



Original paper



3D TOF-PET image reconstruction using total variation regularization

L. Raczyński^{a,*}, W. Wiślicki^a, K. Klimaszewski^a, W. Krzemień^b, P. Kopka^a, P. Kowalski^a, R. Y. Shopa^a, M. Bała^c, J. Chhokar^c, C. Curceanu^d, E. Czerwiński^c, K. Dulski^c, J. Gajewski^e, A. Gajos^c, M. Gorgol^f, R. Del Grande^f, B. Hiesmayr^g, B. Jasińska^f, K. Kacprzak^c, L. Kapłon^c, D. Kisielewska^c, G. Korcyl^c, T. Kozik^c, N. Krawczyk^c, E. Kubicz^c, M. Mohammed^{c,h}, S. z. Niedźwiecki^c, M. Pałka^c, M. Pawlik-Niedźwiecka^c, J. Raj^c, K. Rakoczy^c, A. Ruciński^e, S. Sharma^c, S. Shivani^c, M. Silarski^c, M. Skurzok^{c,d}, E.L. Stepień^c, B. Zgardzińska^f, P. Moskal^c

^a Department of Complex Systems, National Centre for Nuclear Research, 05-400 Otwock-Świerk, Poland

^b High Energy Physics Division, National Centre for Nuclear Research, 05-400 Otwock-Świerk, Poland

^c Marian Smoluchowski Institute of Physics, Jagiellonian University, 31-348 Cracow, Poland

^d INFN, Laboratori Nazionali di Frascati, 00044 Frascati, Italy

^e Institute of Nuclear Physics PAN, Cracow, Poland

^f Institute of Physics, Maria Curie-Skłodowska University, 20-031 Lublin, Poland

^g Faculty of Physics, University of Vienna, 1090 Vienna, Austria

^h Department of Physics, College of Education for Pure Sciences, University of Mosul, Mosul, Iraq

ARTICLE INFO

Keywords:

Total variation regularization
Positron emission tomography
Time-of-flight
Medical imaging

ABSTRACT

In this paper we introduce a semi-analytic algorithm for 3-dimensional image reconstruction for positron emission tomography (PET). The method consists of the back-projection of the acquired data into the most likely image voxel according to time-of-flight (TOF) information, followed by the filtering step in the image space using an iterative optimization algorithm with a total variation (TV) regularization. TV regularization in image space is more computationally efficient than usual iterative optimization methods for PET reconstruction with full system matrix that use TV regularization. The efficiency comes from the one-time TOF back-projection step that might also be described as a reformatting of the acquired data. An important aspect of our work concerns the evaluation of the filter operator of the linear transform mapping an original radioactive tracer distribution into the TOF back-projected image. We obtain concise, closed-form analytical formula for the filter operator. The proposed method is validated with the Monte Carlo simulations of the NEMA IEC phantom using a one-layer, 50 cm-long cylindrical device called Jagiellonian PET scanner. The results show a better image quality compared with the reference TOF maximum likelihood expectation maximization algorithm.

1. Introduction

Time-Of-Flight (TOF) Positron Emission Tomography (PET) has a long history that started in the early 1980s [1,2]. However, the detectors available at that time did not provide the required combination of stopping power, time resolution and light output efficiency necessary for building clinical systems. The introduction of fast scintillators such as LaBr₃:Ce and LSO [3,4] led to the first commercially available TOF-PET system. State-of-the-art TOF-PET scanners operate at a coincidence resolving time (CRT) of the order of 300–400 ps [5,6]. More recent theoretical and experimental studies using small 2 × 2 × 3 mm³ LSO

and LaBr₃:Ce crystals indicate that the CRT limit is expected at about 100 ps [7,8]. Moreover, there are no physical barriers to reaching the resolution of about 10 ps in the future [9]. In this context, it is worth to mention that the Jagiellonian PET (J-PET) Collaboration developed a novel whole-body PET scanner based on plastic scintillators [10–12] capable of simultaneous metabolic and positronium imaging [13,14]. Plastic scintillators are characterized by superior timing properties compared to scintillator crystals; the Cramer-Rao lower limit [15,16] for the CRT achievable with plastic scintillators corresponds to about 70 ps for 100 cm long plastic strips [11].

The signal-to-noise ratio (SNR) improvement due to TOF

* Corresponding author.

E-mail address: lech.raczynski@ncbj.gov.pl (L. Raczyński).

reconstruction is in the first approximation inversely proportional to the square root of the CRT [17]. The TOF gain in SNR allows to reduce the dose to the patient or shorten the overall examination time. The shorter scan time is of high importance, because it reduces the problems related to the movements of the organs in the patients which introduce an additional complication in current reconstruction methods.

Precise TOF information enables new reconstruction approaches; the examples are: TOF Fourier rebinning algorithm (FORE) [18–20] or simultaneous emission-attenuation reconstruction method [21]. In the DIRECT reconstruction algorithm the TOF events are histogrammed directly into images [22,23].

In this contribution we present a 3-dimensional (3-D), semi-analytic TOF reconstruction algorithm with the Total Variation (TV) regularization [24]. The proposed mathematical model describing measurement process with TOF information is inspired by the works of [25,26]. We have gone beyond those investigations to develop an algorithm that operates exclusively in the image space. The recent results [27] suggest that as the CRT improves, the TOF analytic algorithms become more competitive to statistical iterative methods, e.g., TOF Maximum Likelihood Expectation Maximization (TOF-MLEM). In spite of the linear behavior and predictability, analytic methods exhibit higher sensitivity to the data noise, as compared to the statistical iterative approaches. This leads to more noisy images for a low count data and thus affecting the quantitative precision of the imaging studies. Consequently, proper regularization of the analytic reconstruction is of a very practical interest.

Key component of this work is the application of the TV regularization in the image space during the analytic reconstruction filtering step, that is, after the TOF data have been TOF back-projected into the image space. Image space is substantially reduced in size as compared to the TOF data space, making the TV optimization operations much more efficient and practical. The most important part of our investigations is the evaluation of the kernel operator, corresponding to the linear transform mapping an original radioactive tracer distribution into a TOF back-projected image. In this paper the formula for calculations of the kernel operator is introduced and proven. Monte-Carlo (MC) simulation based performance studies of the developed algorithm compared with TOF-MLEM method [28] conclude the paper.

2. Reconstruction methods

2.1. TOF-PET data

The 3-D TOF-PET data can be expressed as [18,29]:

$$p(s, \phi, z_c, \theta, l) = \int_{-\infty}^{\infty} dl f(\vec{x} = l \vec{\omega}_1 + s \vec{\omega}_2 + z_c \vec{\omega}_3) h(l-l') \quad (1)$$

where the function $f(\vec{x} \triangleq (x, y, z))$ describes the radioactive tracer dis-

tribution, s and ϕ are the transaxial sinogram coordinates, z_c is the axial coordinate of the mid-point of the Line-of-Response (LOR), θ is the copolar angle between the LOR and transaxial plane, l is the TOF variable and h is the TOF profile (see Fig. 1). Three unit vectors are defined as:

$$\vec{\omega}_1 = (-\cos\theta\sin\phi, \cos\theta\cos\phi, \sin\theta) \quad (2)$$

$$\vec{\omega}_2 = (\cos\phi, \sin\phi, 0) \quad (3)$$

$$\vec{\omega}_3 = (0, 0, 1). \quad (4)$$

The TOF profile h is often modeled as a Gaussian function [30] with standard deviation $\sigma = c \cdot CRT / (4\sqrt{2}\log 2)$, where c denotes the speed of light. The TOF variable l is related with the TOF time difference Δt between the two arrival times of the two photons by $l = c\Delta t/2$, where $l = 0$ corresponds to the position of the LOR mid-point.

2.2. Notation

An image is represented by a 3-D function $f(\vec{x})$, where $\vec{x} \in \mathbb{R}^3$ denotes the space coordinate. A discretized version of continuous function $f(x, y, z)$ is represented by f_{ijk} , where $(i, j, k) = 1, \dots, n$ correspond to the voxels position in the image matrix. The bold symbol $\mathbf{f} \in \mathbb{R}^N$, where $N = n^3$, represents the vectorized version of the matrix f . In general $n \times n \times n$ size images are stored as N length vectors.

The geometrical arrangement of discrete detectors in a scanner determines a set of samples $(s, \phi, z_c, \theta, l) \triangleq \vec{\Sigma} \in \mathbb{R}^5$ in the projection space. The most common arrangement is a ring scanner: an even number of detectors uniformly spaced along a circle and

$$\vec{\Sigma} = \{(s, \phi, z_c, \theta, l) : |s| \leq R_d, \quad 0 \leq \phi \leq \pi, \\ |z_c| \leq \frac{L_d}{2}, \quad 0 \leq \theta \leq \frac{\pi}{2}, \quad |l| \leq \sqrt{R_d^2 + \frac{L_d^2}{4}}\},$$

where R_d is the detector radius, L_d is the detector length. A discretized TOF projection data are represented by matrix element p_{ijkmq} , where (i, j, k, m, q) correspond to the variables $(s, \phi, z_c, \theta, l)$, respectively. The bold symbol $\mathbf{p} \in \mathbb{R}^P$, represents the vectorized version of the matrix p .

The mathematical operator mapping a function $f(\vec{x})$ into $p(\vec{\Sigma})$, according to Eq. (1), is denoted by \mathcal{R} :

$$p(\vec{\Sigma}) = (\mathcal{R}f)(\vec{\Sigma}) \quad (5)$$

and $K \in \mathbb{R}^{P \times N}$ is a finite-dimensional sampling of the \mathcal{R} transform:

$$\mathbf{p} = K\mathbf{f} \quad (6)$$

and is often called a system matrix.

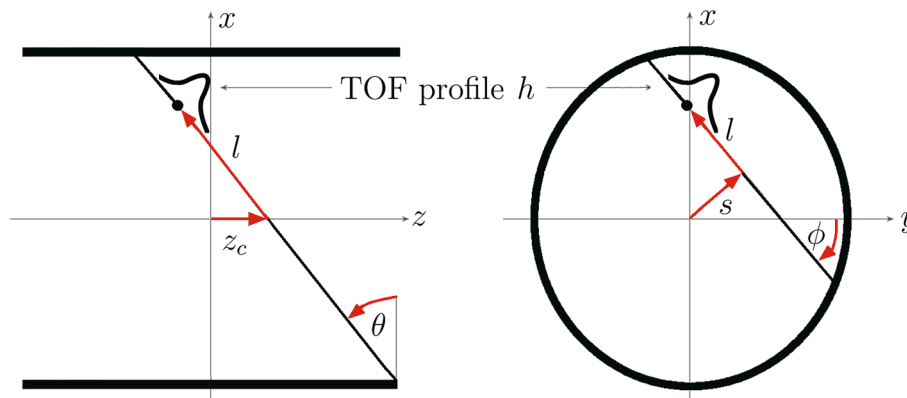


Fig. 1. Schematic view of a cylindrical 3-D PET tomograph in axial cross-section (left panel) and transaxial cross-section (right panel).

2.3. TOF-PET reconstruction

This section describes the proposed TOF algorithm. Procedure of the image reconstruction is as follows:

- **Data pre-correction.** In the first step, list-mode TOF events are pre-corrected. Pre-correction takes into account both the multiplicative factors (detector efficiency and attenuation factors) and the additive contamination of the data (random and scatter events). More details about the pre-correction and pre-processing of the J-PET data may be found in [12,31].
- **TOF back-projection.** Corrected data are back-projected to the image space using the TOF information. The relation between the back-projected data and the unknown radiotracer distribution will be introduced in Section 2.3.1.
- **Reconstruction with regularization.** The radiotracer distribution is reconstructed using the back-projected image by solving unconstrained TV regularization problem. The reconstruction problem based on TV approach will be discussed in Section 2.3.2.

The proposed algorithm is similar to the conventional back-projection filter (BPF) method [32]. However, in the BPF algorithm acquired data are first deposited into projection space. In contrary, in this approach the incorporation of the TOF information allows for direct transformation into the TOF back-projected image, i.e., each measured event is back-projected along the LOR at the estimated location in the image space by the TOF information. In the proposed work, the TV regularization acts as a filtering step, while in standard BPF or TOF-BPF algorithms, regularization is provided via apodizing functions. In the following sections we describe the subsequent steps of the proposed image reconstruction.

2.3.1. TOF back-projection

With the set of correction factors for all LORs in hand, the acquired data are deposited directly into TOF back-projected image. In the following we derive the relation between the TOF back-projected image and the original radioactive tracer distribution. We define a TOF back-projection operator ($\mathcal{H}^\#$) and we provide a linear transform of the projection data p defined in Eq. (5):

$$(\mathcal{H}^\# p)(\vec{x}) = (\mathcal{H}^\# \mathcal{H} f)(\vec{x}) \quad (7)$$

$$b(\vec{x}) = \mathcal{A} f(\vec{x}) \quad (8)$$

where $\mathcal{A} = \mathcal{H}^\# \mathcal{H}$ is an overall TOF forward and back-projection operator and b is TOF back-projected image. The images f and b have the same sizes and one-to-one voxel correspondence. Under the assumption that the TOF profile h is shift invariant, so that the integral in Eq. (1) is a convolution, the operator \mathcal{A} may be described as convolution operator with a kernel $a(\vec{x})$:

$$b(\vec{x}) = a(\vec{x}) * f(\vec{x}). \quad (9)$$

The aspect of the validity of the kernel a shift-invariance assumption will be discussed in the last part of Section 2.3.1.

The kernel $a(\vec{x})$ may be easily derived for a point source $\delta(\vec{x})$ placed in the tomograph center based on Eq. (7), i.e.,

$$a(\vec{x}) = (\mathcal{H}^\# \mathcal{H} \delta)(\vec{x}). \quad (10)$$

In the following, by convention we use subscripts o to denote exact values as opposite to estimated quantities described with subscripts e . A measured TOF event is defined as a set $\vec{g}_e = (x_{u,e}, y_{u,e}, z_{u,e}, t_{u,e}, x_{d,e}, y_{d,e}, z_{d,e}, t_{d,e})$. The $(x_{u,e}, y_{u,e}, z_{u,e})$ and $(x_{d,e}, y_{d,e}, z_{d,e})$ denote the reconstructed position along the upper and lower strips, respectively, see Fig. 2 for details. The $t_{u,e}$ and $t_{d,e}$ are reconstructed hit times such that:

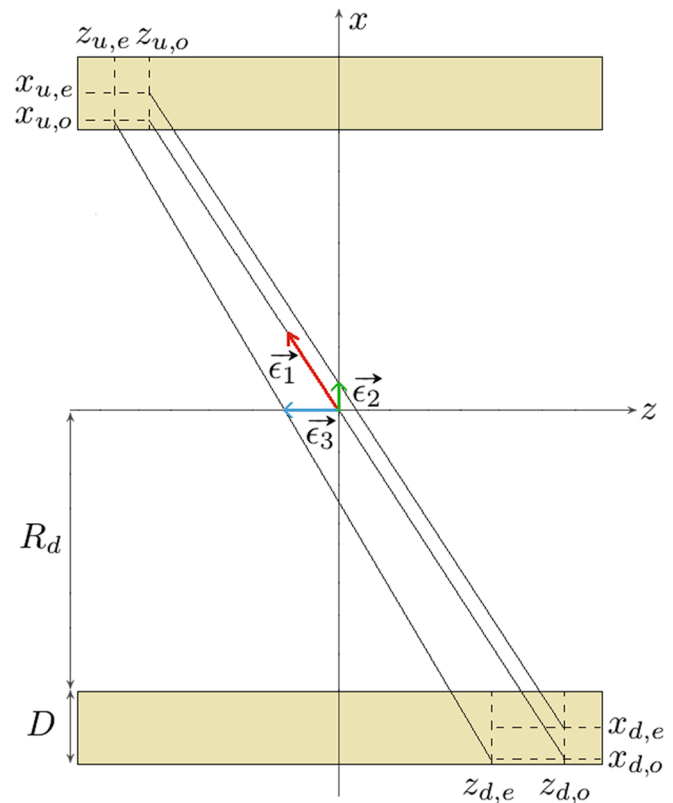


Fig. 2. Example of reconstruction of point source placed in the detector center presented in axial cross-section. The measurement error consists of three components: the TOF uncertainty (red arrow), the unknown DOI (green arrow), and the uncertainty along the axial direction (blue arrow). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

$$l_e = c \frac{t_{d,e} - t_{u,e}}{2}. \quad (11)$$

An annihilation occurring at $\vec{x}_o = (0, 0, 0)$ is measured as the point $\vec{x}_e = \vec{x}_o + \vec{\epsilon}$, where

$$\vec{\epsilon} = l_e \vec{\omega}_{1,e} + s_e \vec{\omega}_{2,e} + z_{c,e} \vec{\omega}_3 \quad (12)$$

$$= \vec{\epsilon}_1 + \vec{\epsilon}_2 + \vec{\epsilon}_3 \quad (13)$$

based on the parametrization introduced in Eq. (1). As it is shown in Fig. 2, the measurement error $\vec{\epsilon}$ is a vector with a component $\vec{\epsilon}_1$ describing TOF uncertainty, a component $\vec{\epsilon}_2$ related to the unknown depth of interaction (DOI) in a single scintillator and a component $\vec{\epsilon}_3$ describing uncertainty along axial direction (note that $\vec{\omega}_3$ is constant (see Eq. (4)) and therefore we do not use the subscripts). From Eq. (12) it may be seen that the measurement error $\vec{\epsilon}$ is evaluated after TOF back-projection of the TOF event \vec{g}_e into the image space. Hence, assuming that the $\vec{\epsilon}$ is independent of the locations and measurement errors of all other annihilations, the $\vec{\epsilon}$ may be considered as a random variable with probability density function (pdf) given by the overall TOF forward and back-projection operator a in Eq. (10).

The proposed mathematical model of measurement errors is inspired by the work of Snyder et al. [25]. In their paper they assumed that $\vec{\epsilon}$ is normally distributed and has two components: parallel and transverse to the LOR. The parallel component error is analogous to $\vec{\epsilon}_1$ in our derivation (see Fig. 2). However, instead of one transverse component error we considered two error vectors $\vec{\epsilon}_2, \vec{\epsilon}_3$ according to the parametrization given in Eq. (1). This approach allows us to include more

information about geometrical arrangement and readout specification of different PET systems, e.g., the J-PET scanner, in our calculations.

In order to simplify the further calculations the following assumption is proposed. Note that in the most interesting case for large detector radius R_d , orientation of component error $\vec{\epsilon}_1$, i.e., vector $\vec{w}_{1,e}$ in Eq. (12), is very close to the true LOR direction $\vec{w}_{1,o}$. Therefore, we assume that $\vec{\epsilon}_1 \approx l_e \vec{w}_{1,o}$, (see Fig. 2 for details) where only l_e is a random variable. Furthermore, the $\vec{\epsilon}_2$ and $\vec{\epsilon}_3$ in Eq. (12) depend only on transaxial ($x_{u,e}, y_{u,e}, x_{d,e}, y_{d,e}$) and axial ($z_{u,e}, z_{d,e}$) uncertainties, respectively. Hence, the error $\vec{\epsilon}$ may be approximated as a sum of three independent random variables and the unknown kernel a is given as the convolution:

$$a(\vec{x}) = (a_{(1)} * a_{(2)} * a_{(3)})(\vec{x}) \tag{14}$$

where $a_{(k)}$ describes pdf of error term $\vec{\epsilon}_k$ in Eq. (12) for $k = 1, 2, 3$. It is shown in appendix A, that

$$a_{(1)}(\vec{x}) = \kappa_1 \frac{h(\|\vec{x}\|) \mathcal{C}(\vec{x}, \theta_{acc})}{\|\vec{x}\| \sqrt{x^2 + y^2}} \tag{15}$$

$$a_{(2)}(\vec{x}) = \kappa_2 \frac{h_2(\sqrt{x^2 + y^2})}{\sqrt{x^2 + y^2}} \tag{16}$$

$$a_{(3)}(\vec{x}) = h_3(z). \tag{17}$$

The parameter θ_{acc} corresponds to the maximal accepted θ angle, κ_1, κ_2 stand for the normalization constants, h_2, h_3 are profile functions given in Eqs. (28), (31), respectively, and function \mathcal{C} is defined in Eq. (26). It should be stressed that the kernel a does not have a finite support due to the Gaussian functions in $a_{(1)}$ and $a_{(3)}$. In order to reduce the reconstruction time, we assessed the truncation of the kernel a and this aspect will be discussed in Section 3.4.

The evaluation of the kernel a is closely related to the modeling of scanner-specific point spread function (PSF). The PSF kernels take into account the effect of image blurring associated with the system resolution either in the image space [33,34] or in the projection space [35,36]. The PSFs are estimated based on reconstructed images of point sources using specific algorithm, e.g., TOF-MLEM. On the other hand, in the proposed approach there is no reconstruction stage during the calculation of the operator; the kernel a is obtained directly after TOF forward- and back-projection of the point source (see Eq. (10)). In this sense PSF kernel and operator a are complementary. Since the inclusion of the all measurement imperfections due to, e.g., the positron range and coincident photons acolinearity, into the operator a is difficult, PSF kernels may be additionally applied to the proposed reconstruction scheme. For this purpose, the PSF kernel may be found by analysis of reconstructed images of a point source and final operator will be given as a convolution of the operator a and the estimated PSF kernel. The parameters of the PSF kernels may be found through fitting different functions, e.g., weighted-exponential function [37], Gaussian model [38] or Gaussian mixture model [39].

In the following section, we discuss the validity of the kernel a shift-invariance assumption (see Eq. (9)). For this purpose we consider the influence of the position of point source on the distributions $a_{(k)}$ of measurement errors $\vec{\epsilon}_k$ for $k = 1, 2, 3$. The pdf $a_{(3)}$ does not depend on the position of point source as the uncertainty of measurement along the axial direction (z) is assumed to be constant (see appendix A.3 for details). The pdf $a_{(2)}$ for central and shifted point sources is not stationary since the efficient DOI changes with different angle in (x,y) cross-section between two detectors in coincidence. Under the assumption that annihilation photons propagate isotropically, only for central point source the angular difference in (x,y) cross-section between two detectors is always $\approx 180^\circ$ (see appendix A.2 for details). However, kernel $a_{(2)}$ contributes only to the (x,y) distribution of the overall kernel a and as the TOF uncertainty is still a major challenge of current PET scanners,

the investigation of influence of pdf $a_{(1)}$ on shift-invariance is of main importance. The main parameter that governs the distribution of kernel $a_{(1)}$ is θ_{acc} angle. Note that the θ_{acc} angle can not be greater than maximum detection angle θ for the point source placed in the tomograph center: $\theta_{max} = \arctan\left(\frac{L_d}{2R_d}\right)$. The solid angle covered by the tomograph is largest in the center (θ_{max}), but it decreases if one moves the point source from the center towards the edge. Consequently, the function describing the kernel a depends on the spatial location of the point source and is said to be spatially variant.

In the 1980s several 3-D analytic reconstruction approaches have been proposed to circumvent the assumption of shift-invariance [40,41]. This has been done by completion of the missing regions by forward-projecting (re-projecting) images initially reconstructed from the subset of projection data for which $\theta \leq \theta_{min}$. The limiting angle θ_{min} imposes shift-invariance everywhere within the field of view specified by the volume of the reconstructed object.

In this work no initial image reconstruction and data re-projection was performed, and the shift-invariance was approximated by rejecting any events with $\theta \geq \theta_{acc}$, where $\theta_{min} \leq \theta_{acc} \leq \theta_{max}$. Increasing θ_{acc} weakens the assumption of shift-invariance of the kernel a , however, on the other hand ensures higher statistics. A trade-off between number of accepted events and size of the area inside the tomograph for which the shift-invariance assumption of kernel a holds is optimized by changing the θ_{acc} parameter.

It is worth noting that the detection angle limitation is required in case of long PET detectors. A more oblique LOR penetrates more scintillating material than a LOR of less axial difference, which coupled with unknown depth of interaction degrades the axial resolution. For instance, for a 200 cm long EXPLORER scanner [42,43], the maximum axial difference between crystal pairs was limited to 115 cm. Moreover, the restriction on maximal θ angle is beneficial due to the photons attenuation effect in patient body; the more oblique the LOR the longer the path through the patient’s body and higher the probability of photon attenuation.

2.3.2. Reconstruction with regularization

The problem in Eq. (9) may be rewritten to the matrix notation:

$$\mathbf{b} = A\mathbf{f} \tag{18}$$

where A is finite-dimensional sampling of operator \mathcal{A} and therefore has a circulant structure (see Eqs. 8,9 for details). The circulant property of matrix A is a critical factor to speed up the algorithm as it allows the use of Fourier transform methods.

Since the TOF back-projected image \mathbf{b} is not a perfect noiseless image, the inverse problem defined in Eq. (18) is ill-posed and the application of some regularization technique is required. The most common class of regularization methods in image processing is based on TV approach. The TV-norm of image f can be defined as:

$$TV(\mathbf{f}) = \sum_i \sqrt{(D_i^{(x)}\mathbf{f})^2 + (D_i^{(y)}\mathbf{f})^2 + (D_i^{(z)}\mathbf{f})^2} \tag{19}$$

where $D^{(x)} \in \mathbb{R}^{N \times N}, D^{(y)} \in \mathbb{R}^{N \times N}$ and $D^{(z)} \in \mathbb{R}^{N \times N}$ are the first-order forward finite-difference operators, that approximate the gradient operators along the x, y, z directions, respectively. Therefore, $D_i^{(u)}\mathbf{f} \in \mathbb{R}$ is the discrete gradient of the image at pixel i along the u direction ($u = x, y, z$). We define $D = (D^{(x)}; D^{(y)}; D^{(z)}) \in \mathbb{R}^{3N \times N}$ as the total first-order forward finite-difference operator. Thus, the TV norm in Eq. (19) can be expressed as:

$$TV(\mathbf{f}) = \sum_i \|D_i\mathbf{f}\|_2 \tag{20}$$

where $D_i\mathbf{f} \in \mathbb{R}^3$ is the discrete gradient of the image at pixel i .

Optimization algorithm finds a solution $\hat{\mathbf{f}}$ of Eq. (18) by solving an

unconstrained regularization problem:

$$\hat{\mathbf{f}} = \operatorname{argmin}_{\mathbf{f}} \operatorname{TV}(\mathbf{f}) + \frac{\mu}{2} \|\mathbf{A}\mathbf{f} - \mathbf{b}\|_2^2, \quad (21)$$

which is known as the TV/L2 minimization. The μ is the regularization parameter. The data fidelity term in Eq. (21) is a L2 norm that constitutes the Gaussian noise model. Note that the image \mathbf{b} is evaluated after the pre-correction of the data (see the beginning of Section 2.3), which are no longer described by the Poisson distribution and it is convenient to assume the Gaussian noise model. The theory of penalty functions implies that the solution of Eq. (21) approaches the solution of Eq. (18) as μ goes to infinity. The proposed algorithm is based on the augmented Lagrangian method [44,45] and is presented in details in appendix B. This algorithm will be denoted hereafter as TOF-BPTV (TOF Back Projection Total Variation regularization).

3. Results

3.1. NEMA IEC body phantom simulation with the J-PET scanner

The performance of the proposed image reconstruction approach was investigated based on the MC simulation studies. The NEMA IEC body phantom, including six spheres and one long cylinder, was simulated according to the specification given in [46]. The four smallest spheres of 10, 13, 17 and 22 mm diameter simulated hot lesions with contrast ratio of 4 : 1 with respect to the activity concentration of the background. The two largest spheres of 28 and 37 mm diameter simulated cold lesions with no radioactivity. The centers of all six spheres were in the same transaxial plane and located at 70 mm from the phantom lid (see Fig. 3). The 180 mm long cold cylinder of 51 mm diameter was inserted on the central axis of the phantom. We modeled an injected activity of 53 MBq of ^{18}F -FDG dissolved in water. MC simulations of the PET imaging of the NEMA IEC phantom were performed with GATE [47]. We modeled the J-PET scanner geometry (see Fig. 3) with 384 strips arranged in one layer with inner radius of 428 mm and length of 500 mm [48,49]. A single plastic strip was built from BC-420 scintillator material [50] with dimensions of $7 \times 19 \times 500 \text{ mm}^3$. More details about the J-PET signal acquisition and processing may be found in [10,51–55].

A coincident event was defined as a set of consecutive interactions of photons, originating from a single annihilation and all interactions of secondary particles. The interactions were considered to originate from the same coincident event if they were detected within the fixed time

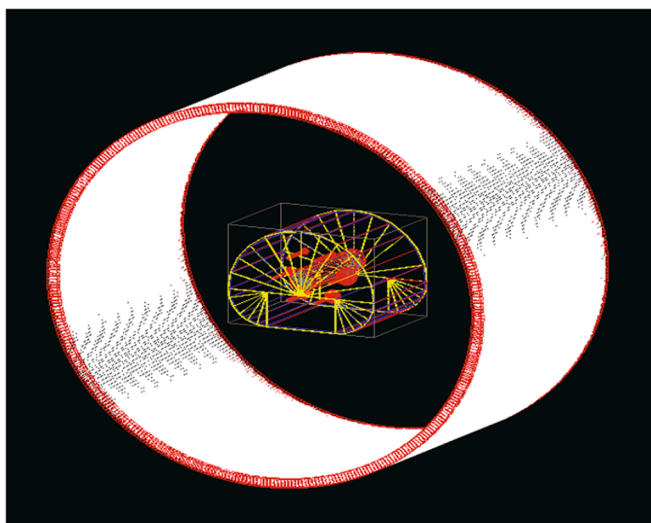


Fig. 3. Schematic view of one layer J-PET scanner with NEMA IEC body phantom inside.

window of 4 ns. This number ensures that the probability that interactions from two different events are ascribed to the same event is below 1 permille. Only events with exactly two interactions registered with an energy loss larger than 200 keV each were accepted [48]. For the NEMA IEC phantom simulation, a total of 60.0 million coincident events were recorded, corresponding approximately to a twenty minutes scan for a real J-PET acquisition. The total number of events included 20.0 million trues, 17.2 million scatters and 22.8 million randoms. During the reconstruction, only true coincidences were considered and the other two types of events were excluded from the event list. Therefore, only multiplicative factors, i.e., attenuation and detector geometric sensitivity, were considered at the pre-correction step. The attenuation correction was performed using the attenuation map used for the GATE simulation.

Reconstructions of simulated data were performed in MATLAB 7.14.0 (R2012a) with the use of chosen procedures from the Image Processing Toolbox. The proposed reconstruction algorithm based on TV minimization was implemented based on the work of [45]. The reconstructed images were represented as 3-D matrices with the voxel size of $2.5 \times 2.5 \times 2.5 \text{ mm}^3$.

As the MC simulation did not take into account the spatial and time uncertainties, an additional smearing using experimental resolutions of the detector was applied. The CRT and axial spatial resolution were simulated for silicon photomultipliers (SiPM) readout solution of the J-PET scanner. The values of CRT and Full Width at Half-Maximum (FWHM) in the z coordinate for SiPMs were estimated based on simulations presented in [11] and were equal to 230 ps and 2 cm, respectively.

3.2. Image quality measures

In order to evaluate the reconstruction performance, the NEMA NU 2–2007 norm for image quality assessment was used [46]. We evaluated two image quality assessments: contrast recovery coefficient (CRC) and background variability (BV). In the transaxial slice through the centers of hot and cold spheres, a circular region of interest (ROI) was defined on each sphere. Twelve circular ROIs of appropriate diameter were then defined on the phantom background. These background ROIs were automatically replicated to four transaxial slices $\pm 10 \text{ mm}$ and $\pm 20 \text{ mm}$ on either side of the central slice. Thus, in total, 60 ROIs were defined on the phantom background for each sphere. The CRC for each hot sphere with diameter d was calculated as:

$$\text{CRC} = \frac{C_{H,d}/C_{B,d} - 1}{4 - 1},$$

where $C_{H,d}$ was the average counts in the hot sphere, $C_{B,d}$ was the average of the background ROI counts and 4 was the true activity ratio between the hot spheres and the warm background. The CRC for each cold sphere with diameter d was calculated as:

$$\text{CRC} = 1 - C_{C,d}/C_{B,d},$$

where $C_{C,d}$ was the average counts in the cold sphere. The BV for each sphere with diameter d was calculated as:

$$\text{BV} = S_d/C_{B,d},$$

where S_d was the standard deviation of the background ROI counts. Additionally, a root mean square error (RMSE) between the full 3-D reconstructed image ($\hat{\mathbf{f}}$) and the true phantom activity image (\mathbf{f}^0) was calculated as a global quality criterion taking into account both bias and variance of the reconstruction algorithms:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{\mathbf{f}}_i - \mathbf{f}_i^0)^2}.$$

3.3. Shift-invariance violation investigation

In the first step of the analysis the influence of the θ_{acc} angle value on the performance of TOF-BPTV method was investigated. For assumed size of cylindrical J-PET detector $\theta_{\text{max}} \approx 30.3^\circ$. Additionally, taking into account the axial (90 mm) and transaxial (150 mm) extent of the NEMA IEC body phantom, θ_{min} that satisfies shift-invariance equals 15.5° [41]. θ_{acc} was changed in the range from 15° to 30° with 2.5° step. Table 1 in 2nd column lists the percentage of total counts as a function of θ_{acc} for a total of 20.0 million true events.

Fig. 4 compares the trends in CRC and BV of the 13-mm hot sphere (Fig. 4a) and the 22-mm hot sphere (Fig. 4b) for three different θ_{acc} values: smallest one (green curves), highest one (blue curves) and middle one (red curves). The error bars indicate standard deviations and were estimated from the 5 realizations of event smearing with assumed J-PET resolutions. The resulting curves for the remaining four cases (see Table 1) were not shown in Fig. 4 for the clarity of presentation. Each particular CRC versus BV curve was obtained after applying different regularization parameter μ (see Eq. (21)). More details about the influence of the parameter μ on the reconstructed image will be discussed in next section. For a quantitative comparison of the results for different θ_{acc} angles, for each CRC versus BV curve a minimal RMSE between reconstructed and true image of activity was calculated. The resulting values of $\min(\text{RMSE})$ are listed in Table 1 in 3rd column. We observe that the best results are obtained for the $\theta_{\text{acc}} = 22.5^\circ$ (see red curves in Fig. 4 and fourth row of Table 1). Hence, during the comparative studies presented in next section we applied $\theta_{\text{acc}} = 22.5^\circ$. This requirement imposes that the TOF-BPTV method uses 90.6% of acquired data (fourth row of Table 1) that corresponds to 18.1 million true events.

3.4. PET reconstruction comparative studies

To perform the comparative studies of the TOF-BPTV algorithm, TOF-MLEM method, implemented in CASToR software [56], was applied. CASToR offers several reconstruction methods for list-mode data as well as several data correction algorithms such as attenuation correction, normalization and PSF modeling. An iterative TOF-MLEM optimization algorithm using 60 iterations with no subsets division was selected. The TOF-MLEM cost function in CASToR software does not include any prior distribution and therefore the algorithm converges to the noisy image. A few approaches can be applied to remedy this problem and improve the quality of reconstructed images. First possibility is to stop the reconstruction algorithm after given number of iterations and use current image estimate as a solution. Moreover, CASToR applies PSF filters. Our preliminary investigations showed that spatial resolution of the J-PET scanner can be approximated with 3D gaussian of 6 mm transaxial FWHM and 12 mm axial FWHM and this kernel was applied to model a shift-invariant PSF in CASToR.

Fig. 5 compares the averages CRC and BV of the hot and cold spheres for both TOF-MLEM and TOF-BPTV image reconstruction. The error bars indicate standard deviations and were estimated from the 5 realizations of event smearing with assumed J-PET resolutions. In case of TOF-BPTV method the curves were obtained after applying various regularization parameter μ values in a range from 10 to 5000. In case of TOF-MLEM

algorithm the curves were obtained after applying different iterations in the range from 1 to 60. In all cases shown in Fig. 5, we observe the typical trade-off between the contrast (CRC) and the noise (BV). The regularization parameter μ trades-off the TV norm and the fidelity term (see Eq. (21)). Small values of regularization parameter favor TV penalty and give less noisy images, but the result may be smoothed with large bias. In that case both CRC and BV have small values (see Fig. 5 for $\mu = 10$ or 25). Increasing value of μ tends to give sharper images, but noise is also amplified (both CRC and BV increase). For each sphere the optimal μ was determined where the CRC reached 95% of its maximum value. In all cases μ was in the range from 200 to 300, and finally the smallest $\mu = 200$ was selected for further presentations in Figs. 7 and 8. Similarly, for TOF-MLEM algorithm, for each sphere the iteration number was extracted where the CRC reached 95% of its maximum value. For 10-, 13-, 17- and 28-mm spheres the 16th, 15th, 11th and 14th iterations were indicated, respectively. Finally, the 15th iteration was selected for further presentations.

In Fig. 6 bias in the region of the two hot spheres of 10 and 17 mm diameter and the background standard deviations values, are shown. The bias is defined as the difference between the true and reconstructed mean value in the ROIs. We normalized the image of true phantom activity to the range from 0 to 1, where 0 means no radioactivity, 1 corresponds to the hot regions and value 0.25 corresponds to the warm background (true activity ratio between the hot spheres and the background is 4). Each reconstructed image was normalized off-line to have the same total sum as the true phantom activity image. It is worth to note that the background standard deviations are approximately the same with TOF-BPTV for selected $\mu = 200$ (see Fig. 6b), while differ with TOF-MLEM for the 15th iteration (see Fig. 6a). The bias of the smallest source is almost the same for both algorithms, while the bias of the 17 mm diameter sphere is smaller in case of TOF-BPTV algorithm.

In Fig. 7 the two exemplary images in the transaxial (top) and coronal (bottom) slices through the centers of all spheres for 15th iteration for TOF-MLEM algorithm (left) and $\mu = 200$ for TOF-BPTV approach (right), are shown. The RMSE between the reconstructed image and the true phantom activity image is equal to 0.024 for TOF-BPTV method and 0.032 for TOF-MLEM algorithm. It can be seen that the structure of the warm phantom background differs and TOF-BPTV image (right) exhibits less intensity variability than the TOF-MLEM image (left). In both images the smallest hot sphere can be distinguished from the warm phantom background.

The reduced intensity variability of the TOF-BPTV image can also be clearly seen in the profiles shown in Fig. 8. In case of TOF-MLEM both cold spheres, observed on 90° and 150° in circular profile, have not reached value 0. The circular profiles further indicate that intensities in the TOF-BPTV image in hot spheres are more flat in comparison to the TOF-MLEM image. Therefore, despite the fact that maximal values for two smallest hot spheres, observed on 30° and 330° in circular profile, are higher for TOF-MLEM images, the mean values and CRCs are slightly better for TOF-BPTV algorithm.

In the last part of this study, the computational speed of both reconstruction methods was compared. Since both algorithms are created with different programming languages, namely C++ in case of TOF-MLEM (CASToR) and MATLAB in case of TOF-BPTV, the estimate of efficiency is only indicative. For comparison purpose, for both algorithms the same total number of 20.0 million events was considered. The reconstruction volume was $400 \times 400 \times 400 \text{ mm}^3$, comprised of $160 \times 160 \times 160$ voxels. On a single CPU (Intel Core i5-5200U @ 2.20 GHz), the total computing time of the TOF-BPTV algorithm was 247 s on average. It should be stressed that the TOF-BPTV evaluation time depends on the overall number of iterations required to solve the TV/L2 minimization problem (see appendix B). We found empirically that the rate of convergence using different values of regularization parameter μ is approximately the same and that 17 iterations is robust to most of the cases. On the same CPU, the computing time of one iteration of the TOF-MLEM method was 121 s on average. In TOF-MLEM algorithm, the

Table 1
Performance of TOF-BPTV algorithm for different values of θ_{acc} angle.

θ_{acc} [°]	% of total	min (RMSE)
15.0	63.1	0.028
17.5	74.2	0.026
20.0	83.4	0.025
22.5	90.6	0.024
25.0	95.7	0.026
27.5	98.8	0.028
30.0	100.0	0.030

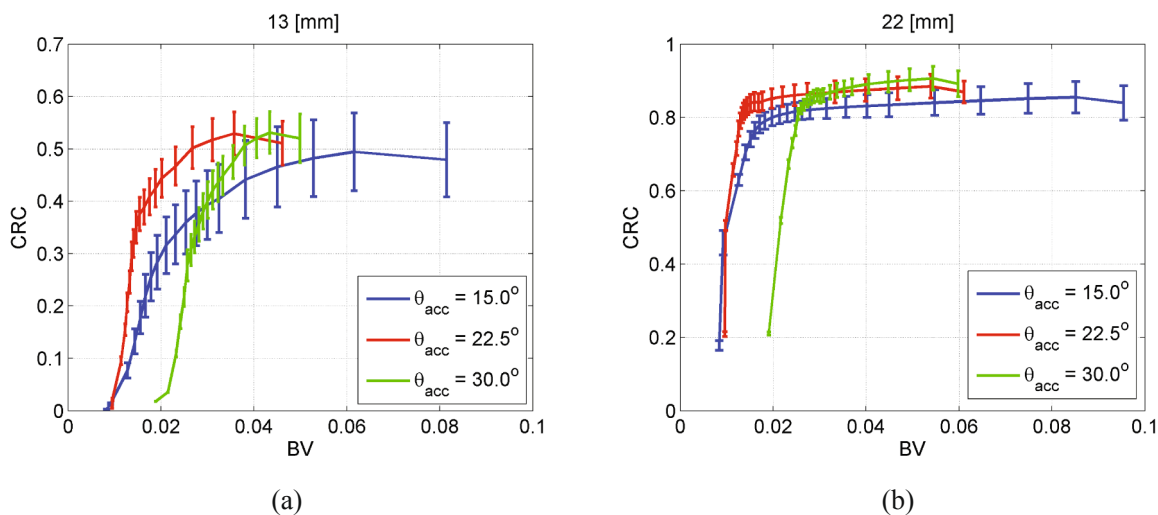


Fig. 4. CRC versus BV in the reconstructed images of 13 mm hot sphere (a) and 22 mm hot sphere (b) computed for different θ_{acc} values.

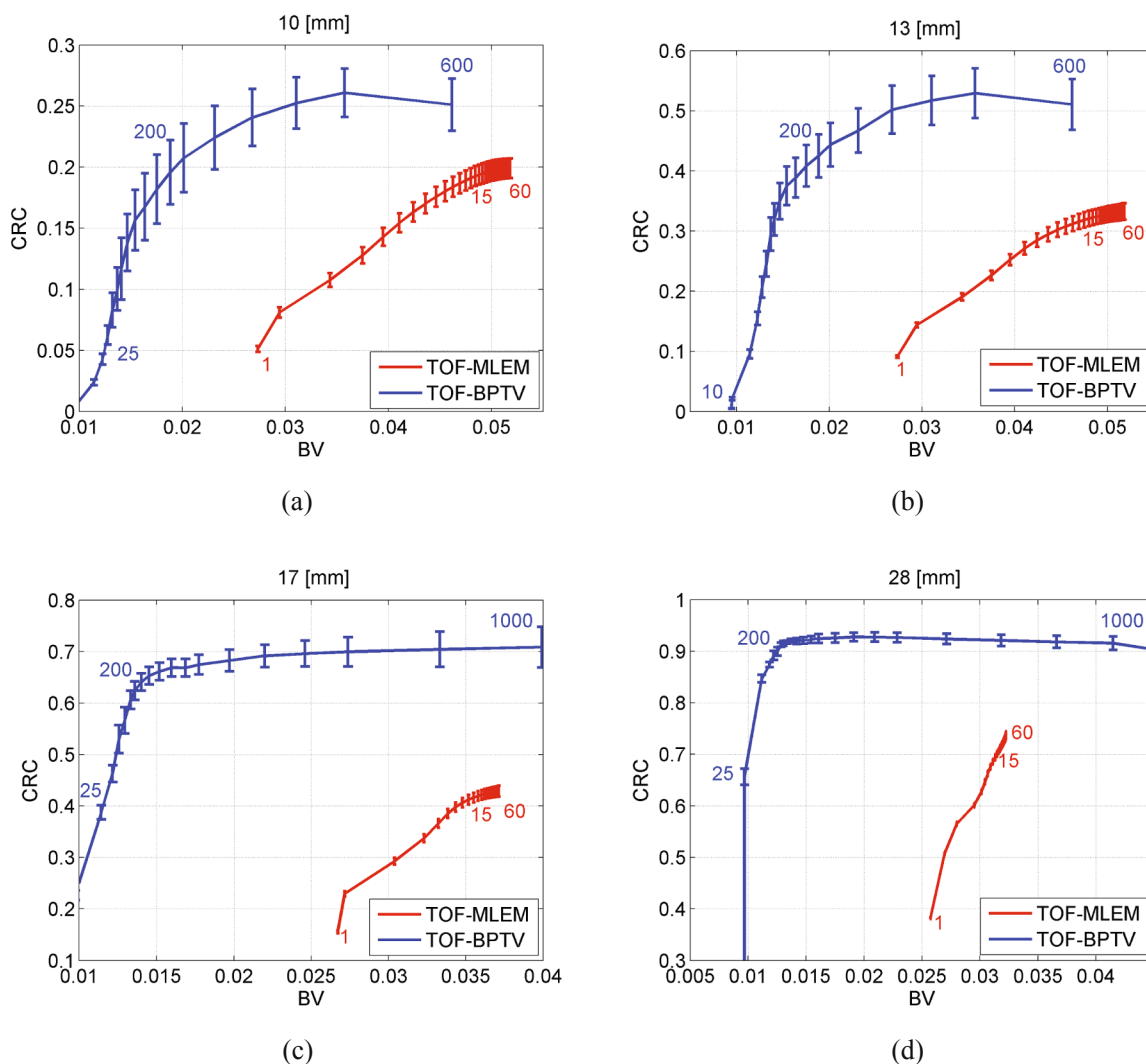


Fig. 5. CRC versus BV in the reconstructed images of 10 mm hot sphere (a), 13 mm hot sphere (b), 17 mm hot sphere (c) and 28 mm cold sphere (d) computed for TOF-BPTV algorithm (blue curves) and TOF-MLEM algorithm (red curves). Values of μ and iteration number are indicated with blue and red colours, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

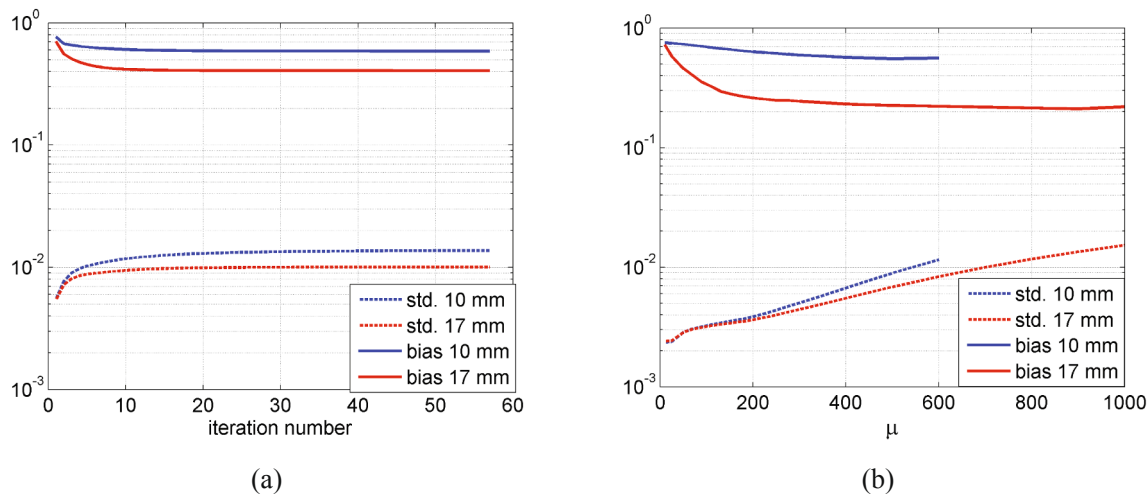


Fig. 6. Biases and standard deviations of the hot sources with 10 mm and 17 mm diameter for TOF-MLEM algorithm (a) and TOF-BPTV algorithm (b).

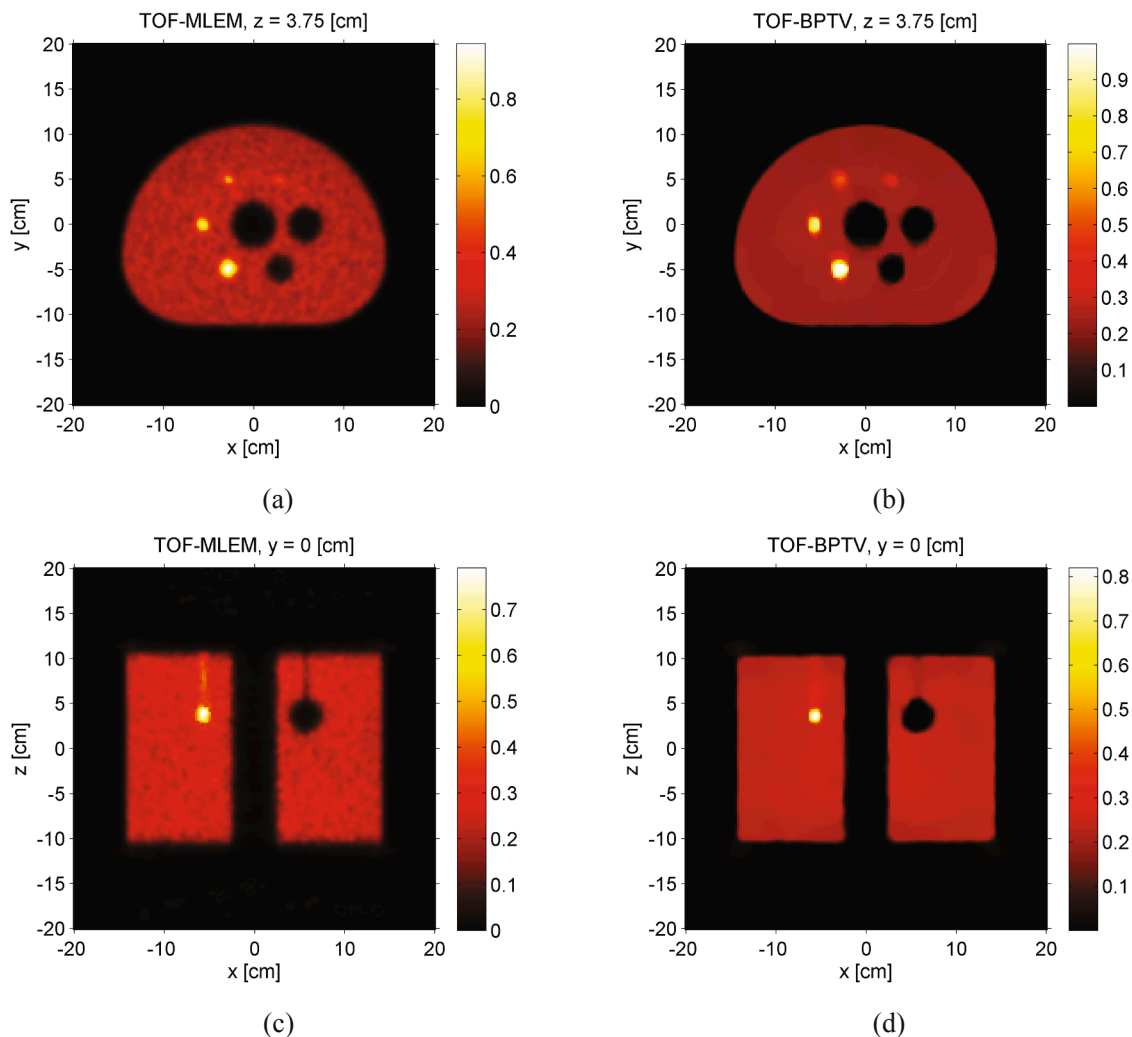


Fig. 7. Transaxial (top) and coronal (bottom) slices through the centers of spheres inside the NEMA IEC body phantom. Images shown for 15th iteration of TOF-MLEM algorithm (a, c) and reconstructed using TOF-BPTV algorithm with $\mu = 200$ (b, d).

convergence criterion was met after 15 iterations, and the total reconstruction time thus amounted to 1815 s. Hence, reconstruction time in TOF-BPTV method was approximately 7.3 times shorter than required

by TOF-MLEM reconstruction.

It should be stressed that the computing time of the TOF-BPTV method strongly depends on the kernel a size. During preliminary

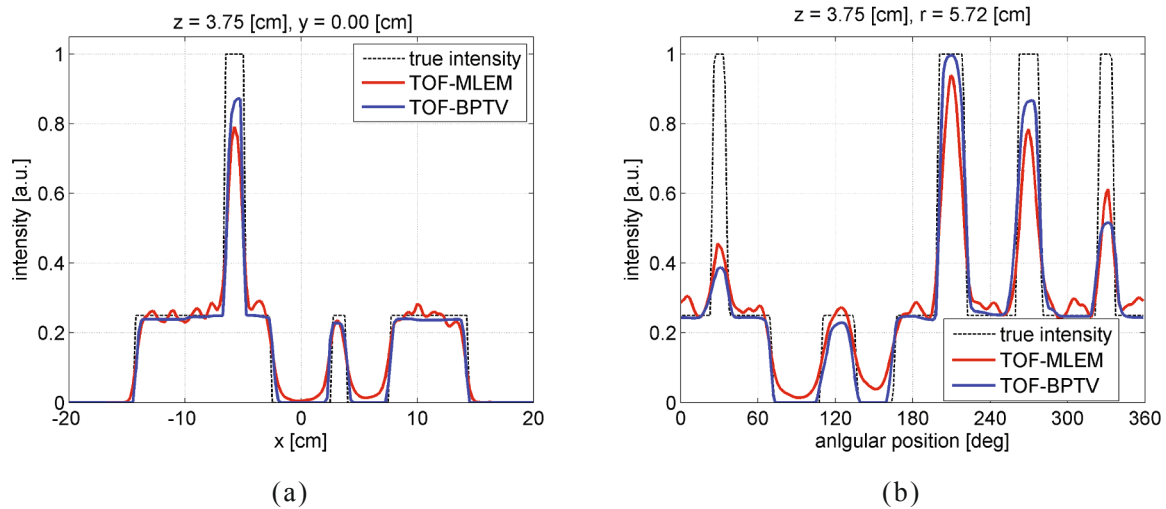


Fig. 8. Emission density profiles through the reconstructed images in transaxial slices of Fig. 7: line profile for $y = 0$ cm (a), circular profile over centers of all spheres for $r = 5.72$ cm (b).

studies an investigation for the potential introduction of error due to truncation on different distances from the center of the kernel a along (x, y, z) directions, from 2.0σ to 4.0σ , was performed. Our tests revealed that 3.0σ seemed to be optimal option for balance between quality degradation and acceleration of reconstruction process. This value was used throughout the studies described in this work.

4. Discussion

During our investigations, we observed that TOF-BPTV algorithm is very sensitive for regularization parameter μ and its value should be kept as low as possible, here $\mu = 200$. In other words, for higher values of μ , the structures appearing in the image have completely different forms than observed in MLEM; the fluctuations are small and slowly changeable. We also noted that the deformation of the initial distribution is greater the smaller is the number of registered events. This is in our opinion one of the main drawback of the TOF-BPTV method.

4.1. Related works

PET reconstruction using TV regularization was investigated by several groups. For instance, imposing the solutions with a small TV norm, leads to an extension of the MLEM method known as the EM-TV algorithm [57,58]. However, our approach is inspired by an alternating regularization procedure for PET first introduced in [59]. Given a noisy projection data \mathbf{p} an image is reconstructed by solving:

$$\min_{\mathbf{f}} \alpha \text{TV}(\mathbf{Kf}) + \beta \|\mathbf{f}\|_1 + \frac{\gamma}{2} \|\mathbf{Kf} - \mathbf{p}\|_2^2 \quad (22)$$

where α, β, γ are positive parameters. In this method TV regularization acts only on the \mathbf{Kf} . As an extension, in [60] an image from PET measurements has been reconstructed by smoothing both in projection and in image space:

$$\min_{\mathbf{f}} \alpha \text{TV}(\mathbf{f}) + \beta \text{TV}(\mathbf{Kf}) + \sum_i \frac{(K_i \mathbf{f} - \mathbf{p}_i)^2}{\mathbf{p}_i} \quad (23)$$

In both methods, the reconstruction problem does not take into account the TOF information, i.e., the projection data \mathbf{p} in Eqs. 22,23 are non-TOF data. With the extra dimension brought by TOF information, the size of the matrix K increases significantly and the problems in Eqs. 22,23 are invariably large-scale. In fact, in that case it becomes impossible to store the matrix K in the memory, even in the sparse format. Moreover, the matrix K does not have a circulant structure as the matrix

A in Eq. (21).

4.2. Choice of parameter μ in the TV regularization problem

The choice of the regularization parameter μ is not known prior to solving the TV minimization problem formulated in Eq. (21). In theory, the optimal value of μ depends on the noise level in the data, so indirectly also on number of true coincidence events or equivalently the time scan of the PET acquisition with a given dose. It is expected that with the increasing number of true events, the signal-to-noise ratio will improve and the optimal μ will also increase. However, our preliminary investigation reveals that a value of μ does not change significantly as the number of events changes from 2.5 million to 25 millions true events; the value of the parameter μ lies in the range from 150 to 250.

4.3. Method limitations

The main limitation of the proposed algorithm is due to the assumption of the operator $a(\vec{x})$ shift-invariance (see Eq. (9)). The kernel operator, evaluated in the appendix A, is derived for a spatially invariant system based on a point source placed in the center of the PET detector.

Since the TV optimization in image space is very efficient we propose to incorporate the shift variance by evaluation of a set of operators $a(\vec{x}, \vec{x}_p)$ for set of point sources \vec{x}_p placed inside the detector volume. The operators $a(\vec{x}, \vec{x}_p)$ may be simulated only for points \vec{x}_p in 2D space, i.e., for point sources along the radial and axial directions, assuming that two points with the same axial and radial distance to the detector center share the same kernel $a(\vec{x}, \vec{x}_p)$ after rotation in the transaxial plane. The reconstruction problem formulated in Eq. (21) may be then modified so that for each kernel $a(\vec{x}, \vec{x}_p)$ a small sub-image centered at position \vec{x}_p could be calculated independently. The final, output image could be evaluated as a weighted sum of the overlapping sub-images. This approach will require specifications of both spacing of the point sources \vec{x}_p and percentage of the sub-images overlapping. This topic is related to existing work on non-stationary image PSF and on TOF spatial inhomogeneity [61] and will be carried on during further studies.

In this section we wish to discuss also the influence of the TOF resolution on the computational cost of the algorithm. The cost of the method is highly dependent on the TOF resolution of the PET scanner; the better the time resolution, the more efficient the reconstruction algorithm. The worsening of the TOF resolution increases the support of the convolution kernel $a_{(1)}$; a TOF profile function h in a nominator has

larger standard deviation (see Eq. (15)). In the limiting case, for the non-TOF data, the absence of TOF information is equivalent to setting $h = 1$ in Eq. (15), and the support of the kernel $a_{(1)}$ goes to infinity. In that case it is more reasonable to follow the reconstruction approach proposed in [59,60], where the TV regularization problem is applied in the projection space and not in the image space (see Eqs. (22) and (23)).

4.4. Extension of the proposed method to conventional PET systems

The proposed image reconstruction algorithm may be applied to state-of-the-art PET scanners equipped with crystal scintillators, e.g., EXPLORER [42,43]. For this purpose, in order to evaluate the matrix A in TV minimization problem formulated in Eq. (21), kernel a , defined in Eq. (14), has to be adapted. In case of PET scanners with crystal scintillators, the only difference concerns the kernel $a_{(3)}$. Compared with J-PET, here the measurement error in scintillator along z direction is uniformly distributed with the width equal to the length of the single crystal L_c . This implies that the pdf $a_{(3)}$, describing the distribution of the positions mean value along the upper and lower strip, has a triangular form with FWHM of $\frac{L_c}{2}$. The pdfs $a_{(1)}$ and $a_{(2)}$ may be calculated in the same way as in the J-PET scanner, i.e., the CRT and θ_{acc} need to be specified to evaluate kernel $a_{(1)}$ and thickness of the crystal scintillator is required to evaluate kernel $a_{(2)}$ (see Eqs. (15)), respectively.

We wish to make last comment about the application of the proposed algorithm to the EXPLORER scanner. This system is based on many modules containing a matrix of crystals, with small but non-negligible gaps between these modules. In that case kernel operator a is much more sensitive to the spatial position of \vec{x}_p , making the system response even more shift variant than in considered J-PET scanner. Therefore, the incorporation of the shift variance to the proposed algorithm is crucial.

Appendix A. Calculation of convolution operator

In the following we will show a derivation of pdfs $a_{(1)}, a_{(2)}, a_{(3)}$ of measurement errors $\vec{\epsilon}_1, \vec{\epsilon}_2, \vec{\epsilon}_3$, respectively. In most cases it is convenient to model the error distributions in projection space and $p_{(k)}^\delta$ describes the pdf of projection data of a point source $\delta(\vec{x})$ affected with errors introduced only by k th components in Eq. (12).

A.1. Calculation of operator

As described in Section 2.3.1 we assume that the $\vec{\epsilon}_1$ depends only on time uncertainties $(t_{u,e}, t_{d,e})$. Therefore, the pdf of projection data $p_{(1)}^\delta$ considers only TOF variable and according to Eq. (1):

$$p_{(1)}^\delta(\vec{\Sigma}) = \int_{-\infty}^{\infty} dl \delta(\vec{x} = l \vec{\omega}_1) h(l-l) = h(l \vec{\omega}_1) \tag{24}$$

where function h describes the TOF profile. The back-projection of the projection data $p_{(1)}^\delta$ onto the image space is given as:

$$a_{(1)}(\vec{x}) = (\mathcal{R}^\# p_{(1)}^\delta)(\vec{x}) = \int_0^{\theta_{acc}} d\theta \int_0^\pi d\phi \int_{-\infty}^{\infty} dl h(l \vec{\omega}_1). \tag{25}$$

It is convenient to convert the spherical coordinates in the above integral to Cartesian coordinates $\vec{u} = (u, v, w)$. The transformation between coordinate systems is given by:

$$dld\phi d\theta = \begin{vmatrix} \frac{dl}{du} & \frac{dl}{dv} & \frac{dl}{dw} \\ \frac{d\phi}{du} & \frac{d\phi}{dv} & \frac{d\phi}{dw} \\ \frac{d\theta}{du} & \frac{d\theta}{dv} & \frac{d\theta}{dw} \end{vmatrix} dudvdw = \frac{d\vec{u}}{\|\vec{u}\| \sqrt{u^2 + v^2}}$$

and finally:

5. Conclusions

The goal of this paper was the introduction of TV regularization for TOF-PET projection data transformed into a TOF back-projected image. In contrast to a more traditional application of the TV regularization for PET data in the projection space, the efficiency our approach comes from the one-time TOF back-projection step. The simulation study demonstrated that the proposed reconstruction method was approximately 7.3 times faster than the TOF-MLEM. Simultaneously, it was shown that the proposed algorithm can reach better performance for PET imaging than TOF-MLEM algorithm.

Future work will address an incorporation of the shift variance technique in the image reconstruction process. We believe that this modification will further improve the performance of the algorithm.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We acknowledge the technical and administrative support of A. Heczko, M. Kajetanowicz, and W. Migdał, and the financial support of the Foundation for Polish Science through MPD and TEAM POIR.04.04.00-00-4204/17 programmes, the National Science Centre through Grant No. 2016/21/B/ST2/01222, the Ministry for Science and Higher Education through Grant No. 7150/E-338/SPUB/2017/1, the EU and MSHE Grant No. POIG.02.03.00-161 00-013/09. B. Hiesmayr acknowledges support from the Austrian Science Fund (FWF-P26783). We thank S. Bass for the careful reading of the manuscript.

$$a_{(1)}(\vec{x}) = \kappa_1 \frac{h(\|\vec{x}\|) \mathcal{E}(\vec{x}, \theta_{acc})}{\|\vec{x}\| \sqrt{x^2 + y^2}}$$

where κ_1 stands for the normalization constant and the function $\mathcal{E}(\vec{x}, \theta_{acc})$ is defined as:

$$\mathcal{E}(\vec{x}, \theta_{acc}) = \begin{cases} 1 & \frac{z}{\|\vec{x}\|} \leq \cos \theta_{acc} \\ 0 & \text{otherwise.} \end{cases} \tag{26}$$

The function $\mathcal{E}(\vec{x}, \theta_{acc})$ originates from the integration limit θ_{acc} of the θ angle in Eq. (25).

A.2. Calculation of operator

As described in Section 2.3.1 the $\vec{\epsilon}_2$ depends only on transaxial uncertainties $(x_{u,e}, y_{u,e}, x_{d,e}, y_{d,e})$. Therefore, the pdf of projection data $p_{(2)}^\delta$ may be described as:

$$p_{(2)}^\delta(\vec{\Sigma}) = \int_{-\infty}^{\infty} dl \delta(\vec{x} = l \vec{\omega}_2^\perp) h_2(l - l) = h_2(l \vec{\omega}_2^\perp) \tag{27}$$

where the profile function $h_2(l)$ has a triangle distribution:

$$h_2(l) = \begin{cases} \frac{2D - 4|l|}{D^2} & |l| \leq \frac{D}{2} \\ 0 & \text{otherwise} \end{cases} \tag{28}$$

where D is a thickness of the plastic scintillator. The function $h_2(l)$ originates from the fact that the depth of interaction is unknown and we assume a mid-point of the strip as the measured position in (x, y) cross-section (see estimates $x_{d,e}, x_{u,e}$ in Fig. 2).

The back-projection of the projection data $p_{(2)}^\delta$ onto the image space is given as:

$$a_{(2)}(\vec{x}) = (\mathcal{R}^\# p_{(2)}^\delta)(\vec{x}) = \int_0^{\theta_{acc}} d\theta \int_0^\pi d\phi \int_{-\infty}^{\infty} dl h_2(l \vec{\omega}_2^\perp) = \int_0^\pi d\phi \int_{-\infty}^{\infty} dl h_2(l \vec{\omega}_2^\perp) \tag{29}$$

and does not depend on the θ angle. It is convenient to convert the polar coordinates in the above integral to Cartesian coordinates. The transformation between coordinate systems is:

$$dl d\phi = \begin{vmatrix} \frac{dl}{du} & \frac{dl}{dv} \\ \frac{d\phi}{du} & \frac{d\phi}{dv} \end{vmatrix} du dv = \frac{du dv}{\sqrt{u^2 + v^2}}$$

and finally:

$$a_{(2)}(\vec{x}) = \kappa_2 \frac{h_2(\sqrt{x^2 + y^2})}{\sqrt{x^2 + y^2}}$$

where κ_2 stands for the normalization constant.

A.3. Calculation of operator

As described in Section 2.3.1 the $\vec{\epsilon}_3$ depends only on axial uncertainties $(z_{u,e}, z_{d,e})$. In this case the calculations do not involve the analysis of pdf of errors in projection space. Note that:

$$\vec{\epsilon}_3 = \frac{z_{d,e} + z_{u,e}}{2} \tag{30}$$

and under the assumption that the uncertainty of the measurement of axial positions $z_{u,e}, z_{d,e}$ are a Gaussian functions with standard deviation σ_z that not depend on the position along the strip, the pdf $a_{(3)}$ is given as:

$$a_{(3)}(\vec{x}) = h_3(z) = \frac{1}{\sqrt{\pi} \sigma_z} \exp\left(-\frac{z}{\sigma_z^2}\right). \tag{31}$$

Appendix B. TV/L2 minimization problem

The TV minimization problem defined in Eq. (21) with definition of TV norm given in Eq. (20) may be expressed as:

$$\min_{\mathbf{f}} \sum_{i=1}^N \|D_i \mathbf{f}\|_2 + \frac{\mu}{2} \|\mathbf{A}\mathbf{f} - \mathbf{b}\|_2^2. \quad (32)$$

The problem in Eq. (32) is convex, but due to the nondifferentiability and nonlinearity of the TV function, the model is computationally difficult to solve. The first step of solving the problem in Eq. (32) is the introduction of an auxiliary variable $\mathbf{w}_i \in \mathbb{R}^3$ to transfer $D_i \mathbf{f}$ out of the nondifferentiable term $\|\cdot\|_2$

$$\min_{\mathbf{f}, \mathbf{w}} \sum_{i=1}^N \|\mathbf{w}_i\|_2 + \frac{\mu}{2} \|\mathbf{A}\mathbf{f} - \mathbf{b}\|_2^2 \quad \text{subject to} \quad D_i \mathbf{f} = \mathbf{w}_i \quad (33)$$

The Lagrangian function $\mathcal{L}(\mathbf{f}, \mathbf{w}, \lambda)$ of problem in Eq. (33) is defined as:

$$\mathcal{L}(\mathbf{f}, \mathbf{w}, \lambda) = \sum_i (\|\mathbf{w}_i\|_2 + \lambda_i (D_i \mathbf{f} - \mathbf{w}_i)) + \frac{\mu}{2} \|\mathbf{A}\mathbf{f} - \mathbf{b}\|_2^2 \quad (34)$$

where λ_i is the Lagrange multiplier associated with the constraint $D_i \mathbf{f} = \mathbf{w}_i$. According to the idea of the quadratic penalty method, it is likely to penalize the violation of constraint $D_i \mathbf{f} = \mathbf{w}_i$. For instance, one may solve the following problem:

$$\mathcal{L}_A(\mathbf{f}, \mathbf{w}, \lambda) = \mathcal{L}(\mathbf{f}, \mathbf{w}, \lambda) + \sum_i \left(\frac{\beta}{2} \|D_i \mathbf{f} - \mathbf{w}_i\|_2^2 \right) \quad (35)$$

where β is a regularization parameter associated with each quadratic penalty term $\|D_i \mathbf{f} - \mathbf{w}_i\|_2^2$. Minimizing the problem in Eq. (35) is known as an augmented Lagrangian method. When the original problem, defined in Eq. (32), is convex, the first-order optimality conditions of augmented Lagrangian function become sufficient for finding optimal solution \mathbf{f} .

The advantage of the introduction of an auxiliary variable \mathbf{w} is that, while either one of the variables (\mathbf{f}, \mathbf{w}) is fixed, minimizing the function \mathcal{L}_A with respect to the other has a closed-form formula with low computational complexity. To this end, the alternating direction method is used to iteratively solve the optimization problem in Eq. (35).

References

- [1] Allemand R, Gresset C, Vacher J. Potential advantages of a cesium uride scintillator for a time of flight positron camera. *J Nucl Med* 1980;21:153–5.
- [2] Snyder DL, Polite DG. Image reconstruction from list-mode data in an emission tomography system having time-of-flight measurements. *IEEE Trans Nucl Sci* 1983; 20:1843–9.
- [3] Moses WW, Derenzo SE. Prospects for time-of-flight PET using LSO scintillator. *IEEE Trans Nucl Sci* 1999;46:474–8.
- [4] Kuhn A, Surti S, Karp JS, Raby PS, Shah KS, Perkins A, Muehlethner PG. Design of a lanthanum bromide detector for time-of-flight PET. *IEEE Trans Nucl Sci* 2004;5: 2550–7.
- [5] Stomka P, Pan T, Germano G. Recent advances and future progress in PET instrumentation. *Semin Nucl Med* 2016;46:5–19.
- [6] van Sluis J, et al. Performance characteristics of the digital biograph vision PET/CT system. *J Nucl Med* 2019;60:1031.
- [7] Schaart DR, Seifert S, Vinke R, van Dam HT, Dendooven P, Lohner H, Beekman FJ. LaBr 3: Ce and SiPMs for time-of-flight PET: achieving 100 ps coincidence resolving time. *Phys Med Biol* 2010;55:N179–89.
- [8] Nemallapudi MV, Gundacker S, Lecoq P, Auffray E, Ferri A, Gola A, Piemonte C. Sub-100 ps coincidence time resolution for positron emission tomography with LSO: Ce codoped with Ca. *Phys Med Biol* 2015;60:4635–49.
- [9] Lecoq P. Pushing the limits in time-of-flight PET imaging. *IEEE Trans Radiat Plasma Med Sci* 2017;1:473–85.
- [10] Moskal P, et al. Test of a single module of the J-PET scanner based on plastic scintillators. *Nucl Instrum Meth Phys Res A* 2014;764:317–21.
- [11] Moskal P, et al. Time resolution of the plastic scintillator strips with matrix photomultiplier readout for J-PET tomograph. *Phys Med Biol* 2016;61:2025–47.
- [12] Niedzwiecki S, et al. J-PET: a new technology for the whole-body PET imaging. *Acta Phys Polon B* 2017;48:1567–76.
- [13] Moskal P, et al. Feasibility study of the positronium imaging with the J-PET tomograph. *Phys Med Biol* 2019;64. 055017.
- [14] Moskal P, Jasińska B, Stepien EL, Bass S. Positronium in medicine and biology. *Nat Rev Phys* 2019;1:527–9.
- [15] DeGroot MH. Probability and statistics. Reading, MA: Addison-Wesley; 1986. p. 420–6.
- [16] Seifert S, van Dam HT, Schaart DR. The lower bound on the timing resolution of scintillation detectors. *Phys Med Biol* 2012;57:1797–814.
- [17] Budinger T. Time-of-flight positron emission tomography: status relative to conventional PET. *J Nucl Med* 1983;24:73.
- [18] Defrise M, Casey ME, Michel C, Conti M. Fourier rebinning of time-of-flight PET data. *Phys Med Biol* 2005;50:2749–63.
- [19] Watson CC. An improved weighting kernel for analytical time-of-flight PET reconstruction. *IEEE Trans Nucl Sci* 2008;55:2551–6.
- [20] Cho S, Ahn S, Li Q, Leahy RM. Exact and approximate Fourier rebinning of PET data from time-of-flight to non time-of-flight. *Phys Med Biol* 2009;54:467–84.
- [21] Defrise M, Rezaei A, Nuyts J. Time-of-flight PET data determine the attenuation sinogram up to a constant. *Phys Med Biol* 2012;57:885–99.
- [22] Matej S, Surti S, Jayanthi S, Daube-Witherspoon ME, Lewitt RM, Karp JS. Efficient 3-D TOF PET reconstruction using view-grouped histo-images: DIRECT – Direct Image Reconstruction for TOF. *IEEE Trans Med Imag* 2009;28:739–51.
- [23] Matej S, Daube-Witherspoon ME, Karp JS. Analytic TOF-PET reconstruction algorithm within DIRECT data partitioning framework. *Phys Med Biol* 2016;61: 3365–86.
- [24] Rudin LI, Osher S, Fatemi E. Nonlinear total variation based noise removal algorithms. *Physica D* 1992;60:259–68.
- [25] Snyder DL, Thomas LJ, Ter-Pogossian MM. A mathematical model for positron-emission tomography systems having time-of-flight measurements. *IEEE Trans Nucl Sci* 1982;28:3575–83.
- [26] Mallon A, Grangeat P. Three-dimensional PET reconstruction with time-of-flight measurement. *Phys Med Biol* 1992;37:717–29.
- [27] Westerwoudt V, Conti M, Eriksson L. Advantages of improved time resolution for TOF-PET at very low statistics. *IEEE Trans Nucl Sci* 2014;61:126–33.
- [28] Jacobson M, et al. Enhanced 3D PET OSEM reconstruction using inter-update Metz filtering. *Phys Med Biol* 2000;45:2417–39.
- [29] Li Y, Matej S, Metzler SD. A unified Fourier theory for time-of-flight PET data. *Phys Med Biol* 2016;61:601–24.
- [30] Conti M, Bendriem B, Casey M, Chen M, Kehren F, Michel C, Panin V. First experimental results of time-of-flight reconstruction on an LSO PET scanner. *Phys Med Biol* 2005;50:4507–26.
- [31] Pawlik-Niedzwiecka M, et al. Preliminary studies of J-PET detector spatial resolution. *Acta Phys Polon A* 2017;132:1645–8.
- [32] Bates R, Peters T. Towards improvements in tomography. *New Zealand J Sci* 1971; 14:883.
- [33] Reader AJ, Julyan PJ, Williams H, Hastings DL, Zweit J. EM algorithm resolution modeling by image-space convolution for PET reconstruction. *IEEE Nucl Sci Symp Conf Rec* 2002;2:1221–5.
- [34] Rahmim A, Lenox M, Michel C, Reader AJ, Sossi V. Space-variant and anisotropic resolution modeling in list-mode EM reconstruction. *IEEE Nucl Sci Symp Conf Rec* 2003;5:3074–7.
- [35] Johnson C, Yan Y, Carson R, Martino R, Daube-Witherspoon M. A system for the 3D reconstruction of retracted-septa PET data using the EM algorithm. *IEEE Trans Nucl Sci* 1995;42:1223–7.
- [36] Panin VY, Kehren F, Michel C, Casey M. Fully 3-D PET reconstruction with system matrix derived from point source measurements. *IEEE Trans Med Imag* 2006;25: 907–21.
- [37] Sureau FC, Reader AJ, Comtat C, Leroy C, Ribeiro MJ, Buvat I, Trebossen R. Impact of image-space resolution modeling for studies with the high-resolution research tomograph. *J Nucl Med* 2008;49:1000–8.
- [38] Barbee DL, Flynn RT, Holden JE, Nickles RJ, Jera R. Partial volume correction of PET-imaged tumor heterogeneity using expectation maximization with a spatially varying point spread function. *Phys Med Biol* 2010;55:221–36.
- [39] Zhang X, Zhou J, Cherry SR, Badawi RD, Qi J. Quantitative image reconstruction for total-body PET imaging using the 2-meter long EXPLORER scanner. *Phys Med Biol* 2017;62:2465–85.

- [40] Cho ZH, Ra JB, Hilal SK. True 3D reconstruction – (TTR) application of algorithm toward full utilization of oblique rays. *IEEE Trans Med Imag* 1983;1:6–18.
- [41] Kinahan PE, Rogers JG. Analytic 3D image reconstruction using all detected events. *IEEE Trans Nucl Sci* 1989;36:964–8.
- [42] Cherry SR, Badawi RD, Karp JS, Moses WW, Price P, Jones T. Total-body imaging: transforming the role of positron emission tomography. *Sci Trans Med* 2017;9:381–7.
- [43] Badawi RD, et al. First human imaging studies with the EXPLORER total-body PET scanner. *J Nucl Med* 2019;60:299–303.
- [44] Huang Y, Ng MK, Wen Y. A fast total variation minimization method for image restoration. *SIAM Multisc Model Simul* 2008;7:774–95.
- [45] Chan SH, Khoshabeh R, Gibson KB, Gill PE, Nguyen TQ. An augmented Lagrangian method for total variation video restoration. *IEEE Trans Imag Process* 2011;20:3097–111.
- [46] NEMA Standards Publication NU 2–2007: performance measurements of Positron Emission Tomographs. Nat. Elect. Manuf. Assoc., Washington, DC, USA.
- [47] Jan S, et al. GATE: a simulation toolkit for PET and SPECT. *Phys Med Biol* 2004;49:4543–61.
- [48] Kowalski P, et al. Scatter fraction of the J-PET tomography scanner. *Acta Phys Pol B* 2016;47:549–60.
- [49] Kowalski P, et al. Estimating the NEMA characteristics of the J-PET tomograph using the GATE package. *Phys Med Biol* 2018;63. 165008.
- [50] Saint Gobain Crystals 2016 <http://www.crystals.saint-gobain.com>.
- [51] Krzemiński W, et al. Analysis framework for the J-PET scanner. *Acta Phys Pol A* 2015;127:1491–4.
- [52] Raczyński L, et al. Compressive sensing of signals generated in plastic scintillators in a novel J-PET instrument. *Nucl Instrum Meth Phys Res A* 2015;786:105–12.
- [53] Pałka M, et al. Multichannel FPGA based MVT system for high precision time (20 ps RMS) and charge measurement. *J Instrum* 2017;12:P08001.
- [54] Korcyl G, et al. Evaluation of single-chip, real-time tomographic data processing on FPGA SoC devices. *IEEE Trans Med Imag* 2018;37:2526–35.
- [55] Krzemiński W, Gajos A, Kacprzak K, Rakoczy K, Korcyl G. J-PET Framework: software platform for PET tomography data reconstruction and analysis. *SoftwareX* 2020;11.
- [56] Merlin T, Stute S, Benoit D, Bert J, Carlier T, Comtat C, Filipovic M, Lamare F, Visvikis D. CASToR: a generic data organization and processing code framework for multi-modal and multi-dimensional tomographic reconstruction. *Phys Med Biol* 2018;63:5505.
- [57] Panin VY, Zeng GL, Gullberg GT. Total variation regulated EM algorithm. *IEEE Trans Nucl Sci* 1999;46:2202–10.
- [58] Bardsley JM. A theoretical framework for the regularization of poisson likelihood estimation problems. *Inv Probl Imag* 2010;4:11–7.
- [59] Barbano PE, Fokas A, Schonlieb C. Alternating regularisation in measurement and image space for PET reconstruction. In *Proc. Int. Conf. Sampta Singapore*; 2011.
- [60] Burger M, Muller J, Papoutsellis E, Schonlieb CB. Total variation regularization in measurement and image space for PET reconstruction. *Inv Probl* 2014;30. 105003.
- [61] Clementel E, Mollet P, Vandenberghe S. Effect of local TOF kernel miscalibrations on contrast-noise in TOF PET. *IEEE Trans Nucl Sci* 2013;60:1578–88.