ according to:

$$\widehat{\boldsymbol{\delta}}^{n+1} = \widehat{\boldsymbol{\delta}}^n + \breve{\varepsilon} \boldsymbol{\gamma}^n \tag{22}$$

- Proceed to the next iteration.

At each iteration n, equations (8*b*) and (10) give the current estimate of the ROI activity and associated variance:

$$\widehat{\Phi}^n = \widehat{\boldsymbol{\delta}}^{n\mathrm{T}} \mathbf{p} \tag{23}$$

$$\mathrm{Var}(\widehat{\Phi}^n) \approx \widehat{\boldsymbol{\delta}}^{n\mathrm{T}} \mathrm{Diag}(\mathbf{p}) \widehat{\boldsymbol{\delta}}^n. \tag{24}$$

In the following, this least-squares dual characterization will be abbreviated as LSD characterization.

2.4. LSD characterization with linear extrapolation

During the iterative process described in sub-section 2.3 one can compute what will be referred to as the overlap coefficient:

$$\rho^n = \frac{\mathbf{k}^{\mathrm{T}}\mathbf{M}^{\mathrm{T}}\widehat{\boldsymbol{\delta}}^n}{\mathbf{k}^{\mathrm{T}}\mathbf{k}} \tag{25}$$

Ideally, when solving the unregularized least-squares problem (i.e., when α is set to 0), this coefficient should tend toward 1, since $\mathbf{M}^{\mathrm{T}}\widehat{\boldsymbol{\delta}}^n$ should tend toward $\mathbf{k}$. In practice, due to the inconsistency of the dual problem, ρ^n will reach a plateau value depending on both $\mathbf{M}$ and $\mathbf{k}$. Let us denote A_{roi} the mean activity in the considered ROI and A_{bkg} the mean activity in the ROI background. In the neighborhood of the studied ROI, the original activity map can be decomposed as:

$$\mathbf{f} \approx A_{\mathrm{roi}}\mathbf{k} + A_{\mathrm{bkg}}(\mathbf{i} - \mathbf{k}) \tag{26}$$

where $\mathbf{i} \in \mathbb{R}^N$ is a vector with all components equal to 1. Using equation (9), one can write the expectation of $\widehat{\Phi}^n$ as:

$$E(\widehat{\Phi}^n) = \mathbf{f}^{\mathrm{T}}\mathbf{M}^{\mathrm{T}}\widehat{\boldsymbol{\delta}}^n \approx A_{\mathrm{roi}}\mathbf{k}^{\mathrm{T}}\mathbf{M}^{\mathrm{T}}\widehat{\boldsymbol{\delta}}^n + A_{\mathrm{bkg}}(\mathbf{i}^{\mathrm{T}} - \mathbf{k}^{\mathrm{T}})\mathbf{M}^{\mathrm{T}}\widehat{\boldsymbol{\delta}}^n \tag{27}$$

Noting $S = \mathbf{k}^{\mathrm{T}}\mathbf{k} = \mathbf{i}^{\mathrm{T}}\mathbf{k}$ the number of pixels in the ROI, and noticing that $\mathbf{i}^{\mathrm{T}}\mathbf{M}^{\mathrm{T}}\widehat{\boldsymbol{\delta}}^n = \mathbf{i}^{\mathrm{T}}\mathbf{k}$ (due to the normalization of $\widehat{\boldsymbol{\delta}}^n$ throughout the iterative process), one has:

$$E(\widehat{\Phi}^n) \approx SA_{\mathrm{roi}}\rho^n + SA_{\mathrm{bkg}} - SA_{\mathrm{bkg}}\rho^n = SA_{\mathrm{bkg}} + S(A_{\mathrm{roi}} - A_{\mathrm{bkg}})\rho^n \tag{28}$$

If the ROI and background activities are sufficiently uniform, $E\left(\widehat{\Phi}^n\right)$ is approximately linear in ρ^n and $E(\widehat{\Phi}^n)$ tends toward the true ROI value $\mathrm{F} = SA_{\mathrm{roi}}$ when ρ^n tends toward 1. Let us consider the series $\{\rho^n, \widehat{\Phi}^n\}$ and $\{\rho^n, \mathrm{Std}(\widehat{\Phi}^n)\}$, with $\mathrm{Std}(\widehat{\Phi}^n)$ the square root of $\mathrm{Var}(\widehat{\Phi}^n)$. When ρ^n reaches its plateau value, $\widehat{\Phi}^n$ also reaches a plateau and the subsequent iterations do not improve the estimate $\widehat{\Phi}^n$ in terms of bias. However, $\|\widehat{\boldsymbol{\delta}}^n\|$ keeps on increasing with the iterations which leads (via equation (24)) to a monotonic rise in $\mathrm{Std}(\widehat{\Phi}^n)$, i.e., in the variance of the estimate. Choosing the optimal value of ρ^n at which to stop the iterative process would require knowing the expected bias in order to compare it with the current variance and evaluate the bias-variance tradeoff, which is not the case in practice. An alternative strategy consists in stopping empirically the iterative process when the convergence of ρ^n becomes slow and extrapolating the values of $\widehat{\Phi}$ and $\mathrm{Std}(\widehat{\Phi})$ for $\rho = 1$ using a linear regression. Let us denote $[\rho^{n_{\min}}\rho^{n_{\max}}]$ the relevant range of ρ values, with $n_{\min} \geqslant 1$ (in practice the first iterations are discarded since they are not representative of the general trend) and $n_{\max}$ corresponding to the iteration index at which the iterative process is stopped. A subset of k ρ values evenly distributed in the relevant range is chosen:

$\{\rho^{n_1} = \rho^{n_{\min}}; \rho^{n_2}; \ldots; \rho^{n_{k-1}}; \rho^{n_k} = \rho^{n_{\max}}\}$, corresponding to k ROI estimates $\{\widehat{\Phi}^{n_1}; \ldots; \widehat{\Phi}^{n_k}\}$ and k variance estimates $\{\mathrm{Std}(\widehat{\Phi}^{n_1}); \ldots; \mathrm{Std}(\widehat{\Phi}^{n_k})\}$. A linear regression is then performed on the two series $\{\rho^{n_t}, \widehat{\Phi}^{n_t}\}$ and $\{\rho^{n_t}, \mathrm{Std}(\widehat{\Phi}^{n_t})\}$ with $t = 1 \ldots \mathrm{k}$:

$$\Phi \approx A(\rho^{n_1}; \ldots; \rho^{n_k}; \widehat{\Phi}^{n_1}; \ldots; \widehat{\Phi}^{n_k})\rho + B(\rho^{n_1}; \ldots; \rho^{n_k}; \widehat{\Phi}^{n_1}; \ldots; \widehat{\Phi}^{n_k}) \tag{29a}$$

where A and B are the first and zero order coefficients of the linear regression respectively,

$$\mathrm{Std}(\Phi) \approx C(\rho^{n_1}; \ldots; \rho^{n_k}; \mathrm{Std}(\widehat{\Phi}^{n_1}); \ldots; \mathrm{Std}(\widehat{\Phi}^{n_k}))\rho + D(\rho^{n_1}; \ldots; \rho^{n_k}; \mathrm{Std}(\widehat{\Phi}^{n_1}); \ldots; \mathrm{Std}(\widehat{\Phi}^{n_k})) \tag{29b}$$

where C and D are the first and zero order coefficients of the linear regression respectively. For illustrative purposes, figure 1 shows an example for the regression $\Phi \approx A\rho + B$. We denote LSD-ex the LSD characterization obtained when extrapolating the two linear regressions for

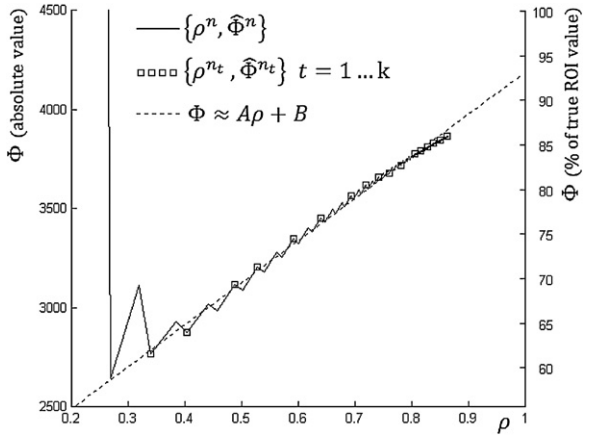

Figure 1. Example of a linear regression over a series $\{\rho^{n_t}, \widehat{\Phi}^{n_t}\}, t = 1 \ldots k$.

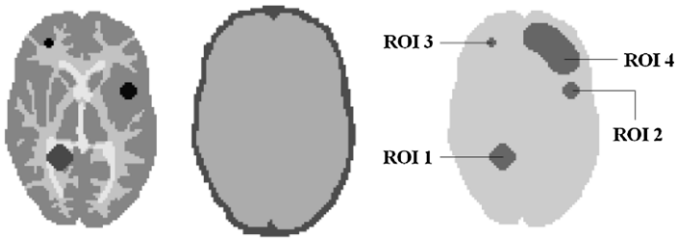

Figure 2. Hoffman brain phantom. From left to right are displayed the activity map, the density map, and the ROI subdivision.

Table 1. 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
ROI 1	1.5	0.1
ROI 2	2	0.1
ROI 3	3	0.1

the value $\rho = 1$:

$$\widehat{\Phi}_{\text{ex}} = A(\rho^{n_1}; \ldots; \rho^{n_k}; \widehat{\Phi}^{n_1}; \ldots; \widehat{\Phi}^{n_k}) + B(\rho^{n_1}; \ldots; \rho^{n_k}; \widehat{\Phi}^{n_1}; \ldots; \widehat{\Phi}^{n_k}) \tag{30a}$$

$$\begin{aligned}\text{Std}(\widehat{\Phi}_{\text{ex}}) = {} & C(\rho^{n_1}; \ldots; \rho^{n_k}; \text{Std}(\widehat{\Phi}^{n_1}); \ldots; \text{Std}(\widehat{\Phi}^{n_k})) \\ & + D(\rho^{n_1}; \ldots; \rho^{n_k}; \text{Std}(\widehat{\Phi}^{n_1}); \ldots; \text{Std}(\widehat{\Phi}^{n_k}))\end{aligned} \tag{30b}$$

2.5. Monte Carlo validation

We studied a Hoffman brain phantom. Figure 2 shows the 2D activity and density maps and the ROIs used in our study. Table 1 gives the relative uptakes and attenuation coefficients of the different tissues. Three uniform hot spots of various sizes and uptakes mimicking tumors were

included inside the cerebral parenchyma and corresponded to ROI 1–3. A fourth non-uniform ROI was defined inside the frontal lobe, with boundaries not correlated with anatomical region borders. Regions 1–3 were relatively small ROIs with high contrast with respect to their background and were typical of clinical situations when one wants to estimate the statistical properties of a hot tumor, whereas region 4 was a large ROI with no particular contrast with respect to its neighborhood and was representative of a non specific area serving as a reference for ROI uptake comparison (for instance in dopaminergic neuroimaging where an occipital ROI is used to assess the striatal uptake through the calculation of the binding potential (Soret *et al* 2003)). From a statistical point of view, the estimation errors regarding ROI 1–3 will mainly be caused by bias due to partial volume effects. In ROI 4, the bias is expected to be small and the estimation error will mainly reflect the estimate variability.

The projection data were produced using the GATE (Geant4 Application in Emission Tomography) platform (Jan 2004, Jan 2011). The numerical Hoffman phantom we used was an axially invariant 3D phantom whose 2D activity and density maps are displayed in figure 2 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) (Lamare *et al* 2006). Positron range and non-collinearity were not modeled. A 3D data set was produced corresponding to a total activity of 80 MBq (^{18}FDG) and an acquisition time of 10 min. The simulated prompts, estimated scatter (Bailey and Meikle 1994), and estimated random (Kadrmas 2004) were stored. The 29 transaxial sinograms of each data set (prompts, scatter and random) were then summed and re-sampled to produce three high-statistics 2D 128×128 sinograms. The prompts 2D sinogram included about 10 million counts, the scatter 2.4 million counts, and the random 1.8 million counts. These 2D sinograms were considered as noise-free (Fu *et al* 2009), i.e., as the exact distributions of the prompts, scatter and random. Let us note respectively $\bar{\mathbf{p}}$, $\bar{\mathbf{s}}$ and $\bar{\mathbf{r}}$ these three distributions. The distribution of the trues (i.e., the corrected projections) was computed as $\bar{\mathbf{t}} = \bar{\mathbf{p}} - \bar{\mathbf{s}} - \bar{\mathbf{r}}$. Realistic noisy data sets were simulated using the following method:

- Choose a count rate C.
- Scale the exact distributions using the appropriate scaling factors μ and η:

$$\bar{\bar{\mathbf{t}}} = \mu\bar{\mathbf{t}}; \quad \bar{\bar{\mathbf{s}}} = \mu\bar{\mathbf{s}}; \quad \bar{\bar{\mathbf{r}}} = \eta\mu\bar{\mathbf{r}} \tag{31}$$

  where μ is such that $\bar{\bar{\mathbf{t}}}$ includes C counts, and η accounts for the quadratic variation of the number of random counts with the count rate.
- Simulate the noisy measured prompts $\mathbf{p}$ as:

$$\mathbf{p} = \Pi(\bar{\bar{\mathbf{t}}}) + \Pi(\bar{\bar{\mathbf{s}}}) + \Pi(\bar{\bar{\mathbf{r}}}) \tag{32}$$

  where Π stands for the Poisson noise operator (modeled using Knuth's algorithm (Knuth 1969)).
- Simulate the estimated random $\widehat{\mathbf{r}}$ as:

$$\widehat{\mathbf{r}} = \Pi(\bar{\bar{\mathbf{r}}}) \tag{33}$$

- Simulate the estimated scatter $\widehat{\mathbf{s}}$ as:

$$\widehat{\mathbf{s}} = \Pi(\bar{\bar{\mathbf{s}}}) * \mathbf{G}_2 \tag{34}$$

  where the convolution with a 2D Gaussian kernel of FWHM 2 pixels accounts for the smooth nature of the estimated scatter.

Three count rates were considered: 50 000, 200 000, and 800 000 counts. For each count rate, 1000 realistic replicates of the measured prompts $\mathbf{p}$, estimated scatter $\widehat{\mathbf{s}}$, and estimated random $\widehat{\mathbf{r}}$, were generated. The system matrix $\mathbf{M}$ employed for the reconstructions and ROI

estimations was built using a uniformly distributed pixel activity model and corrected for attenuation and normalization using the appropriate correction factors. The images were reconstructed on a 128 × 128 grid with a pixel size of 2 mm. The following methods were evaluated for the task of ROI characterization:

- Maximum-likelihood expectation-maximization (MLEM) (Shepp and Vardi, 1982, Lange and Carson, 1984) was run on 400 replicates for each count rate, using the measured prompts $\mathbf{p}$. The estimated random $\widehat{\mathbf{r}}$ and scatter $\widehat{\mathbf{s}}$ were taken into account by adding them to the projections of the current image estimate at each iteration, according to the ordinary Poisson model proposed by Politte and Snyder (Politte and Snyder, 1991). The total number of iterations was 100 for the 50 000 count sinograms, 150 for the 200 000 count sinograms, and 200 for the 800 000 count sinograms. These values were set empirically so as to ensure sufficient convergence of the algorithm. At each iteration, the ROI estimates were deduced by pixel summation. The bias was computed as the absolute difference between the mean of the current estimates (over the 400 replicates) and the true ROI value. The variability was computed as the standard deviation of the current estimates (over the 400 replicates).
- Huesman's method was run on the 1000 replicates for each count rate. Here, the system matrix was a simple geometric projector and the projections were corrected for attenuation and normalization. The ROI estimates were produced using equation (6) as:

$$\widehat{\Phi} = \mathbf{k}^{\mathrm{T}}\mathbf{BCAp} - \mathbf{k}^{\mathrm{T}}\mathbf{BCA}\widehat{\mathbf{r}} - \mathbf{k}^{\mathrm{T}}\mathbf{BCA}\widehat{\mathbf{s}}. \tag{35}$$

  The bias was computed as the absolute difference between the mean of the estimates (over the 1000 replicates) and the true ROI value. The variability was computed as the standard deviation of the estimates (over the 1000 replicates). As this has already been investigated elsewhere, the agreement between the true variance and the estimated variance provided by equation (7) was not studied.
- LSD characterization. The iterative process described in section 2.3 was run on the 1000 replicates for each count rate. The total number of iterations was 100 for the 50 000 count sinograms, 150 for the 200 000 count sinograms, and 200 for the 800 000 count sinograms. These values are identical to those chosen for MLEM: the computation time of an iteration step of MLEM being very similar to that of an iteration step of LSD, the same total number of iterations was used for both algorithms so that their performance can be compared on the basis of identical computational costs. The ROI activity was computed using equations (23):

$$\widehat{\Phi}^n = \widehat{\boldsymbol{\delta}}^{n\mathrm{T}}(\mathbf{p} - \widehat{\mathbf{r}} - \widehat{\mathbf{s}}) \tag{36}$$

  The LSD method does not require the correction of the projections prior to the ROI characterization, thus avoiding the need to set negative projection values to zero (as is sometimes done before MLEM reconstruction). The ROI variability was then computed using equations (24):

$$\mathrm{Std}(\widehat{\Phi}^n) = \sqrt{\mathrm{Var}(\widehat{\Phi}^n)} \approx \sqrt{\widehat{\boldsymbol{\delta}}^{n\mathrm{T}}\mathrm{Diag}(\mathbf{p} + \widehat{\mathbf{r}})\widehat{\boldsymbol{\delta}}^n} \tag{37}$$

  where the variance of the estimated scatter was assumed to be negligible. As discussed above, for ROI 1–3, as the goal was to minimize the bias, LSD characterization was performed with non smoothed ROI characteristic functions, i.e., $\alpha = 0$ in equation (15). In ROI 4, as the main error source is expected to be the variance, LSD characterization was performed with three smoothing levels, corresponding to $\alpha = 0$, $\alpha = 2$ and $\alpha = 4$ pixels. In addition, LSD characterization was also run using a classical gradient minimization to justify the choice of the Newton-like descent direction employed in the recipe described in section 2.3.

- LSD-ex characterization. This method was only applied to ROI 1–3 and used the 1000 noisy replicates for each count rate. The total number of iterations was set to $n_{\max} = 40$. The subset of ρ values included $k = 15$ points: $\{n_1 \dots n_{15}\} = \{4, 5, 6, 7, 8, 9, 10, 12, 15, 18, 21, 25, 30, 35, 40\}$. The corresponding range of ρ values was [0.58 0.91] for ROI 1, [0.48 0.87] for ROI 2, and [0.34 0.77] for ROI 3. The bias was computed as the absolute difference between the mean of the estimates $\widehat{\Phi}_{\text{ex}}$ (over the 1000 replicates) and the true ROI value. The variability was computed as the standard deviation of the estimates $\widehat{\Phi}_{\text{ex}}$ (over the 1000 replicates). The variance estimates given by equation (30*b*) were computed and assessed through their 95% confidence interval ($\text{CI}_{95\%}$) over the 1000 replicates.

The reason why MLEM was run on a restricted subset (400 replicates) of the whole simulated data set (1000 replicates) is that MLEM reconstruction is highly time-consuming. For Huesman's method, $\mathbf{k}^{\text{T}}\mathbf{BCA}$ was pre-computed and stored for a given ROI then applied to the 1000 noisy replicates. Similarly, for LSD, the successive values of the dual characteristic function $\widehat{\boldsymbol{\delta}}^n$ $(n = 1 \dots 200)$ were pre-computed and stored for a given ROI then applied to the 1000 noisy replicates, which significantly quickened the ROI estimation and allowed to process the 1000 replicates within a reasonable computation time. To each bias and variance computation corresponded a root mean square (RMS) error value according to RMS error $= \sqrt{\text{bias}^2 + \text{variance}}$. In the result section, these three figures of merit (bias, standard deviation and RMS error) are normalized and expressed in percent of the true ROI value.

3. Results

Figure 3 shows the intrinsic convergence properties of LSD characterization. The iterative process was performed with $\alpha = 0$. Figures 3(A/D), (B/E) and (C/F) correspond to ROI 1 to 3 respectively. Figures 3(A)–(C) show the evolution of the objective function defined in equation (14) with the iteration number and figures 3(D)–(F) represent the evolution of the overlap coefficient defined in equation (25) with the iteration number. The solid curves correspond to the use of the approximate Newton descent described in section 2.3, the dotted curves correspond to the classical gradient. Figures 4–6 compare the four estimation methods and correspond respectively to the 50 000 count, 200 000 count, and 800 000 count simulations. In each figure, ROIs 1–4 are shown from left to right. For each of the four ROIs, row A shows the estimation bias, row B displays the associated standard deviation, and row C displays the resulting RMS error. The thick dotted curves correspond to the results of MLEM, the thin solid curves to the results of LSD given by equation (36), the gray dashed lines to the results of Huesman's method given by equation (35), and the crosses to the results of LSD-ex given by equation (30*a*). The thick solid curves labeled '2' and '4' refer to the results of LSD using respectively $\alpha = 2$ and $\alpha = 4$ pixels. In row B, the thin dotted curves indicate the 95% confidence interval of the LSD variability estimate provided by equation (37). Table 2 shows the adequacy between the variability of the LSD estimates at iteration 100 (obtained using the noisy replicates) and the predicted value of that variability computed using equation (37). Last, tables 3 and 4 summarize the performances of LSD-ex. Table 3 compares the LSD-ex estimates with the optimal MLEM estimates (MLEM-opt, corresponding to the iteration number for which the RMS error reaches its minimum) in terms of bias, standard deviation, and RMS error. Table 4 shows the adequacy between the variability of the LSD-ex estimates (obtained using the noisy replicates) and the predicted value of that variability computed using equation (30*b*).

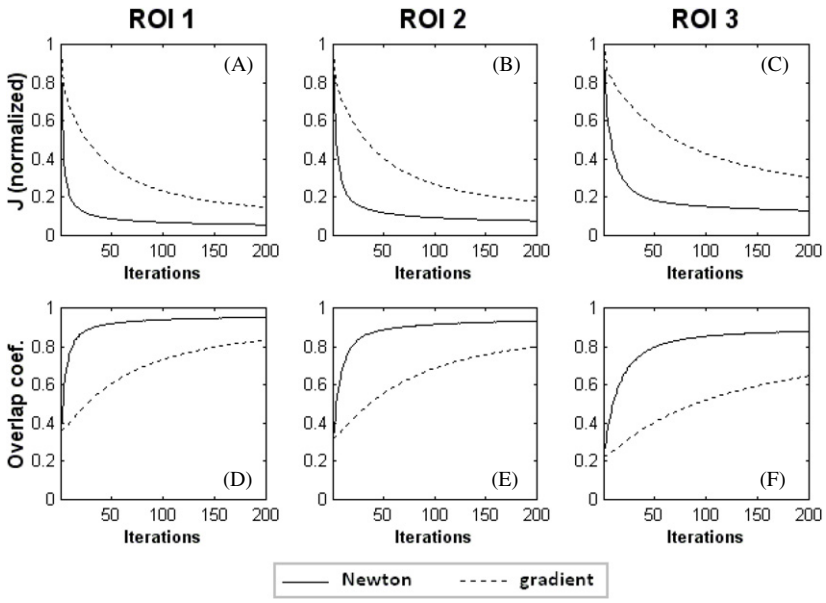

Figure 3. Convergence properties of LSD characterization for ROI 1 (left), ROI 2 (middle), and ROI 3 (right) in terms of the objective function J (top) and the overlap coefficient ρ (bottom).

Table 2. Variability of the LSD estimates at iteration 100: based on the noisy replicates (first column), and predicted using equation (37) (second column).

		Std($\hat{\Phi}$) (%)	Std($\hat{\Phi}$) predicted $CI_{95\%}$ (%)	Error
50 000 counts	ROI 1	10.31	9.75–10.01	−0.43%
	ROI 2	12.48	12.06–12.38	−0.26%
	ROI 3	21.54	20.53–21.33	−0.61%
200 000 counts	ROI 1	5.24	5.01–5.08	−0.20%
	ROI 2	6.54	6.19–6.27	−0.32%
	ROI 3	10.48	10.32–10.52	−0.06%
800 000 counts	ROI 1	3.41	3.22–3.24	−0.18%
	ROI 2	4.08	3.92–3.94	−0.15%
	ROI 3	6.79	6.62–6.67	−0.15%

4. Discussion

The convergence properties of the LSD algorithm shown in figure 3 clearly justify the use of our Newton-like descent for the minimization of the objective function since the overlap coefficient reaches a plateau within an acceptable number of iterations depending on the ROI size (between 25 for the large ROI 1 and 50 for the small ROI 3). When using the classical gradient, the convergence of the objective function is very slow and after 200 iterations the plateau of the overlap coefficient is far from being reached. These results further justify the choice of $n_{\max} = 40$ as a reasonable number of iterations after which to stop the iterative process before performing the linear extrapolation in LSD-ex. In practice, this choice can be adapted on a case-by-case basis since the trend of the overlap coefficient is perfectly known. Furthermore, this trend is independent of the measured data in general and of the

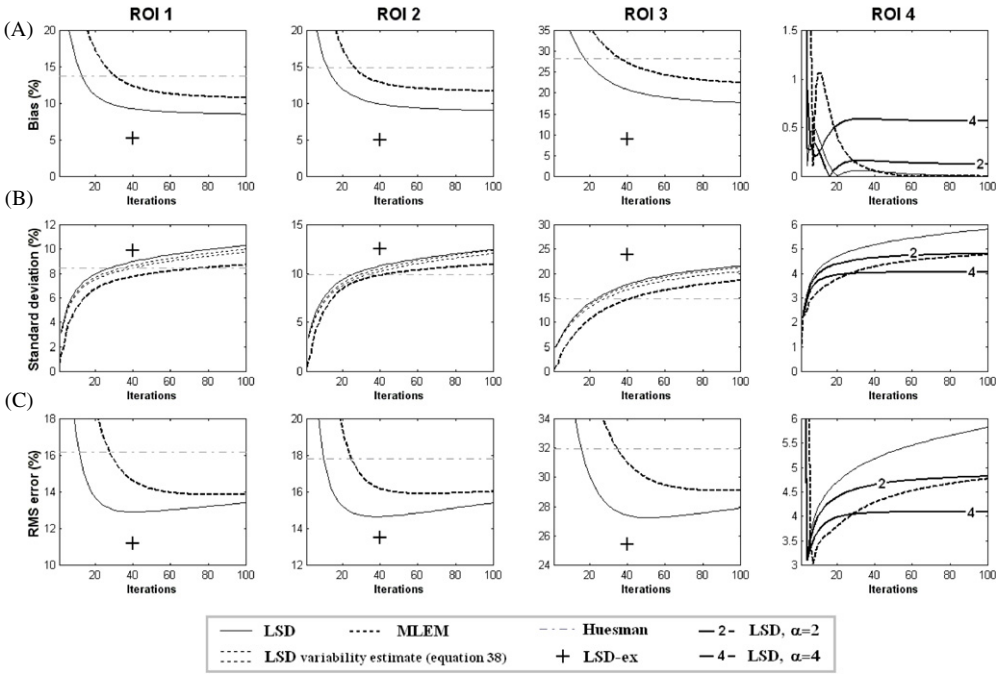

Figure 4. 50 000 count simulations— comparison in terms of bias, standard deviation and RMS error of the performances of LSD (thin solid curves), MLEM (thick dotted curves), Huesman's method (gray dashed lines), LSD-ex (crosses), and LSD with $\alpha = 2$ and $\alpha = 4$ pixels (thick solid lines labeled '2' and '4'). Row B, the thin dotted curves indicate the 95% confidence interval of the LSD variability estimate provided by equation (38).

Table 3. Comparison between MLEM-opt and LSD-ex estimations in terms of bias, standard deviation and RMS error.

		50 000 counts			200 000 counts			800 000 counts		
		ROI 1	ROI 2	ROI 3	ROI 1	ROI 2	ROI 3	ROI 1	ROI 2	ROI 3
Bias	MLEM	10.9%	12.0%	22.9%	9.0%	11.0%	18.8%	8.5%	9.9%	18.5%
	LSD-ex	5.2%	5.0%	8.9%	4.8%	5.5%	9.2%	4.9%	5.5%	8.2%
	Difference	**−5.7%**	**−7%**	**−14%**	**−4.2%**	**−5.5%**	**−9.6%**	**−3.6%**	**−4.4%**	**−10.3%**
Std. dev.	MLEM	8.6%	10.4%	18.1%	4.8%	5.9%	10.2%	3.1%	3.9%	6.4%
	LSD-ex	9.9%	12.5%	23.8%	5.1%	6.5%	11.4%	3.3%	4.0%	7.5%
	Difference	**+ 2.3%**	**+ 2.1%**	**+ 5.7%**	**+ 0.3%**	**+ 0.6%**	**+ 1.2%**	**+ 0.2%**	**+ 0.1%**	**+ 1.1%**
RMS error	MLEM	13.8%	15.9%	29.1%	10.3%	12.5%	21.5%	9.1%	10.6%	19.6%
	LSD-ex	11.2%	13.5%	25.4%	7.0%	8.5%	14.6%	5.9%	6.8%	11.1%
	Difference	**−2.6%**	**−2.4%**	**−3.7%**	**−3.3%**	**−4%**	**−6.9%**	**−3.2%**	**−3.8%**	**−8.5%**

count rate in particular (it only depends on the system matrix and ROI characteristic function), which should facilitate an automated design of the extrapolation procedure when working on successive recordings of the same ROI.

Comparing the four methods used for ROI characterization (figures 4–6), as expected, the bias is rather independent of the count level. This is strictly true for LSD, LSD-ex, and Huesman's method since they are linear. For MLEM, the bias slightly increases at low count rate due to the nonlinear nature of the method. Regarding the standard deviation, as expected,

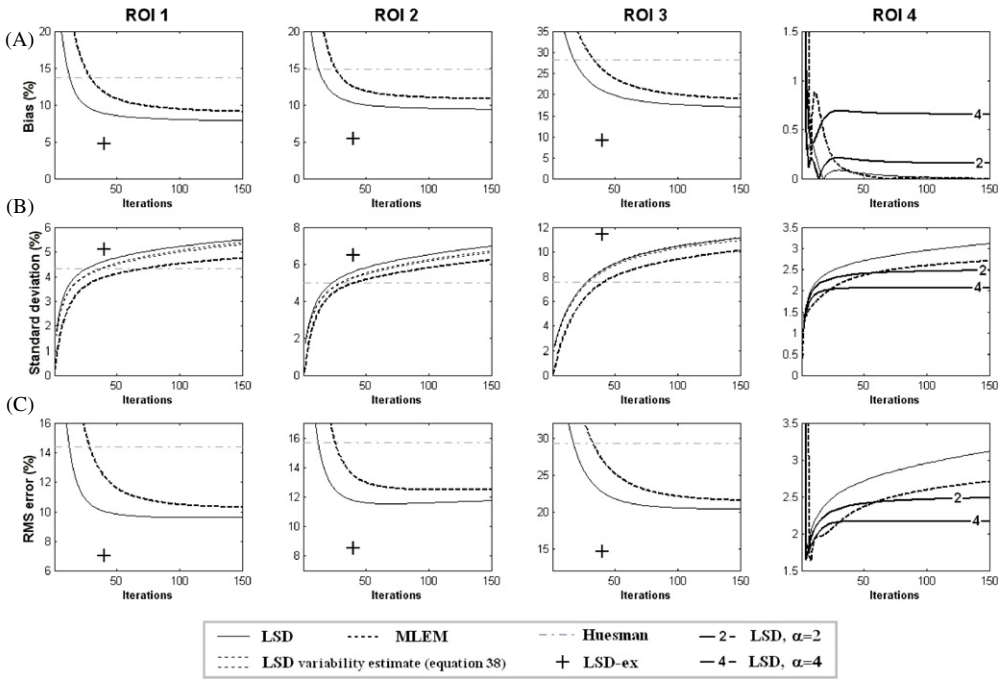

Figure 5. Same as figure 4 for the 200 000 count simulations.

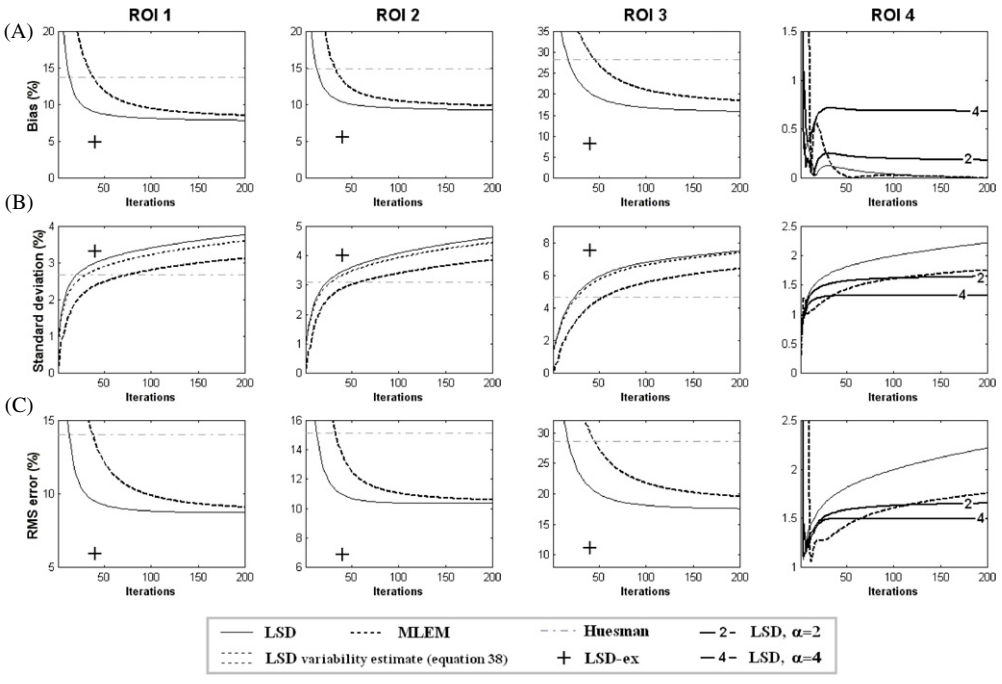

Figure 6. Same as figure 4 for the 800 000 count simulations.

Table 4. Adequacy between the real and predicted variability of the LSD-ex estimates.

		Std($\hat{\Phi}_{ex}$) (%)	Std($\hat{\Phi}_{ex}$) predicted $CI_{95\%}$ (%)	Error
50 000 counts	ROI 1	9.92	9.26–9.47	−0.56%
	ROI 2	12.50	11.71–12.00	−0.65%
	ROI 3	23.83	22.50–23.40	−0.88%
200 000 counts	ROI 1	5.13	4.75–4.80	−0.35%
	ROI 2	6.50	6.01–6.08	−0.46%
	ROI 3	11.43	11.08–11.30	−0.24%
800 000 counts	ROI 1	3.32	3.06–3.07	−0.26%
	ROI 2	4.01	3.80–3.82	−0.19%
	ROI 3	7.52	7.11–7.16	−0.39%

it is approximately reduced by a factor of 2 (whatever the method) when changing from 50 000 to 200 000 counts then from 200 000 to 800 000 counts. As discussed above, the RMS error in ROI 1–3 estimates depends on the relative proportion of the bias and the variance. At high count rate, the RMS error is mainly influenced by the bias, whereas at low count rate the variance has a significant and early impact on the evolution of the RMS error. In ROI 4, the RMS error mostly results from the noise since the bias tends to zero as the iterations proceed.

Huesman's method results in estimation biases that are significantly higher than the two iterative methods (that is why the corresponding dashed line does not appear in ROI 4 plots for which the bias is around 5%). On the contrary, the variance of Huesman's estimates remains moderate compared with MLEM and LSD. The resulting RMS error is significantly higher in Huesman's method than with MLEM and LSD, whatever the ROI and count rate.

The performance of MLEM and LSD are rather similar in terms of bias, variance and RMS error. Likely due to the appropriate definition of the descent direction employed for the minimization of the objective function, the convergence of LSD is faster than that of MLEM (figures 4–6, rows A), yielding a faster bias decrease. For a given iteration, the variance of the LSD estimates is always slightly higher than that of the MLEM estimates, due to the higher convergence rate (figures 4–6, rows B). The resulting RMS errors are comparable for LSD and MLEM and follow roughly the same trends whatever the studied ROI and count level (figures 4–6, rows C). For the three small tumor ROIs, it appears that the RMS error is always significantly lower with LSD than with MLEM. Regarding the non-specific large ROI, the RMS error is lower with MLEM, but the difference between the two methods remains small and weakly significant from a practical point of view. As confirmed by table 2, the LSD variance estimates are fairly accurate since the error after 100 iterations is always below 1% (the systematic under-estimation follows likely from the fact that the variance of the estimated scatter is neglected in equation (37)). The LSD variance estimates also prove to be highly robust as confirmed by the narrow 95% confidence intervals.

In ROI 4, the minimum RMS error is reached early in the iterative process and the subsequent RMS error rise follows the trend of the variance increase. For this kind of ROI with large size and low contrast, a regularized estimation is hence desirable to avoid a large increase in variance. As described in section 2.3, the regularization can be performed by smoothing the characteristic function of the ROI with an adjustable convolution kernel. Figure 7 shows the effect of regularization on the convergence properties of the LSD algorithm applied to ROI 4. The top row displays the characteristic function $\breve{\mathbf{k}}_\alpha$ which, as discussed in section 2.3, can be interpreted as the target resolution kernel for LSD characterization. The second row shows the exact object smoothed using the corresponding Gaussian filter $\mathbf{G}_\alpha$. It clearly appears that the filtering definitely alters the intrinsic resolution of the small tumor ROIs due to their

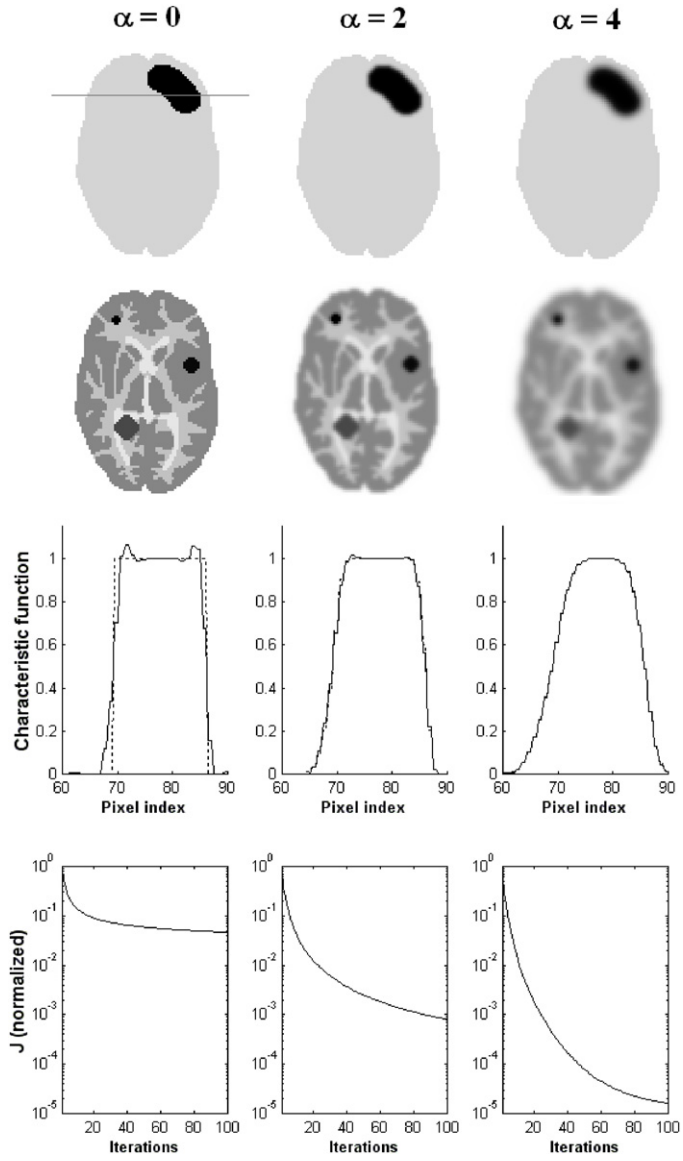

Figure 7. Unregularized LSD ($\alpha = 0$, left column) and regularized LSD with $\alpha = 2$ (middle column) and $\alpha = 4$ (right column). Top row: ROI 4 characteristic function $\breve{\mathbf{k}}_\alpha$. Second row: phantom smoothed using $\mathbf{G}_{\boldsymbol{\alpha}}$. Third row: line profiles through $\breve{\mathbf{k}}_\alpha$ (dotted) and $\mathbf{M}^T\hat{\boldsymbol{\delta}}$ (solid) at iteration 100. Bottom row: convergence of the objective function J.

high contrast. In ROI 4, the smoothing and the related loss of resolution are not expected to significantly bias the ROI estimate. The third row displays line profiles through the target resolution kernel $\breve{\mathbf{k}}_\alpha$ (dotted) and through the resolution kernel $\mathbf{M}^T\widehat{\boldsymbol{\delta}}$ (solid) at iteration 100. Whereas the match between the two is perfect for $\alpha = 4$ and almost perfect for $\alpha = 2$, the solution of the unregularized dual problem fails in perfectly matching $\mathbf{M}^T\widehat{\boldsymbol{\delta}}$ with $\breve{\mathbf{k}}_\alpha$. The bottom row shows the convergence of the objective function throughout the iterations and illustrates that the roughness of the regularization strongly influences the speed and quality of

the convergence. These observations explain the results shown in figures 4–6 for ROI 4. In terms of bias, the regularization does not yield a significant degradation of the ROI estimates (almost 0% at convergence for $\alpha = 0$, about 0.2% for $\alpha = 2$ and less than 0.7% for $\alpha = 4$). In terms of variance, the use of $\alpha = 2$ significantly reduces the variance compared to $\alpha = 0$, yet the variance still follows a slow monotonic increase while the iterations proceed. Choosing $\alpha = 4$ yields a further reduction along with a fast and sustainable stabilization of the variance (a plateau is reached within about 30 iterations). As a consequence, the strongly regularized LSD algorithm produces the most accurate and stable ROI estimate in terms of RMS error. So far, the adjustment of parameter α has been performed intuitively, based on basic *a priori* knowledge regarding the considered problem such as the size of the ROI, its expected contrast with respect to the background, the SNR in the recorded projections. Further work has to be undertaken in order to design an automated procedure for the optimization of parameter α.

In the small tumor ROIs, LSD-ex characterization was also used to study to what extent it could reduce the bias and RMS error. It appears from figures 4–6 and table 3 that the bias was lowered compared with LSD and MLEM due to the compensation of sampling-induced partial volume effects. The residual bias affecting LSD-ex estimates is mainly due to the re-sampling of the projection data and to the modeling errors affecting the system matrix which does not account for the detector response function. When compared with the optimal MLEM estimate in terms of the RMS error (MLEM-opt), the LSD-ex estimate reduces the bias from 4% to 14% (depending on the ROI and count level). The variance of the LSD-ex estimate is similar to that of the LSD estimate around 100 iterations (except for the smallest ROI for which the variance is slightly higher, likely due to a wider range of ρ values over which the extrapolation is performed, producing a higher variability in the linear regression). The resulting reduction in terms of RMS error with comparison to MLEM-opt ranged from 2.5% to 8.5%. Furthermore, table 4 shows that the LSD-ex estimate variance predicted using equation (30*b*) is highly robust and accurate since the associated 95% confidence interval is always narrow and close to the true LSD-ex estimate variance (the variance is mostly under-estimated with an error always below 1%). Yet, before the LSD-ex method can be widely used, an automatic criterion for the definition of the series to be extrapolated will have to be designed, so that the procedure could be fully user-independent.

Regarding the computational requirements of the compared algorithms, the LSD characterization of one ROI requires approximately the same time as that needed by an MLEM iterative reconstruction followed by pixel summation. When several ROIs are studied, the LSD algorithm has to be run for each ROI. Yet, when performing successive characterizations of the same ROI (as in dynamic imaging), the dual characteristic function has only to be estimated once.

As a summary, one can list the advantages of dual ROI characterization versus classical direct ROI quantification:

- No need for image reconstruction.
- A similar computational complexity.
- A higher convergence speed allowed by the Newton-like descent.
- A linear expression of the ROI value enabling a straightforward estimation of the ROI variance.
- The ability to support non-uniform ROIs.
- An intuitive handling of the regularization of the inverse problem using a smoothed version of the ROI characteristic function. The resulting estimation bias is directly interpretable in terms of resolution loss.
- An innovative extrapolation scheme allowing a significant reduction in the estimation bias induced by sampling partial volume effects. When LSD-ex is applied to the ill-posed

unregularized least-squares problem, the iterative process can be suitably stopped very early in the iterations (based on the knowledge of the overlap coefficient), hence limiting the estimate variance increase without requiring any stopping criterion.

- Performances at least as good as those of optimal direct quantification (in terms of RMS error).
- An inverse problem that is independent of the measured data. When ROI estimation is required on successive recordings of the same object, the dual problem needs only to be solved once. This should be particularly helpful in the frame of dynamic imaging.

5. Conclusion

This paper presents an original method for region of interest characterization in emission tomography relying on a dual formulation of the ROI estimation problem inspired by the works of Louis on the approximate inverse. The dual formulation yields estimations for the ROI activity and variance directly from the measured data without any image reconstruction. It requires the definition of an ROI characteristic function that can be derived from a co-registered morphological image. When this characteristic function is properly smoothed, it acts as a target resolution kernel for the characterization, and makes it possible to intuitively tune the resolution-variance tradeoff. The dual problem is solved in the least-squares sense (LSD characterization) by means of an iterative descent method. At each iteration, an LSD estimate is produced while having a precise knowledge of the current resolution kernel (within system modeling errors) and estimate variance. Furthermore, a linear extrapolation scheme is described to reduce the estimation bias (LSD-ex). Both LSD and LSD-ex were compared with classical ROI estimation using pixel summation after MLEM image reconstruction and with Huesman's reference method through Monte Carlo simulations. We found that the performances of LSD characterization are at least as good as those of the classical methods in terms of RMS error. The convergence of LSD is faster than that of direct MLEM estimation and the LSD variance estimates is highly robust and accurate. For three small tumor ROIs, LSD-ex allows a reduction in the estimation bias of up to 14%, resulting in a reduction in the RMS error of up to 8.5%, compared with the optimal MLEM estimation. For a large non specific region, LSD using appropriate smoothing makes it possible to intuitively tune the resolution-variance tradeoff.

References

Alpert N M, Chesler D A, Correia J A, Ackerman R H, Chang J Y, Finklestein S, Davis S M, Brownell G L and Taveras J M 1982 Estimation of the local statistical noise in emission computed tomography *IEEE Trans. Med. Imaging* **1** 142–6

Bailey D L and Meikle S R 1994 A convolution-subtraction scatter correction method for 3D PET Phys *Med. Biol.* **39** 411–24

Barrett H H, Wilson D W and Tsui B M 1994a Noise properties of the EM algorithm: I. theory Phys *Med. Biol.* **39** 833–46

Barrett H H, Wilson D W and Tsui B M 1994b Noise properties of the EM algorithm: II. Monte-Carlo simulations Phys *Med. Biol.* **39** 847–72

Budinger T F, Derenzo S E, Gullberg G T, Greenberg W L, William L and Huesman R H 1977 Emission computer assisted tomography with single-photon and positron annihilation photon emitters *J. Comput. Assist. Tomogr.* **1** 131–45

Buvat I 2002 A bootstrap approach for analyzing the statistical properties of SPECT and PET images Phys *Med. Biol.* **47** 1761–75

Carson R E 1986 A maximum likelihood method for region-of-interest evaluation in emission tomography *J. Comput. Assist. Tomogr.* **10** 654–63

Carson R E, Yan Y, Daube-Witherspoon M E, Freedman N, Bacharach S L and Herscovitch P 1993 An approximation formula for the variance of PET region-of-interest values IEEE Trans. Med *Imaging* **12** 240–50

Fessler J A 1996 Mean and variance of implicitly defined biased estimators (such as penalized maximum l ikelihood): applications to tomography IEEE Trans *Image Process.* **5** 493–506

Formiconi A R 1993 Least squares algorithm for region-of-interest evaluation in emission tomography *IEEE Trans. Med. Imaging* **1** 90–100

Fu L, Stickel J R, Badawi R D and Qi J 2009 Quantitative accuracy of penalized-likelihood reconstruction for ROI activity estimation *IEEE Trans. Nucl. Sci.* **56** 167–72

Gillen G J 1992 A simple method for the measurement of local statistical noise levels in SPECT *Phys. Med. Biol.* **37** 1573–9

Golub G H and von Matt U 1997 Generalized cross-validation for large-scale problems *J. Comput. Graph. Stat.* **6** 1–34

Hansen P C and O'Leary D P 1993 The use of the L-curve in the regularization of discrete ill-posed problem SIAM *J. Sci. Comput.* **14** 1487–503

Haynor D R and Woods S D 1989 Resampling estimates of precision in emission tomography *IEEE Trans. Med Imaging* **8** 337–43

Huesman R H 1977 Effects of a finite number of projection angles and finite lateral sampling of projections on the propagation of statistical errors in transverse section reconstruction *Phys. Med. Biol.* **22** 511–21

Huesman R H 1984 A new fast algorithm for the evaluation of regions of interest and statistical uncertainty in computed tomography *Phys. Med. Biol.* **29** 543–52

Jan S *et al* (for the OpenGATE collaboration) 2004 GATE: a simulation toolkit for PET and SPECT *Phys. Med. Biol.* **49** 4543–61

Jan S *et al* 2011 GATE V6: a major enhancement of the GATE simulation platform enabling modelling of CT and radiotherapy *Phys. Med. Biol.* **56** 881–901

Kadrmas D J 2004 LOR-OSEM: statistical PET reconstruction from raw line-of-response histograms *Phys. Med. Biol.* **49** 4731–44

Kim H J, Zeeberg B R and Reba R C 1993a Evaluation of reconstruction algorithms in SPECT neuroimaging: I. Comparison of statistical noise in SPECT neuroimages with 'naive' and 'realistic' predictions *Phys. Med. Biol.* **38** 863–73

Kim H-J, Zeeberg B R and Reba R C 1993b Evaluation of reconstruction algorithms in SPECT neuroimaging: II. Computation of deterministic and statistical error components *Phys. Med. Biol.* **38** 881–95

Knuth D E 1969 Seminumerical Algorithms *The Art of Computer Programming* vol 2 (Reading, MA: Addison-Wesley)

Lamare F, Turzo A, Bizais Y, Cheze Le Rest C and Visvikis D 2006 Validation of a Monte Carlo simulation of the Philips Allegro/GEMINI PET systems using GATE *Phys. Med. Biol.* **51** 943–62

Lange K and Carson R 1984 EM reconstruction algorithms for emission and transmission tomography *J. Comput. Assist. Tomogr.* **2** 306–16 PMID: 6608535

Lartizien C, Aubin J B and Buvat I 2010 Comparison of bootstrap resampling Methods for 3-D PET imaging *IEEE Trans. Med. Imaging* **29** 1442–54

Liew S C, Hasegawat B H, Brown J K and Lang T F 1993 Noise propagation in SPECT images reconstructed using an iterative maximum-likelihood algorithm *Phys. Med. Biol.* **38** 1713–26

Louis A K 1996 Approximate inverse for linear and some nonlinear problems *Inverse Problems* **12** 175–90

Louis A K and Maass P 1990 A mollifier method for linear operator equations of the first kind *Inverse Problems* **6** 427–40

Louis A K, Weber T and Theis D 2008 Computing reconstruction kernels for circular 3D cone beam tomography *IEEE Trans. Med. Imaging* **27** 880–6

Mariano-Goulart D, Fourcade M, Bernon J L, Zanca M and Rossi M 2003 Experimental study of stochastic noise propagation in SPECT images reconstructed using the conjugate gradient algorithm Comput *Med. Imaging Graph.* **27** 53–63

Moore S C, Kijewski M F, Müller S P and Holman B L 1988 SPECT image noise power: effects of nonstationary projection noise and attenuation compensation *J. Nucl. Med.* **29** 1704–9 PMID: 2981029

Moore S C, Southekal S, Park M, McQuaid S J, Kijewski M F and Müller S P 2012 Improved regional activity quantitation in nuclear medicine using a new approach to correct for tissue partial volume and spillover effects *IEEE Trans. Med. Imaging* **31** 405–16

Muzic R F, Chen C H and Nelson D 1998 A method to correct for scatter, spillover and partial volume effects in region of interest analysis in PET *IEEE Trans. Med. Imaging* **17** 202–13

Nickerson L D, Narayana S, Lancaster L, Fox P T and Gao J H 2003 Estimation of the local statistical noise in positron emission tomography revisited: practical implementation *NeuroImage* **19** 442–56

Nuyts J 2002 On estimating the variance of smoothed MLEM images *IEEE Trans. Nucl. Sci.* **40** 714–21

Pan X and Metz C E 1995 Analysis of noise properties of a class of exact methods of inverting the 2D exponential Radon transform *IEEE Trans. Med. Imaging* **14** 659–68

Politte D G and Snyder D L 1991 Corrections for accidental coincidences and attenuation in maximum-likelihood image reconstruction for positron-emission tomography *IEEE Trans. Med. Imaging* **10** 82–9

Rico A, Strauss O and Mariano-Goulart D 2009 Choquet integrals as projection operators for quantified tomographic reconstruction *Fuzzy Sets Syst.* **160** 198–211

Shepp L A and Vardi Y 1982 Maximum likelihood reconstruction for emission tomography *IEEE Trans. Med. Imaging* **1** 113–22

Soret M, Koulibaly P M, Darcourt J, Hapdey S and Buvat I 2003 Quantitative accuracy of dopaminergic neurotransmission imaging with ^{123}I SPECT *J. Nucl. Med.* **44** 1184–93 PMID: 12843235

Strauss O, Lahrech A, Rico A, Mariano-Goulart D and Telle B 2009 NIBART: a new interval based algebraic reconstruction technique for error quantification of emission tomography images *MICCAI'09: Int. Conf. on Medical Image Computing and Computer-Assisted Intervention* pp 148–55

Wang W and Gindi G 1997 Noise analysis of MAP-EM algorithms for emission tomography *Phys. Med. Biol.* **42** 2215–32

Wells R G, Gifford H C and King M A 2000 The impact of ROI definition on estimator performance on FBP reconstructed Ga-67 SPECT images *IEEE Trans. Nucl. Sci.* **47** 1210–7

Wilson D W and Tsui B W 1993 Noise properties of filtered-backprojection and ML-EM reconstructed emission tomography images *IEEE Trans. Nucl. Sci.* **40** 1198–203