INSTITUTE OF PHYSICS PUBLISHING PHYSICS IN MEDICINE AND BIOLOGY

Phys. Med. Biol. **51** (2006) 5455–5474 doi:10.1088/0031-9155/51/21/005

Joint estimation of dynamic PET images and temporal basis functions using fully 4D ML-EM

Andrew J Reader[1], Florent C Sureau[2,3], Claude Comtat[2], Régine Trébossen[2] and Irène Buvat[4]

[1] School of Chemical Engineering and Analytical Science, The University of Manchester, PO Box 88, Manchester M60 1QD, UK
[2] Service Hospitalier Frédéric Joliot, CEA/DSV/DRM, Orsay, France
[3] Siemens Medical Solutions, Saint-Denis, France
[4] UMR 678 INSERM-UPMC, CHU Pitié-Salpêtrière, Paris, France

E-mail: andrew.j.reader@manchester.ac.uk

Received 23 May 2006, in final form 28 July 2006
Published 6 October 2006
Online at stacks.iop.org/PMB/51/5455

Abstract

A fully 4D joint-estimation approach to reconstruction of temporal sequences of 3D positron emission tomography (PET) images is proposed. The method estimates both a set of temporal basis functions and the corresponding coefficient for each basis function at each spatial location within the image. The joint estimation is performed through a fully 4D version of the maximum likelihood expectation maximization (ML-EM) algorithm in conjunction with two different models of the mean of the Poisson measured data. The first model regards the coefficients of the temporal basis functions as the unknown parameters to be estimated and the second model regards the temporal basis functions themselves as the unknown parameters. The fully 4D methodology is compared to the conventional frame-by-frame independent reconstruction approach (3D ML-EM) for varying levels of both spatial and temporal post-reconstruction smoothing. It is found that using a set of temporally extensive basis functions (estimated from the data by 4D ML-EM) significantly reduces the spatial noise when compared to the independent method for a given level of image resolution. In addition to spatial image quality advantages, for smaller regions of interest (where statistical quality is often limited) the reconstructed time–activity curves show a lower level of bias and a lower level of noise compared to the independent reconstruction approach. Finally, the method is demonstrated on clinical 4D PET data.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Four-dimensional (4D), or dynamic, positron emission tomography (PET) is able to produce parametric images of physiological function which often have greater clinical utility than conventional static 3D PET images. For example, when using the positron emitter [^{18}F]-fluoro-deoxy-glucose (FDG), images of glucose metabolic rate can be obtained through appropriate kinetic modelling of the time series of 3D images—which is not possible if only a single static 3D PET image is available. However, 4D PET is constantly challenged by the limited statistical quality of the 3D images obtained for each time frame of the temporal image sequence, due to the continuing demand for improved spatiotemporal resolution. This problem of low counts is further accentuated if small regions of interest (ROIs) are selected, from which time–activity curves (TACs) are derived for kinetic analysis and physiological parameter estimation within small structures. However, the principal reason for the poor statistical quality is the use of independent 3D image reconstruction for each time frame, each of which only draws information from a fraction of the acquired data set, ignoring the data both before and after the time frame in question.

This work seeks to treat the reconstruction as a truly 4D problem, whereby an entire 4D (spatiotemporal) distribution of a PET isotope will be reconstructed directly from all of the acquired data. Such a reconstruction requires a choice of spatiotemporal basis functions, and whilst for 3D imaging the choice of spatial basis function has often been intuitively selected as the voxel, the choice of temporal basis function is not obvious. The default choice for the temporal basis functions is a set of non-overlapping top-hat functions (corresponding to a series of frame-by-frame independent 3D reconstructions)—which imposes time-frame independence, and hence significant loss in statistical quality. Use of singular value decomposition (SVD) on the binned emission data can provide a set of useful temporal basis functions (Matthews *et al* 1997), but these functions usually contain negative values, and the precise number to actually use (the 'cut-off') needs to be determined.

In contrast, this work proposes estimation of the temporal basis functions as part of the reconstruction process itself—requiring only that the number of such basis functions be specified in advance. Specifically, a joint-estimation methodology is proposed which alternately estimates (i) the *coefficients* for the temporal basis functions (i.e. estimation of a set of 3D images of coefficients (or 'weights') for the basis functions) and (ii) the *temporal basis functions themselves* which are used in conjunction with these 3D images of coefficients.

Direct reconstruction of 4D images has been considered by other researchers. For example, Carson and Lange (1985) first proposed the expectation maximization (EM) algorithm for direct parametric reconstruction (2D/3D), which was the motivator for the work of Kamasak *et al* (2005) for direct reconstruction of parametric images. Matthews *et al* (1997) looked at reconstruction of 3D image sequences directly, using filtered backprojection and SVD to first determine the temporal basis functions. Gunn *et al* (2002) used a dictionary of basis functions to determine parametric images directly from the data. Nichols *et al* (2002) and Asma and Leahy (2006) used splines for temporal basis functions. Walledge *et al* (2004) considered various inter-frame smoothing strategies. The key difference of this work in relation to all these approaches is the estimation of a limited set of temporal basis functions from the measured emission data, carried out in conjunction with the estimation of the coefficient images (which specify the coefficient for each of the temporal basis functions for each spatial location within the field of view). As two sets of inter-dependent parameters are estimated, the phrase 'joint estimation' is used.

The concept of joint estimation for PET reconstruction has previously been applied to the case of estimation of both the emission activity and the attenuation map (Erdogan and Fessler

2000). Likewise, Nuyts *et al* (1999) developed a simultaneous maximum likelihood method. The methodology proposed here is in the same spirit as those approaches (although applied to a very different estimation problem): one set of unknown parameters is held constant while the other set is optimized, and vice versa.

2. Theory

2.1. Parameterization of the likelihood for a Poisson data model

For an I-dimensional measured data vector $\mathbf{m}$ (where each element m_i is a sample from a Poisson distribution with mean equal to q_i), the likelihood of obtaining vector $\mathbf{m}$ had the mean vector been $\mathbf{q}$ is defined by the product of all the individual Poisson probabilities of obtaining m_i for a given mean q_i:

$$\Pr(\mathbf{m}|\mathbf{q}) = \prod_{i=1}^{I} \exp[-q_i] \frac{q_i^{m_i}}{m_i!}. \tag{1}$$

In 4D PET, the measured data are given by $\mathbf{m}$ (the number of counts contained in each possible measurement element i), which will still be regarded as a single I-D vector, even though the measured spatiotemporal data are usually in 5D (two projection angles (ϕ, θ), two coordinates on the rotated projection plane (x', y') and time t). For the case of PET data, the mean vector $\mathbf{q}$ for the measured data vector $\mathbf{m}$ can be modelled by a system matrix $\mathbf{A}$ operating on a vector of parameters $\boldsymbol{\theta}$:

$$\mathbf{q} = \mathbf{A}\boldsymbol{\theta}. \tag{2}$$

In equation (2), the mean vector $\mathbf{q}$ has been factorized into a matrix $\mathbf{A}$ for which all the elements are assumed to be known (which typically represent and model the known PET measurement process, but the matrix can include more than this, as will be seen) and a vector $\boldsymbol{\theta}$ of unknown parameters defining the particulars of the object being imaged. Continuing in the context of PET, the object is the PET isotope's spatiotemporal (4D) distribution. There is no theoretical restriction on the way that $\mathbf{q}$ has been decomposed into $\mathbf{A}$ and $\boldsymbol{\theta}$, equation (2) is nothing more than a model of the mean of the data. In designing the parameterization of the mean of the data, the only questions to be posed are: (i) is the choice of the vector $\boldsymbol{\theta}$ (when all possible $\boldsymbol{\theta}$ are considered) able to represent all feasible objects? and (ii) is the combined choice of $\mathbf{A}$ and $\boldsymbol{\theta}$ adequate to represent all the feasible vectors $\mathbf{q}$? A chosen decomposition of $\mathbf{q}$ is merely a modelling issue, and not a theoretical or statistical concern. It is this freedom and flexibility in the choice of $\mathbf{A}$ and $\boldsymbol{\theta}$ which will be exploited in this work for estimation of temporal basis functions. In general, the number of parameters used to represent the spatiotemporal object (i.e. the number of elements in $\boldsymbol{\theta}$) does not equal the number of elements in the measured data vector $\mathbf{m}$. The number of elements in $\boldsymbol{\theta}$ will be set to J, making $\mathbf{A}$ an $I \times J$ matrix and $\boldsymbol{\theta}$ a J-D vector.

2.1.1. Choice of the matrix A. For an object representation, a set of spatiotemporal basis functions is often employed (covering the spatial field of view for all the time points of the scan duration), such that $\boldsymbol{\theta}$ specifies the coefficients for each of these spatiotemporal basis functions in the set. In such a case, the matrix $\mathbf{A}$ is factorized according to

$$\mathbf{A} = \mathbf{HB}, \tag{3}$$

where $\mathbf{B}$ is an $S \times J$ matrix of spatiotemporal basis functions, such that each of the J columns contains one such spatiotemporal basis function (each represented discretely as a set of S

spacetime samples, where the sample value could be obtained from a definite integral over a small 4D volume element of the continuous basis function), and $\mathbf{H}$ is an $I \times S$ system matrix, giving the probability that a positron emission from within a given sample location s (in spacetime) gives rise to an event being detected in a given element i of the measured data vector $\mathbf{m}$. A simple choice of $\mathbf{B}$ is the unit matrix $\mathbf{I}$, corresponding to the case where the spatiotemporal basis functions are 4D top-hat functions (3D voxels with independent time frames matching the time sampling of $\mathbf{m}$). With regard to $\mathbf{H}$, the assumption is often made that the system matrix is time invariant, in which case $\mathbf{H}$ would consist of numerous copies of the same static system matrix $\mathbf{H}_{\text{static}}$, duplicated according to the number of time samples in the measured data vector $\mathbf{m}$ (hence, when $\mathbf{B} = \mathbf{I}$, this corresponds to the simple case of a series of separate time-frame reconstructions, each using $\mathbf{H}_{\text{static}}$).

Conventionally, the size and number of voxels (3D top-hat functions) are chosen according to the known spatial resolution limits of the PET scanner, whereas the size and number of time frames (1D top-hat functions) are often chosen so as to integrate over many time sample points in the measured vector $\mathbf{m}$ whilst attempting to not overly compromise the desired temporal resolution. This integration over many time points aims to increase the amount of data in $\mathbf{m}$ which contributes to a given 3D image in the time series—thus improving spatial image quality at the cost of temporal resolution.

Overlapping spatial basis functions can be considered (such as blobs (Lewitt 1992)), as well as overlapping temporal basis functions (such as non-zero-order spline basis functions, as used by Nichols *et al* (2002)). However, these choices appear somewhat arbitrary and cannot usually be regarded as necessarily optimal. There is of course immense flexibility in the choice of the matrix $\mathbf{B}$ and considerable scope for the inclusion of *a priori* information—such that considerable regularization can be included at the modelling stage of the reconstruction process.

In this work, shift-invariant voxels (or clusters of voxels) will be used as the spatial basis functions, on the understanding that the possible benefits of using other spatial basis functions would equally apply to the investigations carried out here. The key innovation in this work is the avoidance of any *a priori* selection of temporal basis functions. Instead, the time-basis functions will be estimated directly from the acquired data vector $\mathbf{m}$.

2.1.2. Choice of $\boldsymbol{\theta}$. $\boldsymbol{\theta}$ is the all-important parameter vector which models the mean vector $\mathbf{q}$ through its description of the object—it is the elements of $\boldsymbol{\theta}$ which will be estimated so as to maximize the likelihood (equation (1)). Often these parameters specify the coefficients of spatiotemporal basis functions to model the object of interest, but, as previously discussed, there is no theoretical reason whatsoever which restricts the elements of $\boldsymbol{\theta}$ to be a set of coefficients for basis functions. In this work, $\boldsymbol{\theta}$ will be used in the conventional way (by specifying the coefficients of temporal basis functions), but also in a separate model $\boldsymbol{\theta}$ will be used to hold a set of parameters which serve to define the temporal basis functions.

2.2. Algorithm to maximize the likelihood

Whatever the choice of $\boldsymbol{\theta}$, an algorithm is needed to estimate the values contained in $\boldsymbol{\theta}$ such that the likelihood (equation (1)) is maximized (for a given fixed measured data set $\mathbf{m}$). First, equation (1) is no longer expressed as a direct function of the mean $\mathbf{q}$, but rather as a function of the chosen set of parameters $\boldsymbol{\theta}$ which describe the mean (through use of equation (2)):

$$\Pr(\mathbf{m}|\boldsymbol{\theta}) = \prod_{i=1}^{I} \exp[-(\mathbf{A}\boldsymbol{\theta})_i] \frac{[(\mathbf{A}\boldsymbol{\theta})_i]^{m_i}}{m_i!}. \tag{4}$$

Taking the natural logarithm (to simplify the derivative which will follow) gives

$$\ln[\Pr(\mathbf{m}|\boldsymbol{\theta})] = \sum_{i=1}^{I} \{-(\mathbf{A}\boldsymbol{\theta})_i + m_i \ln[(\mathbf{A}\boldsymbol{\theta})_i] - \ln m_i!\}. \tag{5}$$

To find a maximum, the partial derivative with respect to each one of the unknown parameters θ_j is taken and set to zero:

$$\frac{\partial}{\partial \theta_j} \ln[\Pr(\mathbf{m}|\boldsymbol{\theta})] = \sum_{i=1}^{I} \left\{ -a_{ij} + \frac{m_i}{(\mathbf{A}\boldsymbol{\theta})_i} a_{ij} \right\} = 0 \qquad \text{for } j = 1, \ldots, J, \tag{6}$$

where $\boldsymbol{\theta} = \{\theta_j\}_J$, $\mathbf{A} = \{a_{ij}\}_{I \times J}$ and $(\mathbf{A}\boldsymbol{\theta})_i = \sum_{j=1}^{J} a_{ij}\theta_j$. Rearranging gives

$$\frac{1}{\sum_{i=1}^{I} a_{ij}} \sum_{i=1}^{I} \left\{ \frac{m_i}{(\mathbf{A}\boldsymbol{\theta})_i} a_{ij} \right\} = 1. \tag{7}$$

Multiplying both sides by θ_j gives

$$\frac{\theta_j}{\sum_{i=1}^{I} a_{ij}} \sum_{i=1}^{I} \left\{ \frac{m_i}{(\mathbf{A}\boldsymbol{\theta})_i} a_{ij} \right\} = \theta_j. \tag{8}$$

Following the description of Barrett and Myers (2003), an iterative algorithm is obtained by replacing $\boldsymbol{\theta}$ by a succession of estimates $\hat{\boldsymbol{\theta}}^k$ and the fixed point iteration procedure is used to obtain the well-known maximum likelihood expectation maximization (ML-EM) algorithm (Shepp and Vardi 1982):

$$\hat{\theta}_j^{k+1} = \frac{\hat{\theta}_j^k}{\sum_{i=1}^{I} a_{ij}} \sum_{i=1}^{I} \left\{ \frac{m_i}{(\mathbf{A}\hat{\boldsymbol{\theta}}^k)_i} a_{ij} \right\}, \tag{9}$$

where the caret denotes a parameter estimate. Equation (9) is a very general algorithm (retaining all the previously mentioned flexibility in the choice of $\mathbf{A}$ and $\boldsymbol{\theta}$) able to offer maximum likelihood estimates of a given set of parameters $\boldsymbol{\theta}$ which have been chosen to model the mean $\mathbf{q}$ of the measured data vector $\mathbf{m}$.

2.3. Alternating optimization

The basic strategy of the proposed fully 4D reconstruction is to use equation (9) to first estimate the coefficients for a set of temporal basis functions (which can have an initially arbitrary form), and then to use equation (9) with a different model of the mean in order to adapt/update the basis functions whilst holding the coefficients constant.

For the first step of the alternating algorithm, the chosen model for the mean data vector $\mathbf{q}$ is

$$\mathbf{q} = \mathbf{H}\mathbf{B}\boldsymbol{\theta}_c, \tag{10}$$

where the subscript c has been added to the parameter vector $\boldsymbol{\theta}$ to indicate that in this case $\boldsymbol{\theta}_c$ specifies the set of *coefficients* for each temporal basis function at each voxel. The number of elements in $\boldsymbol{\theta}_c$ corresponds to the number of spatial (3D) sampling locations (i.e. the number of voxels) multiplied by the number of different temporal basis functions. Hence, $\boldsymbol{\theta}_c$ specifies one coefficient image for each temporal basis function, and so the unique set of coefficients associated with each voxel specifies the temporal behaviour of the radioactivity for that voxel. By way of analogy, $\boldsymbol{\theta}_c$ can be regarded as specifying a set of coefficients for a truncated Fourier series for each and every voxel, giving the TAC for each voxel. However, instead of

the Fourier basis functions, the temporal basis functions will be estimated in the second step of the alternating optimization process.

Using equation (10) to parameterize the mean, from equation (9) an algorithm to obtain a series of estimates $\hat{\boldsymbol{\theta}}_{\mathrm{c}}^{k}$ of the vector of coefficients can be easily obtained:

$$\hat{\boldsymbol{\theta}}_{\mathrm{c}}^{k+1} = \frac{\hat{\boldsymbol{\theta}}_{\mathrm{c}}^{k}}{\mathbf{B}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{1}}\mathbf{B}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\frac{\mathbf{m}}{\mathbf{HB}\hat{\boldsymbol{\theta}}_{\mathrm{c}}^{k}}, \tag{11}$$

where $\mathbf{1}$ is an I-D vector containing a value of one for all elements and k is the iteration number. Matrix–vector multiplication in equation (11) is conventional, but the products and ratios of vectors are performed element by element. Equation (11) is of course nothing more than the familiar ML-EM algorithm of equation (9), but with an overall system matrix given by $\mathbf{HB}$ and the vector of unknowns given by $\boldsymbol{\theta}_{\mathrm{c}}$.

For the second step in the alternating estimation algorithm, the same overall expression for the mean data $\mathbf{q}$ (equation (10)) is used, but the arrangement of the known and unknown parameters within the model is modified in order to place the sampled temporal basis functions into a vector $\boldsymbol{\theta}_{\mathrm{b}}$ of unknown parameters. This requires the coefficients of the basis functions to be known, regarded now as constants within the system matrix component of $\mathbf{q}$. In this case, the decomposition of $\mathbf{q}$ is now given by

$$\mathbf{q} = \mathbf{HC}\boldsymbol{\theta}_{\mathrm{b}}, \tag{12}$$

where the definition of the matrix $\mathbf{C}$ is such that

$$\mathbf{C}\boldsymbol{\theta}_{\mathrm{b}} = \mathbf{B}\boldsymbol{\theta}_{\mathrm{c}}, \tag{13}$$

where both sides of equation (13) yield the same 4D object. As an option, in order to encourage a minimal level of smoothness in the temporal basis functions, the parameterization of each temporal basis function, as given in $\boldsymbol{\theta}_{\mathrm{b}}$, can be used to specify a set of coefficients for a series of shift-invariant sub-basis functions (in the tests below this is only used with one data set). So instead of equation (13), the following relation between the vectors and matrices can be optionally used:

$$\mathbf{CT}\boldsymbol{\theta}_{\mathrm{b}} = \mathbf{B}\boldsymbol{\theta}_{\mathrm{c}}, \tag{14}$$

where $\mathbf{T}$ is a tridiagonal matrix with the same three non-zero values passing down its diagonal (0.25, 0.5, 0.25), defining a three-point kernel (i.e. $\mathbf{T}$ performs a 1D convolution). In effect, for T time sampling points, this corresponds to each temporal basis function being defined by a set of T coefficients for T sub-basis functions, where the sub-basis functions are the three-point kernels. Of course, setting $\mathbf{T}$ equal to the unit matrix $\mathbf{I}$ removes the use of sub-basis functions. When voxels ($v = 1, \ldots, V$) are used for the spatial basis functions, the matrix $\mathbf{C}$ which satisfies equation (14) is given by

$$\mathbf{C} = \begin{bmatrix} \mathrm{diag}\{\theta_{c_{(1,1)}}\} & \mathrm{diag}\{\theta_{c_{(1,2)}}\} & \cdots & \mathrm{diag}\{\theta_{c_{(1,B)}}\} \\ \mathrm{diag}\{\theta_{c_{(2,1)}}\} & \mathrm{diag}\{\theta_{c_{(2,2)}}\} & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \mathrm{diag}\{\theta_{c_{(V,1)}}\} & \cdots & \cdots & \mathrm{diag}\{\theta_{c_{(V,B)}}\} \end{bmatrix}, \tag{15}$$

where $\mathrm{diag}\{\theta_{c_{(v,b)}}\}$ is defined as a square $T \times T$ diagonal matrix with the central diagonal elements all set equal to the scalar value $\theta_{c_{(v,b)}}$ and all other elements zero. The element $\theta_{c_{(j,b)}}$ corresponds to the coefficient of temporal basis function b for voxel v. Hence, the 4D EM algorithm for ML estimation of the parameter vector $\boldsymbol{\theta}_{\mathrm{b}}$ for the temporal basis functions is given by

$$\hat{\boldsymbol{\theta}}_{\mathrm{b}}^{k+1} = \frac{\hat{\boldsymbol{\theta}}_{\mathrm{b}}^{k}}{\mathbf{T}^{\mathrm{T}}\mathbf{C}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{1}}\mathbf{T}^{\mathrm{T}}\mathbf{C}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\frac{\mathbf{m}}{\mathbf{HCT}\hat{\boldsymbol{\theta}}_{\mathrm{b}}^{k}}, \tag{16}$$

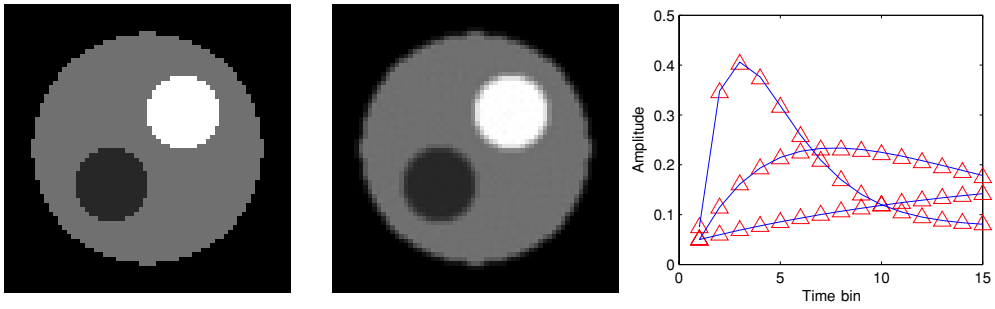

Figure 1. The 3D (2D space with time) simulation phantom. Left: true distribution at time frame 4. Middle: reconstructed distribution at time frame 4. Right: true (solid line) and reconstructed (triangles) TACs for the three regions. The reconstruction used the joint-estimation method with 15 temporal basis functions with random initialization (150 cycles, each cycle consisting of 16 iterations of coefficient estimation and 16 iterations of basis function estimation).

where $\mathbf{T}$ is usually set equal to the unit matrix $\mathbf{I}$. The overall reconstruction procedure consists of running equation (11) for a series of iterations (fixing the temporal basis functions at some initial values), then using the estimate $\boldsymbol{\theta}_c$ to create the $\mathbf{C}$ matrix for use in equation (16). Equation (16) can then be run for a fixed number of iterations to estimate $\boldsymbol{\theta}_b$, which in turn allows creation of a new matrix $\mathbf{B}$ for use in equation (11), and so on. Hence, this method requires up to three parameters to be selected: the number of iterations for each algorithm to use before switching between the two, and then also the number of such cycles to perform overall, although investigations to date indicate very little sensitivity to the number of updates before switching.

3. Methods

The new fully 4D ML-EM based joint-estimation approach to dynamic reconstruction was assessed using two sets of simulated phantom data (2D space with time and then 3D space with time) as well as clinical 4D data and compared with the conventional independent time-frame reconstruction method (3D ML-EM).

3.1. Simulated 3D data and reconstructions

To assess the overall characteristics of the proposed joint-estimation approach, a 3D simulation (2D space with time) was developed. The simulated phantom consisted of three regions, each with a different temporal behaviour (figure 1). Fifteen time frames were used, and 64 $\times$ 64 pixels, using a 2D sinogram (64 $\times$ 64) for each time bin. To encourage high-quality reconstructions on a pixel grid, a spatial basis consisting of groups of five pixels was used: a first-order neighbourhood with weighting factors of 1.0 for the centre pixel and 0.5 for the four neighbouring pixels. The EM algorithm was used to estimate the coefficients for an overlapping set of these identical five-pixel cluster spatial basis functions, each of which was centred on a pixel on the 64 $\times$ 64 grid.

The noise-free sinogram data for the 15 time frames were reconstructed using the conventional independent-frame method (2D ML-EM) and the proposed joint-estimation method. For the joint-estimation method, three different initializations for the temporal basis functions were tested (Gaussians, top-hat functions and random values), each for the case of

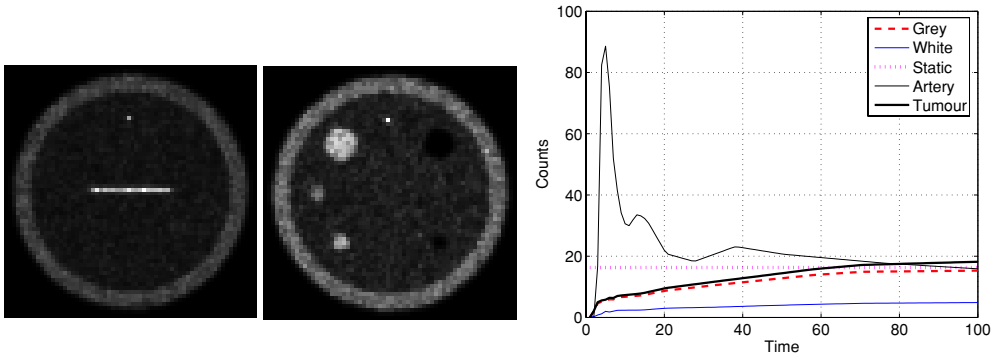

Figure 2. The 4D (3D space with time) simulation phantom. Left: two example transverse slices through the true 3D radioactivity distribution. Right: the true time–activity curves (TACs) used for the Monte Carlo simulation (1 time unit = 12 s). Note the 'counts' is an arbitrary scale, just used to illustrate the simulated TACs.

using either three or 15 temporal basis functions (hence six different reconstructions using the joint method were assessed). The reconstructions used equations (11) and (16) (with $\mathbf{T} = \mathbf{I}$). This group of six reconstructions was designed to test the method's sensitivity to initialization and to the number of temporal basis functions selected (for the extreme cases of using (i) only three temporal basis functions when there are three different TACs in the data and (ii) the maximum of 15 basis functions—when there are 15 time frames in the data). The behaviour of the log-likelihood objective was tracked with iteration number for all the reconstructions.

3.2. Simulated 4D data and reconstructions

A basic Monte Carlo simulation for the high-resolution research tomograph (HRRT) (Wienhard *et al* 2002) was created to generate list-mode data, incorporating the scanner geometry (with detector gaps), but excluding effects such as positron range, attenuation, scatter and randoms. This permitted a *relative* assessment of the fully 4D ML-EM method in comparison to the independent 3D ML-EM method, without the further complication of data corrections. The phantom used for evaluation of the algorithms (figure 2) was a cylinder of length 50 mm and diameter 62 mm, with an additional outer rim of 5 mm thickness (total cylinder diameter 72 mm). The simulated radioactivity in the main cylinder body followed a time–activity curve corresponding to brain white matter and the outer rim followed a time–activity curve corresponding to grey matter. The cylinder contained six spheres of radius 2.5 mm (three were cold spheres (representing ventricles) and three mimicked an arterial TAC) and six spheres of radius 5 mm (again, three were cold and three followed an arterial TAC). A tumour, in the form of a sphere of radius 2.5 mm, was included in the background (white matter) region. Finally, the phantom included two line sources of length 25 mm and diameter 1 mm (one parallel to the scanner axis, displaced by 22.5 mm, and one perpendicular to the axis, displaced by 8 mm from the central transverse plane). The activity of the line sources was held constant in time. Sixty million events were generated over a time interval of 20 min (simulating a mean count rate of 50 kcps). This short time scale and phantom were selected to allow a more computationally practical means of assessing the performance of the reconstruction methods. Images were reconstructed into $64 \times 64 \times 50$ matrices (1.2 mm voxel side length) for 100 time frames (interval 12 s). For the independent reconstruction method, this gave a time series of 100 3D images each of $64 \times 64 \times 50$ voxels, whereas for the 4D method this gave

a set of C coefficient images, each of size 64 × 64 × 50, where C corresponds to the chosen number of temporal basis functions. C was chosen to be 6 (based on previous studies of the number of basis functions (Reader *et al* 2005b)). Note that the joint-estimation (4D) method used equations (11) and (16) (with $\mathbf{T}$ containing the three-point kernel described in section 2.3).

The images obtained from both methods were also assessed for varying levels of post-reconstruction spatial smoothing (convolution with 3D Gaussians of $\sigma = 0.5$, 0.7 and 0.9 (voxels)) and post-reconstruction temporal smoothing (convolution with 1D Gaussians of $\sigma = 0.5$, 1, 1.5, 2, 2.5, 3 and 3.5 (time bin units)).

3.3. Clinical data

To obtain an initial assessment of the methodology on clinical data, 15 min of a 60 min [^{11}C]-PE2I brain scan (carried out with the HRRT) were reconstructed using the alternating optimization methodology. Fifteen time frames, each of 1 min duration, were used, with six temporal basis functions and hence also six images of coefficients (each of size 170 × 170 × 208, 1.2 mm voxels). During the first 15 min of the scan, 287 million events were acquired. For each 1 min frame, the scatter sinograms were estimated using the single scatter simulation method (Watson *et al* 1996) and the randoms were estimated using smoothed delayed coincidence window data (Byars *et al* 2005). A model of the point response of the system was included as an image-space convolution matrix $\mathbf{R}$, in order to achieve resolution recovery and noise reduction (Reader *et al* 2002, 2005a). The shift-invariant kernel used to construct $\mathbf{R}$ was an offset mono-exponential ($\beta + \exp(-\alpha r)$, see Reader *et al* (2005a) for more details). To accelerate reconstruction, 16 subsets were used, such that for each time frame under consideration just one sixteenth of the events were used for any given update of the parameters being estimated. The first part of the alternating algorithm was implemented according to

$$\hat{\theta}_{\mathrm{c}}^{k+1} = \frac{\hat{\theta}_{\mathrm{c}}^{k}}{\mathbf{B}^{\mathrm{T}}\mathbf{R}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{N}^{\mathrm{T}}\mathbf{L}^{\mathrm{T}}\mathbf{1}}\mathbf{B}^{\mathrm{T}}\mathbf{R}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{N}^{\mathrm{T}}\mathbf{L}^{\mathrm{T}}\frac{\mathbf{m}}{\mathbf{LNXRB}\hat{\theta}_{\mathrm{c}}^{k} + \mathbf{Ns} + \mathbf{r}}, \tag{17}$$

where $\mathbf{R}$ is the resolution modelling convolution matrix for image space (assumed shift invariant in this initial assessment), $\mathbf{L}$ is a diagonal matrix containing the inverse of the attenuation correction factors, $\mathbf{N}$ is a diagonal matrix containing the inverse of the normalization correction factors, $\mathbf{X}$ is the 3D x-ray transform and $\mathbf{s}$ and $\mathbf{r}$ are the vectors containing the unique scatter and the randoms estimate in measurement space for each of the 15 time frames ($\mathbf{L}$, $\mathbf{N}$, $\mathbf{X}$ and $\mathbf{R}$ are assumed to be time invariant). Note that $\mathbf{Ns}$ is used, as the single scatter simulation method produces normalized scatter $\mathbf{s}$. Combining the scatter and randoms vectors ($\mathbf{Ns}$ and $\mathbf{r}$) into a single offset vector $\mathbf{b}$, and then recognizing the cancellation of the diagonal $\mathbf{L}$ and $\mathbf{N}$ matrices, one obtains

$$\hat{\theta}_{\mathrm{c}}^{k+1} = \frac{\hat{\theta}_{\mathrm{c}}^{k}}{\mathbf{B}^{\mathrm{T}}\mathbf{R}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{w}}\mathbf{B}^{\mathrm{T}}\mathbf{R}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\frac{\mathbf{m}}{\mathbf{XRB}\hat{\theta}_{\mathrm{c}}^{k} + \frac{\mathbf{b}}{\mathbf{w}}}, \tag{18}$$

where $\mathbf{w}$ is the vector obtained from the matrix–matrix–vector product $\mathbf{LN1}$ ($= \mathbf{N}^{\mathrm{T}}\mathbf{L}^{\mathrm{T}}\mathbf{1}$). The second and final part of the alternating estimation method is correspondingly given by

$$\hat{\theta}_{\mathrm{b}}^{k+1} = \frac{\hat{\theta}_{\mathrm{b}}^{k}}{\mathbf{C}^{\mathrm{T}}\mathbf{R}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{w}}\mathbf{C}^{\mathrm{T}}\mathbf{R}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\frac{\mathbf{m}}{\mathbf{XRC}\hat{\theta}_{\mathrm{b}}^{k} + \frac{\mathbf{b}}{\mathbf{w}}}, \tag{19}$$

which estimates the temporal basis functions (note that for the measured data sub-basis functions were not used, i.e. $\mathbf{T} = \mathbf{I}$ in equation (16)). The overall joint-estimation algorithm

consisted of repeated application of equations (18) and (19), performing 16 updates of the parameters each time (either the coefficients or the basis functions), corresponding to the 16 different subsets of the data which were available for each 1 min time frame.

3.4. Evaluation of 4D simulation results

For the 4D simulation phantom data, the spatial properties of the reconstruction methods were evaluated for the mid-point time frame 50 (10 min). The image for this point in time was used to assess background noise level, using a simple spatial standard deviation figure of merit (FOM), which calculated the standard deviation of the voxel values belonging to the background region (white matter) of the cylinder, divided by the mean value. The resolution of the two line sources was also assessed, by summing all the profiles for all the length of each line source. This allowed an average full width at half maximum (FWHM) resolution measure to be made for the line source. These spatial noise and spatial resolution measures were performed for images from both methods obtained at two iteration numbers (early on in the iterative process and at the end of reconstruction), for three post-reconstruction spatial-smoothing levels (3D Gaussians of $\sigma = 0.5$, 0.7 and 0.9 voxels).

The temporal properties of the images were assessed using an average bias FOM, evaluated for each region of interest (ROI) (white matter (background), grey matter (rim), artery (sphere), ventricle (sphere) and tumour (sphere)). This FOM determined the mean of the differences at each time point between the reconstructed TAC and the corresponding true TAC. The resulting mean value over all time points gave an overall indication of bias for the entire TAC. As an indicator of the noise for each point within each TAC, the spatial standard deviation of the voxel values within the ROI contributing to a given point on a TAC was determined. These standard deviations (calculated as the standard error of the mean), each normalized by the region mean value and expressed as a percentage, were then summed and normalized for the number of contributions, to give an overall mean noise indicator for a given TAC. These bias and noise assessments were made for every ROI, for both reconstruction techniques and for seven different levels of post-reconstruction temporal smoothing.

4. Results

4.1. Simulated data (3D) results

Figure 1 shows an example reconstruction of the phantom and the TACs. Figures 3 and 4 show the converged temporal basis functions for the three different initializations and for the choice of three or 15 basis functions. It is evident that the ML solution for the joint estimation of the coefficients and temporal basis functions is not unique, as the basis functions found for each different initialization are different (there are only minor differences when estimating three temporal basis functions, but major differences for the estimation of 15 basis functions). For noise-free data, the basis functions converge to temporally extensive functions and the same log likelihood (see figure 5) for all cases considered (whether three or 15 basis functions) and for all initializations tested. (Note that for the case of initializing with top-hat functions, the tails adjacent to the top hat itself were set to 10% of the peak top-hat value.) For noisy data, benefits are only seen when fewer than 15 basis functions are used (since with 15 basis functions the number of parameters to estimate matches that of the 15-frame independent reconstruction, yielding 15 top-hat temporal basis functions at convergence, giving the same log likelihood as the independent reconstruction). For three temporal basis functions with noisy data, temporally extensive basis functions are obtained, which allow each time frame

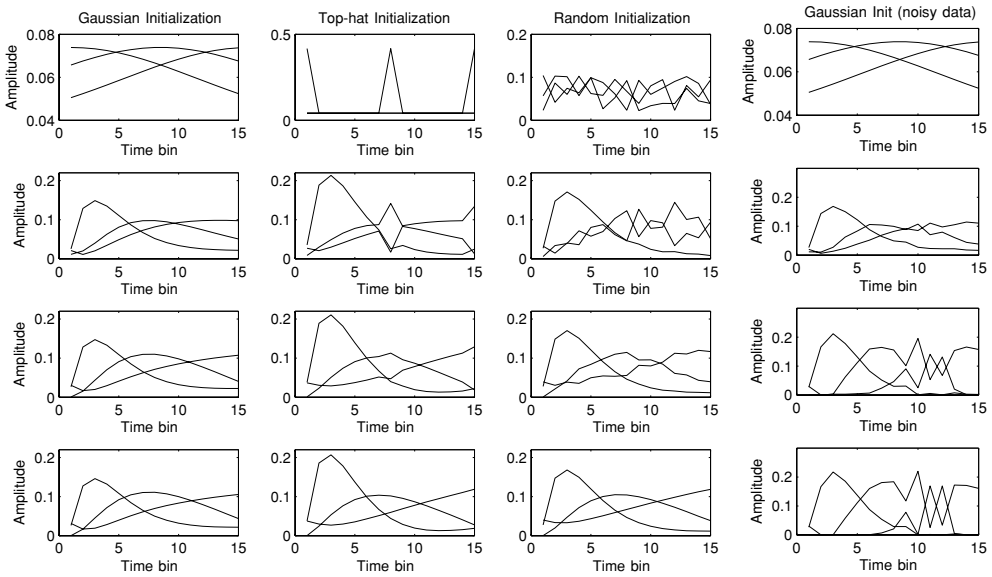

Figure 3. Estimated set of *three* temporal basis functions for the 2D space with time simulation data set for four different reconstructions (corresponding to noise-free data with three different initializations (Gaussian, top hat and random) and noisy data with Gaussian initialization). Top row: cycle 0 of the joint-estimation method (i.e. the initial temporal basis functions are shown). Rows 2, 3 and 4 correspond to cycles 15, 75 and 150, respectively (by cycle 150 the joint-estimation method has converged).

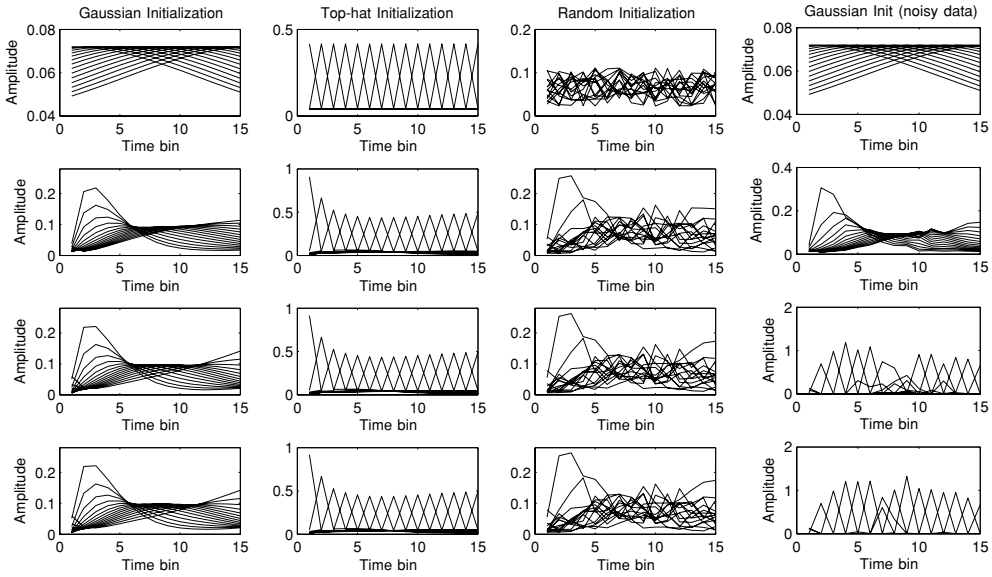

Figure 4. Estimated set of fifteen temporal basis functions for the 2D space with time simulation data set for four different reconstructions (corresponding to noise-free data with three different initializations (Gaussian, top hat and random) and noisy data with Gaussian initialization). Top row: cycle 0 of the joint-estimation method (i.e. the initial temporal basis functions are shown). Rows 2, 3 and 4 correspond to cycles 15, 75 and 150, respectively (by cycle 150 the joint-estimation method has converged).

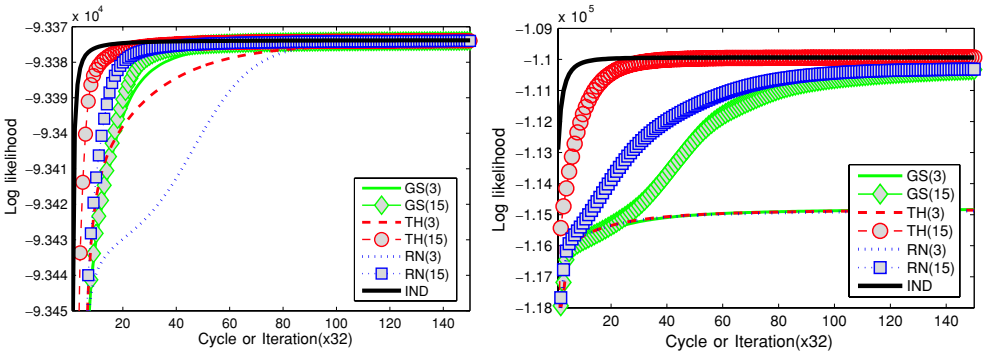

Figure 5. Log likelihood for noise-free data (left) and for noisy data (right). The joint-estimation method is assessed for three different initializations (Gaussians (GS), top hats (TH) and random values (RN)) and for the choice of three or 15 temporal basis functions. The independent time-frame method is also included. Since one cycle for the joint-estimation method included 16 iterations of coefficient estimation and 16 iterations of basis function estimation, the independent method is shown for multiples of 32 iterations.

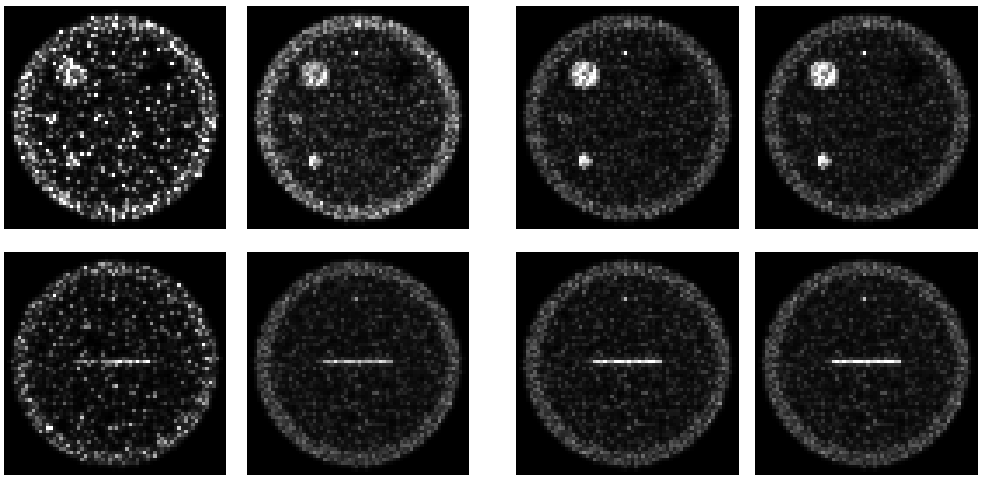

Figure 6. Example transverse slices of the simulated 3D phantom. Column 1: 3D ML-EM reconstruction for time frame 50 (out of 100). Column 2: 4D ML-EM reconstruction for time frame 50 (out of 100). Column 3: 3D ML-EM reconstruction *summed* over all 100 time frames. Column 4: 4D reconstruction *summed* over all 100 time frames. All images were post-reconstruction smoothed spatially, using 3D convolution with a Gaussian of $\sigma = 0.5$ voxels, which yields approximately the same level of image resolution (as measured by the line source, see figure 7) for both reconstruction methods (for 24 cycles of the alternating 4D algorithm and 50 iterations of the independent method). Note that in this case, one cycle consists of either 16 updates of the images of coefficients or 16 updates of the temporal basis functions.

to benefit from many neighbouring frames. The results from the noise-free data indicate that regularizing the noisy-data reconstruction through use of just three temporal basis functions in the system modelling (a reduction in the number of unknowns), whilst giving a lower log likelihood than the independent reconstruction, does not introduce bias for the case of three TACs. Furthermore, considering row 2 of figures 3 and 4, it can be seen that the temporal basis

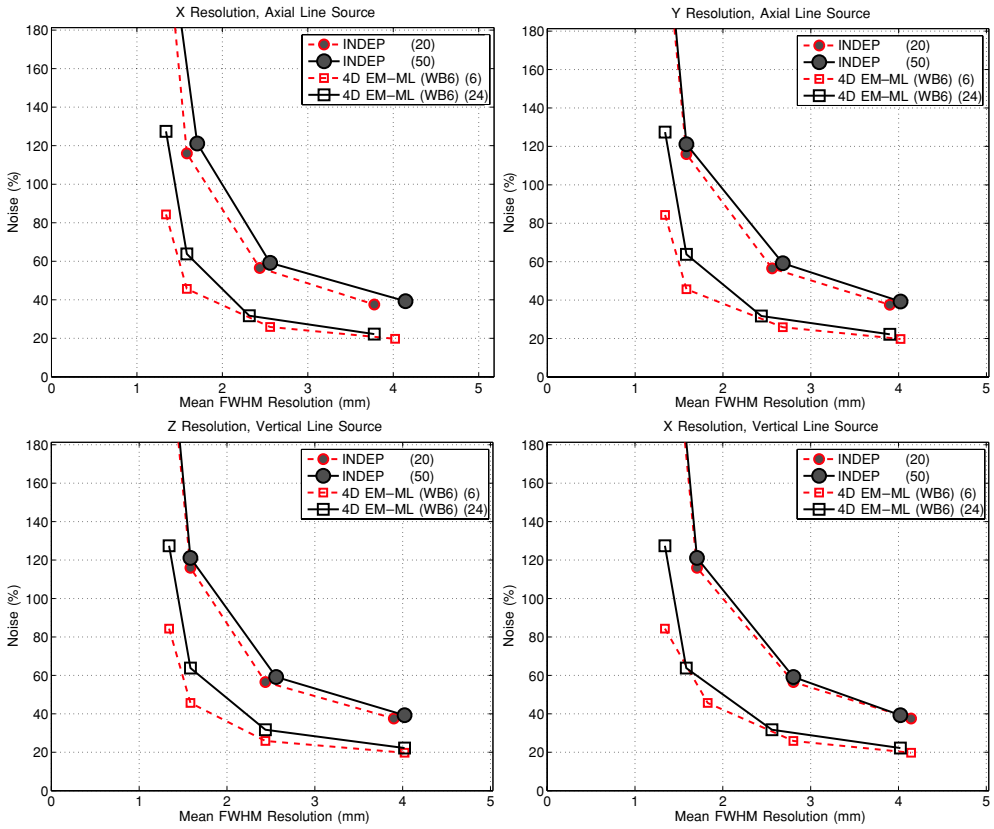

Figure 7. Noise–resolution trade-off for the independent (INDEP) time-frame reconstruction method (3D EM-ML, shown for iterations 20 and 50) and for the fully 4D reconstruction method (4D EM-ML). 4D EM-ML is shown for six and 24 cycles of the alternating method (where in this case one cycle consists of either 16 iterations for the estimation of the coefficients/weights (W) or 16 iterations for the estimation of the six bases (B6)). Both reconstruction methods are shown for four different levels of post-reconstruction smoothing (no smoothing and convolution with a 3D Gaussian, $\sigma = 0.5$, 0.7 and 0.9 voxels) to obtain the differing levels of spatial image noise. The line source diameter was 1 mm.

functions estimated from the noisy data match well with those estimated for the noise-free data, suggesting that noisy-data reconstructions can be regularized through early stopping of the iterative process.

Figure 5 indicates that the joint-estimation method, with its initialization-dependent time-basis functions, achieves the same log likelihood as the conventional independent-frame ML-EM method for the case of noise-free data, emphasizing the non-uniqueness of the ML estimate. Hence, for the case of a limited number of temporal behaviour patterns, there are many ways to *accurately* parameterize the mean of the data $\mathbf{q}$, in such a way that no bias is introduced. Hence, it is clear that the proposed joint-estimation methodology with temporally extensive initialization of the basis functions is able to exploit this non-uniqueness by choosing a more favourable (and yet still bias free) parameterization of the mean of the data, $\mathbf{q}$, such that a greater level of temporal integration occurs compared to the conventional independent-frame reconstruction.

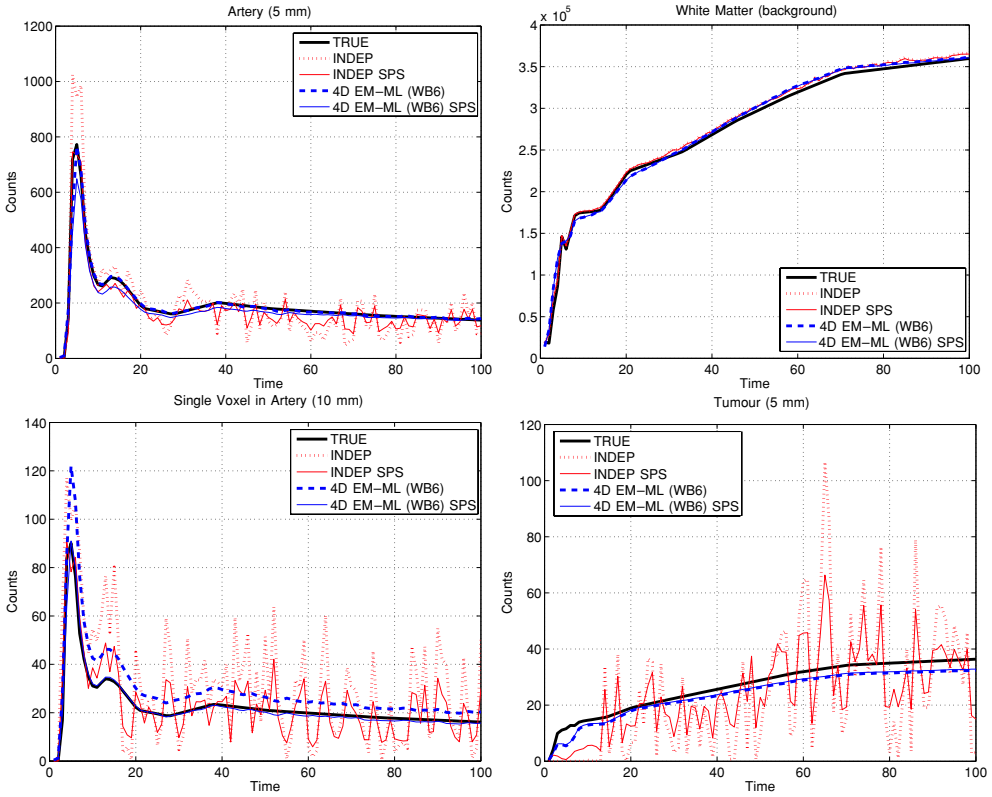

Figure 8. TACs for the two methods with and without post-reconstruction spatial smoothing. The TACs shown correspond to those from 50 iterations of the independent method (3D EM-ML) and 24 cycles (each of 16 iterations for both the weights and the bases) for the 4D EM-ML method. The TACs from both reconstruction methods are also shown for the case where a spatial post-reconstruction smooth (SPS) with a $\sigma = 0.5$ 3D Gaussian kernel has first been carried out (1 time unit = 12 s).

4.2. Simulated data (4D) results

Figure 6 shows example transverse slices (from 3D images of a single time frame) of the independent 3D ML-EM reconstruction and the 4D ML-EM reconstruction. The 4D method demonstrates visually reduced image noise. Figure 6 also shows the same slices for the case where all 100 time frames have been summed, which indicates that the differences between the independent (3D) ML-EM and the 4D method for the time-summation case are minimal. The similarity between the time-frame summed slices and the single time-frame slices for the 4D approach indicates the significant temporal data integration achieved by the 4D method—retaining a level of image quality comparable to the static (summed) case, but retaining the required temporal information (as will be shown). This result arises from the use of coefficient images with the 4D method: each coefficient image voxel specifies the coefficient for a temporal basis function and each of these basis functions is able to extend through all time frames. Figure 7 gives the resolution-noise trade-off for the independent (3D) and the 4D method, for two different iteration numbers and for three levels of spatial post-reconstruction smoothing. It is clear that the 4D reconstruction method consistently gives markedly lower

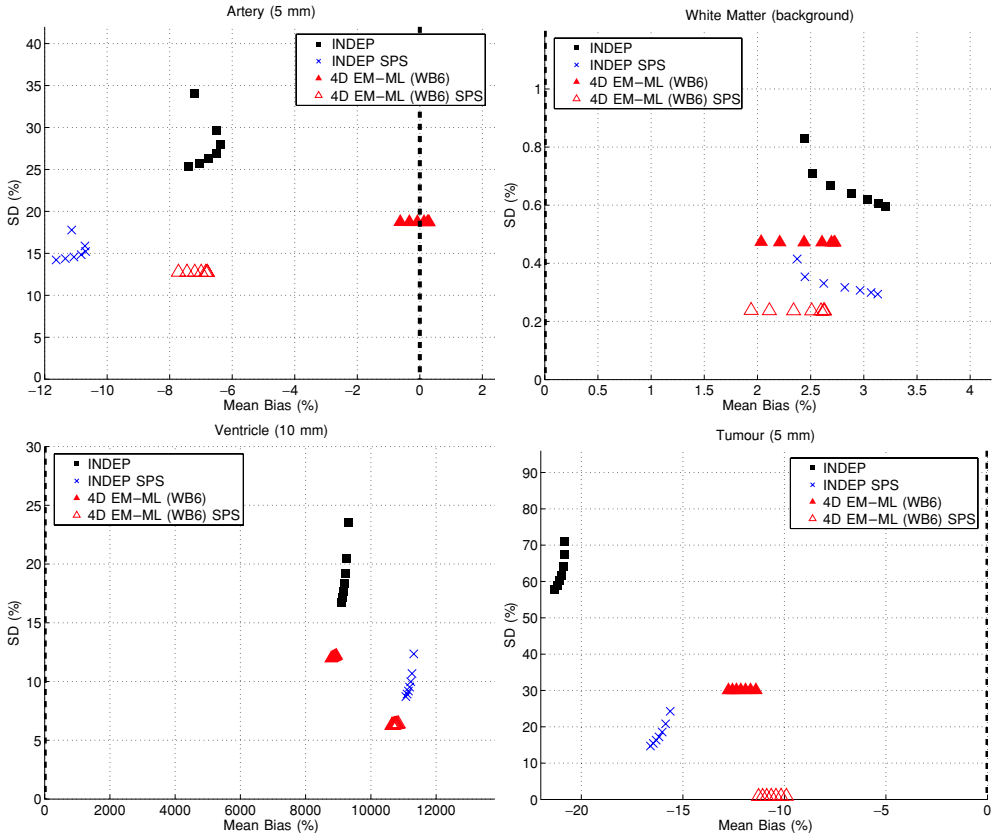

Figure 9. Mean bias and mean noise for the two methods (with and without spatial post-reconstruction smoothing (SPS)) for seven levels of post-reconstruction temporal smoothing. The temporal smoothing was carried out by 1D Gaussian kernel convolution of each voxel's TAC, using Gaussians of $\sigma = 0.5$, 1, 1.5, 2, 2.5, 3 and 3.5 (time bin units).

spatial noise levels (often more than a factor of 2 better) for a given spatial resolution when compared to the independent time-frame reconstruction, supporting the visual evidence of figure 6.

Figure 8 shows the TACs for all methods for a selection of regions—the 4D method outperforms the independent approach, giving less noisy and more accurate TACs in all cases. This visual assessment is confirmed in figure 9, which shows that both the mean bias and the mean noise level in each TAC are lower when using the 4D reconstruction, for all of the ROIs. Note that post-reconstruction temporal smoothing only modifies the bias of the TACs for the 4D method (the triangles in figure 9 only displace horizontally on the graphs), since the temporal noise of the TACs is already very low through use of just six time-basis functions which extend through all time frames. Hence, any temporal smoothing of the 4D reconstruction only gives rise to increasing bias, and so only spatial smoothing should be considered for the 4D joint-estimation methodology. Figure 10 shows the development of the six temporal basis functions with respect to iteration and figure 11 shows the six images of coefficients along with their corresponding temporal basis functions for the fully 4D method.

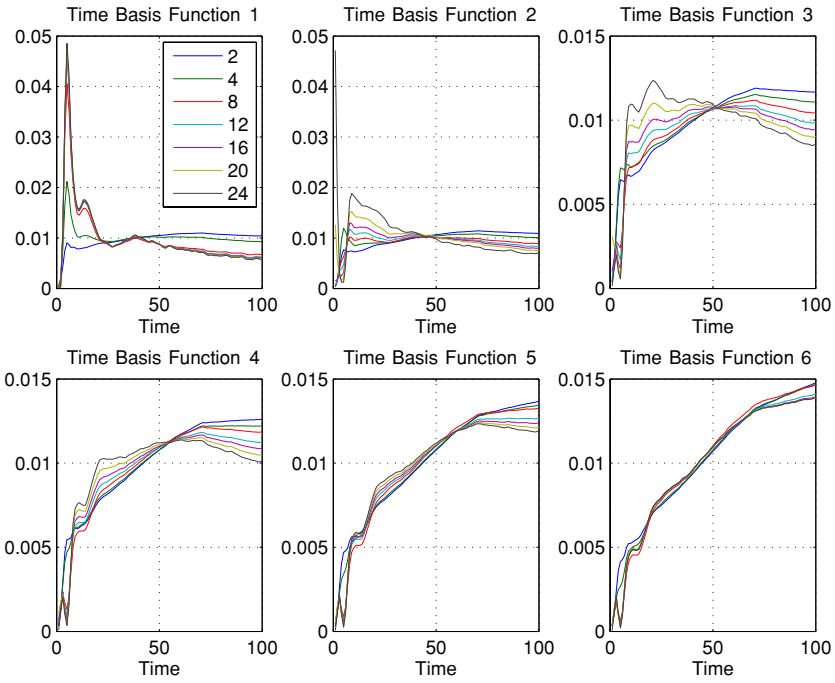

Figure 10. The six temporal basis functions after 2, 4, 8, 12, 16, 20 and 24 cycles. In this case, one cycle consisted of either 16 iterations for the coefficients estimation or 16 iterations for the basis function estimation. Initialization was with six Gaussian functions, centred at six evenly spaced points in time.

4.3. Clinical data results

Figure 12 shows the reconstructions of frame 15 (60 s of data, between the 14th and 15th min of acquisition) for the 4D joint-estimation method and for the independent method. It is apparent that the 4D method significantly reduces noise, through weighted integration of all of the data permitted by the temporally extensive basis functions. Figure 13 compares the TACs obtained by the two approaches, demonstrating again the noise reduction which is achieved by the fully 4D approach. Figure 14 compares the temporal basis functions obtained when estimating six or 15 such functions. Even after 60 cycles, when using 15 temporal basis functions, all time frames are covered by the estimated temporal basis functions, meaning that similar levels of noise reduction are observed at this number of cycles, whether estimating six or 15 basis functions. (This is confirmed by figure 12, which showed very similar results for six and 15 basis functions after 24 cycles.) These results suggest that if the joint-estimation algorithm is terminated before convergence, choosing the maximum number of temporal basis functions is the best option. However, if run until convergence, fewer basis functions than frames must be chosen to achieve noise reduction.

5. Discussion

The proposed joint-estimation fully 4D reconstruction algorithm is based on a change of parameterization of the mean of the measured PET data vector. Conventionally, time frames are treated independently (analogous to a choice of top-hat functions for the temporal basis) or

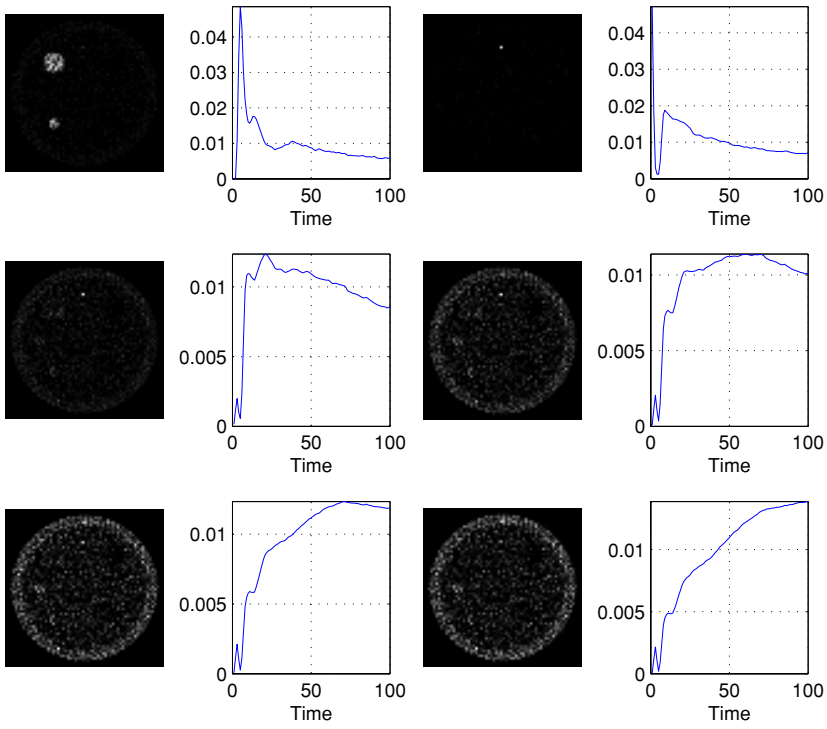

Figure 11. The six coefficient images and their respective time-basis functions after 24 cycles.

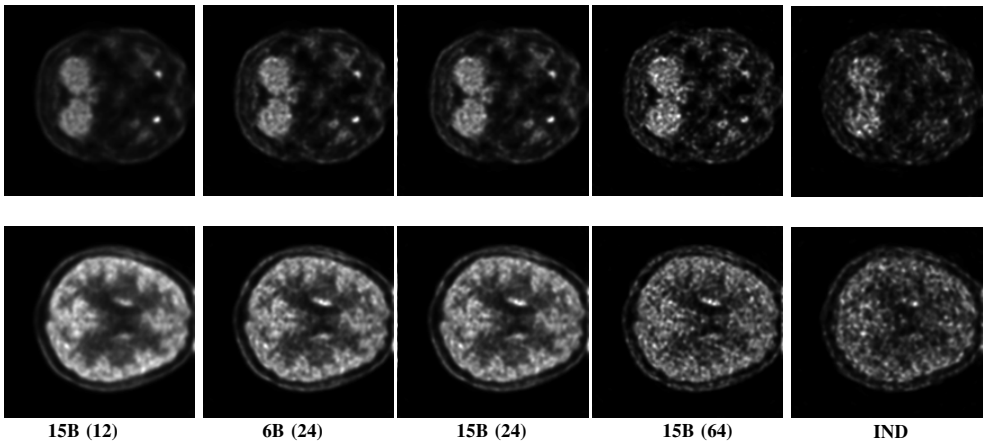

Figure 12. Single slices of frame 15 of the reconstructions. The first four columns correspond to the joint-estimation 4D ML-EM method when using 15 bases (12 cycles), 6 bases (24 cycles), 15 bases (24 cycles), 15 bases (64 cycles) and the final column corresponds to the independent reconstruction (6 iterations, 16 subsets). (The joint-estimation method also used 16 subsets.)

completely summed (static reconstruction—analogous to a choice of a flat constant function for the temporal basis). In contrast, the proposed approach offers an intermediate option between these two extremes, which is able to preserve the temporal behaviour whilst simultaneously improving the signal to noise through summation of data frames—provided there is only

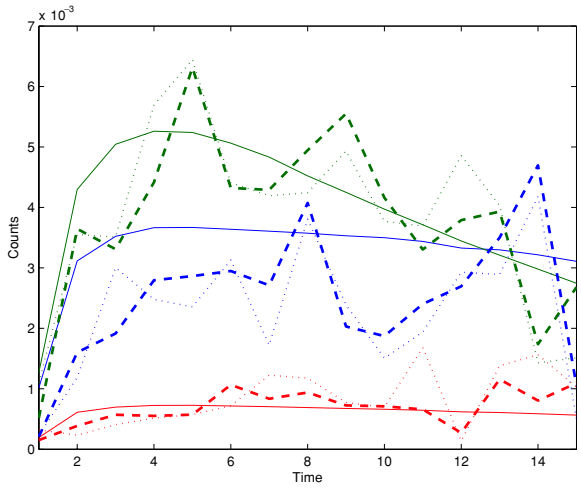

Figure 13. TACs for three single voxels for the 4D joint-estimation method (solid line), the independent method (bold dashed line) and the independent method if no resolution model is used (dotted line). Twenty-four cycles were used, with six temporal basis functions for the 4D method.

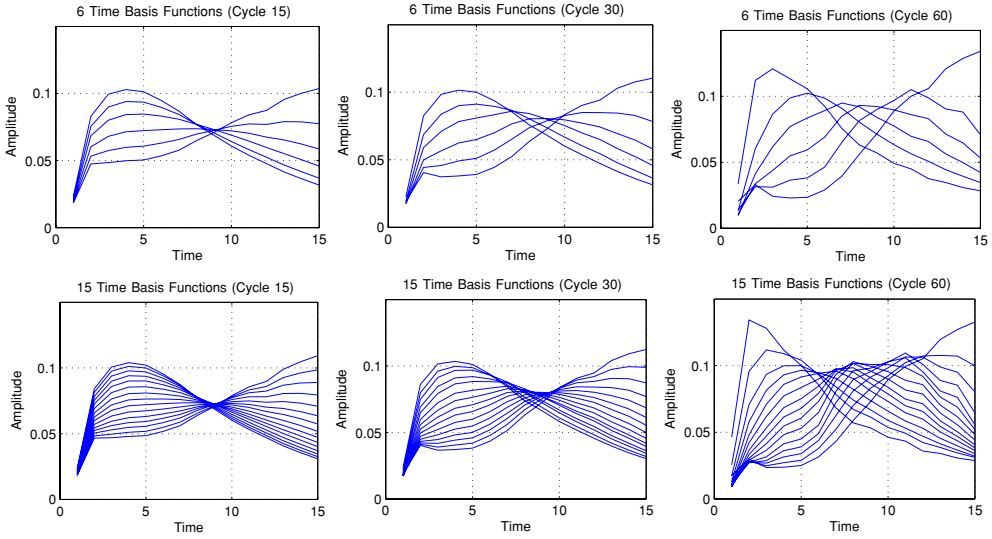

Figure 14. Dependence on the number of temporal basis functions for the clinical data. Top row: the six basis functions obtained after 15, 30 and 60 cycles, respectively. Bottom row: the case where the 4D method is used with 15 basis functions. Since each basis function is initialized as a shifted Gaussian with the sum of counts equal to 1, and since the coefficient images are estimated as the first step in the alternating 4D estimation, the amplitudes for the two different cases of using six and 15 basis functions are virtually the same. For the case of 15 basis functions, this means that the 15 coefficient images have correspondingly smaller values to compensate.

a limited set of TAC types present in the measured data. The approach exploits the fact that accurate modelling of the mean of the measured data can in fact be done in more than one way and uses this redundancy to find a parameterization of the mean which integrates more of the temporal data together. At present, the technique does not exploit modelling of

tracer kinetics, rendering the method suitable for studies where the precise tracer kinetics are unknown. The estimated temporal basis functions obtained with the current method should not be assigned physiological meaning, since the resulting functions are non-unique, dependent on the initialization and the number of functions selected.

In its current form, the proposed joint-estimation approach is computationally demanding, taking over an order of magnitude more time to reconstruct a 4D image compared to the independent time-frame reconstruction approach. In terms of memory requirements, the new method is not significantly more demanding than the conventional approach—although special measures need to be taken with regard to data corrections, as all time frames are accessed for a single iterative update of the coefficients or temporal basis functions. In this work, all the required scatter and random values ($\mathbf{b}/\mathbf{w}$ in equations (18) and (19)) were stored in a separate list-mode file, to avoid the need for random access to fully 3D sinogram data sets for every single time frame. It should also be noted that the use of variable frame lengths is of course compatible with the methodology.

There is evidently much scope for accelerating the methodology, through use of subsets, over-relaxation and even a simultaneous updating procedure. A further possibility is the implementation of the algorithm in image space, using the independent-frame reconstruction as the source data, to determine a set of coefficient images and temporal basis functions directly in image space which model the result of the independent-frame reconstruction.

6. Conclusion

A new approach to 4D reconstruction is proposed, which estimates both a set of temporal basis functions and the coefficients for each of these functions at each spatial location. The alternating estimation method, which uses a fully 4D ML-EM algorithm, permits each time frame to benefit from substantially more of the measured data than the conventional independent (3D ML-EM) time-frame approach, leading to image quality advantages in both the space and time dimensions. This benefit is drawn from the assumption that a limited number of time-basis functions are sufficient to describe the temporal behaviour of the radioactivity in each voxel. This level of improvement cannot be matched by spatiotemporal smoothing strategies, which lower noise at the cost of increasing bias in the images.

References

Asma E and Leahy R M 2006 Mean and covariance properties of dynamic PET reconstructions from list mode data *IEEE Trans. Med. Imaging* **25** 42–54

Barrett H H and Myers K J 2003 *Foundations of Image Science* (New York: Wiley)

Byars L G, Sibomana M, Burbar Z, Jones J, Panin V, Barker W C, Liow J-S, Carson R E and Michel C 2005 Variance reduction on randoms from delayed coincidence histograms for the HRRT *Proc. IEEE Nucl. Sci. Symp. and Medical Imaging Conf. (Puerto Rico, 2005)* pp 2622–6

Carson R E and Lange K 1985 The EM parametric image reconstruction algorithm *J. Am. Stat. Assoc.* **80** 20–2

Erdogan H and Fessler J A 2000 Algorithms for joint estimation of attenuation and emission images in PET *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing* vol 6 pp 3783–6

Gunn R N, Gunn S R, Turkheimer F E, Aston J A D and Cunningham V J 2002 Positron emission tomography compartmental models: a basis pursuit strategy for kinetic modeling *J. Cereb. Blood Flow Metab.* **22** 1425–39

Kamasak M E, Bouman C A, Morris E D and Sauer K 2005 Direct reconstruction of kinetic parameter images from dynamic PET data *IEEE Trans. Med. Imaging* **24** 636–50

Lewitt R M 1992 Alternatives to voxels for image representation in iterative reconstruction algorithms *Phys. Med. Biol.* **37** 705–15

Matthews J, Bailey D, Price P and Cunningham V 1997 The direct calculation of parametric images from dynamic PET data using maximum-likelihood iterative reconstruction *Phys. Med. Biol.* **42** 1155–73

Nichols T E, Qi J, Asma E and Leahy R M 2002 Spatiotemporal reconstruction of list-mode PET data *IEEE Trans. Med. Imaging* **21** 396–404

Nijran K S and Barber D C 1985 Towards automatic analysis of dynamic radionuclide studies using principal components factor analysis *Phys. Med. Biol* **30** 1315–25

Nuyts J, Dupont P, Stroobants S, Benninck R, Mortelmans L and Suetens P 1999 Simultaneous maximum *a posteriori* reconstruction of attenuation and activity distributions from emission sinograms *IEEE Trans. Med. Imaging* **18** 393–403

Reader A J, Ally S, Bakatselos F, Manavaki R, Walledge R, Jeavons A P, Julyan P J, Zhao S, Hastings D L and Zweit J 2002 One-pass list-mode EM algorithm for high-resolution 3D PET image reconstruction into large arrays *IEEE Trans. Nucl. Sci.* **49** 693–9

Reader A J, Sureau F C, Comtat C, Buvat I and Trebossen R 2005a List-mode reconstruction with system modelling derived from projections *Proc. IEEE Nucl. Sci. Symp. and Medical Imaging Conf. (Puerto Rico, 2005)* pp 1832–6

Reader A J, Sureau F C, Comtat C, Trebossen R and Buvat I 2005b Direct fully 4D list-mode reconstruction with temporal basis function estimation *Proc. IEEE Nucl. Sci. Symp. and Medical Imaging Conf. (Puerto Rico, 2005)* pp 1955–9

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

Walledge R J, Manavaki R, Honer M and Reader A J 2004 Inter-frame filtering for list-mode EM reconstruction in high resolution 4D PET *IEEE Trans. Nucl. Sci.* **51** 705–11

Watson C C, Newport D and Casey M E 1996 A single-scatter simulation technique for scatter correction in 3D PET *Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine* ed P Grangeat and J L Amans (Dordrecht: Kluwer)

Wienhard K *et al* 2002 The ECAT HRRT: performance and first clinical application of the new high resolution research tomograph *IEEE Trans. Nucl. Sci.* **49** 104–10