THE SPARSITY-PROMOTED SOLUTION TO THE UNDERSAMPLING TOF-PET IMAGING: NUMERICAL SIMULATIONS

D. Lao[1, *], M. W. Lenox[2], and G. Akabani[1, 2]

[1]Department of Nuclear Engineering, Texas A&M University, USA

[2]Texas A&M Institute for Preclinical Studies, Texas A&M University, USA

Abstract—Recently, the limited-angle TOF-PET system has become an active research topic due to the considerable reduction of hardware cost and potential applicability for performing needle biopsy on patients while in the scanner. This undersampling measurement configuration oftentimes suffers from the deteriorated reconstructed images. However, the established theory of Compressed Sampling (CS) provides a potential framework for undertaking this problem, given that the imaged object can be sparse in some transformed domain. In here, we studied using numerical simulations the application of sparsity-promoted framework to TOF-PET imaging for two undersampling configurations. From these simulations, a relationship was obtained between the number of detectors (or the range of angle) and TOF time resolution, which provided an empirical guide of designing a low-cost TOF-PET systems while ensuring good reconstruction quality. Another contribution is the exploration of p-TV regularization, which showed that RMSE (Root of Mean Square Error) and SSIM (Structural Similarity) were optimized when $p = 0.5$. Several sets of representative numerical experiments were executed to validate the proposed methodology, which demonstrates the promising applicability of undersampling TOF-PET imaging.

1. INTRODUCTION

Positron emission tomography (PET) is an imaging modality where the distribution of the positron emitting radionuclide inside the body is determined based on the detection of photons emitted from positron

annihilation [1]. In contrast to computed tomography (CT) and magnetic resonance imaging (MRI) that provide detailed anatomical or morphological information, PET shows great superiority in monitoring the functional metabolism in normal and neoplastic tissues [2]. The applicability of time-of-flight (TOF) information in upgrading PET imaging has been realized since Brownell et al., especially for the purpose of whole-body imaging [3]. The fundamental improvement rendered by TOF is an increase in signal-to-noise ratio (SNR) [3, 4]; therefore, TOF can help improve the quality of reconstructed images. TOF brought resurgence to PET with the development of fast detector system, (i.e., LSO used).

Recently, the limited-angle TOF-PET system has been proposed for breast-dedicated imaging [5], and it turned to be an active area of research due to its distinct advantages, in particular, the capacity of simultaneously performing a needle biopsy on patients while in the scanner. Furthermore, the limited-angle configuration can significantly reduce the hardware manufacturing cost by reducing the number of detectors. Note that this kind of system only can be applicable under the condition that the TOF time resolution is high enough. Nonetheless, artifacts will inevitably appear in the reconstructed images because of the incomplete coverage of imaged object by a partial PET ring. From a mathematical point of view, this reconstruction scheme is a serious ill-posed inverse problem due to the deficient observations compared with unknowns. Consequently, there are an infinite number of solutions, which can match the observed data, but cannot provide meaningful interpretation for clinical purposes.

One of the popular methods to address this problem is regularization, whose idea is to restrict the data fidelity by augmenting the penalty term characterizing prior information of the object. Furthermore, a large number of medical images are shown to be piecewise constant, and they also have sparse representations in some transformed domain like discrete cosine transformation (DCT) and wavelet [6], or some trained over-complete dictionary [7–9]. It means that sparse regularizations are good for the reconstruction of medical imaging. As a matter of fact, since the invention of compressed sampling (CS) developed by Candes et al. [15], the sparsity-promoted solution [15, 16] has been widely explored in medical imaging problems, such as, MRI [8, 9, 17], CT [18, 19], and PET [20–22]. Encouraged by these observations, the sparsity-promoted regularization was employed to address our undersampling TOF-PET imaging problem mentioned above.

It is well known that CS is a revolutionary signal sampling paradigm, where a N-dimensional signal can be faithfully recovered

from much fewer measurements than N, given that the signal of interest is itself sparse or sparse represented in some transformed domain such as DCT, wavelet and so on [15, 16]. Due to the potential improvement on reconstruction performance over traditional approaches and dramatic cost reduction in data acquisition, CS methods have been extensively investigated in many fields covering medical imaging [8, 9, 17–22], communication [23, 24], and image processing [6, 7, 10–12].

As for the convex optimization involved in CS, the objective function to be minimized consists of two components, the first represents the data fidelity, and the second is the sparsity-promoted regularized term. Traditionally, to impose the piecewise-constant property of medical images the role of sparse-promoted penalty is usually assigned to the total variation (TV) regularization. More generic statements on sparsity through l_p-norm constraint have also been made (see the details in [13, 15, 16]).

Regarding the application of sparse reconstruction in PET imaging, the research has been focused on the development of efficient reconstruction algorithms [20–22]. To the best of our knowledge, the first work within the context of the sparsity-promoted reconstruction can be traced to that made by Harmany et al. [20]. In [20], the photon-limited collections are modeled as Poisson distribution and the choice of wavelet-based sparse transformation is discussed. Furthermore, to correct the estimation bias caused by the mandatory introduction of the sparsity-promoted term into the original data fidelity, Wang and Qi proposed a two-step reconstruction strategy [21], where the sparsity-promoted algorithm is first implemented to determine the basis components with significant coefficients, and then the final solution is straightforward obtained by solving the reduced-dimensional least-square problem. In [22], the authors briefly introduce the application of the first-order primal-dual algorithm to deal with real PET data. In summary, these studies show that CS is capable of breaking the bottleneck of PET imaging with fewer observed data than unknowns.

In here we made the first attempt to investigate the potential benefit of sparsity-promoted reconstruction for the undersampling TOF-PET imaging. The combined use of TV and l_1-norm regularizations is preferred for sparsity-promoted medical imaging problems, in particular, the former is to make use of the property of piecewise constant enjoyed by most of medical images, and the latter is to take advantage of their intrinsic correlated structured property. Additionally, such combination renders us another benefit that the l_1-norm regularization can obviously suppress the artifacts introduced by TV [17]. Our first contribution is the development of an efficient

algorithm that deals with the combination of l_1-norm and p-TV, where p-TV is the direct extension of traditional TV regularization. Specifically, p-TV becomes the standard TV when $p = 1$. Our numerical simulations show that 0.5-TV ($p = 0.5$) is optimal for our problem. The second contribution is to extensively investigate two undersampling TOF PET configurations within the framework of sparse reconstruction. One is similar as the current traditional PET configuration, except that fewer detectors are uniformly and sparsely distributed over a complete ring, and the other one is the so-called limited-angle PET system [5], where the detectors with density of $384/360°$ are distributed over two opposite partial rings. We carried out numerous numerical experiments to study the reconstruction quality dependent on the number of detectors (or the range of angle) and the TOF time resolution, which provides an empirical guide for designing a low-cost undersampling TOF-PET system. The results show that there is a trade-off between the choices of number detectors (or the range of angle) and the TOF-PET time resolution.

The remainder of this paper is organized as follows: In Section 2, the methodology for solving the optimization problem with combination of p-TV and l_1-norm regularizations is described, and its associated reconstruction algorithm is introduced. In Section 3, four sets of numerical experiments are discussed. The first set gives a comparison between the traditional EM (Expected Maximum) algorithm and the proposed sparsity-promoted algorithm to show the advantages of applying sparsity-promoted framework to the TOF-PET imaging; the second set is used to study the effect of p-TV on the reconstruction quality; the third and fourth sets are used to discuss the relationship between the number of detectors (or the range of angle), TOF time resolution and the reconstruction performance in terms of RMSE and 1-SSIM. Finally some conclusions are summarized in Section 4.

2. METHODOLOGY

Formally expressed, the TOF-PET measurement can be approximately represented by the following convolution of the scene intensity f and a TOF kernel function h along the line-of-response (LOR) (referring to Figure 1), in particular [25, 26],

$$y(s, \theta, t) = \int_{-\infty}^{\infty} f(s\cos\theta - l\sin\theta, s\sin\theta + l\cos\theta)h(t - l)dl. \quad (1)$$

Here, the variables s and θ are the radial and angular coordinates, respectively, while t is the so-called TOF variable. The kernel h is
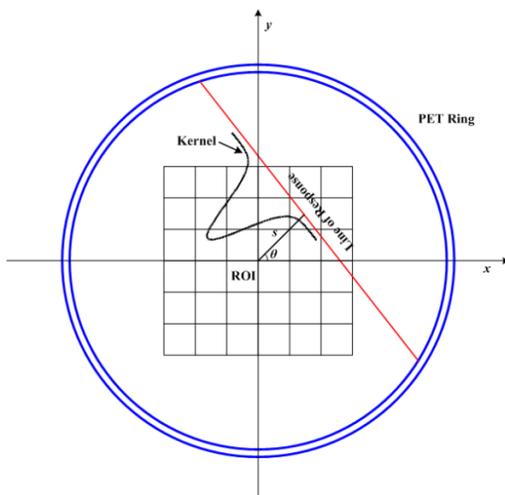
Figure 1. Sketch map of TOF-PET imaging system where the red line denotes the LOR (line of response). The curve marked as kernel is the TOF Gaussian function and the ROI (region of interest) is discretized into $n \times n$ pixels.

often considered to be time-shift invariant and modeled as a Gaussian function [26].

To facilitate the implementation of numerical simulation, the continuous Equation (1) is usually casted into a series of compact forms, i.e.,

$$\mathbf{y}_t = \mathbf{A}\mathbf{H}_t\mathbf{f} = \mathbf{A}_t\mathbf{f},$$
$$t = \{-T\Delta t, -(T-1)\Delta t, \dots, (T-1)\Delta t, T\Delta t\} \tag{2}$$

where Δt is known as TOF bin, and the number of bins is $(2T+1)$. $\mathbf{A}$ is the projection matrix of PET system, whose entries characterize the probability $p(i, j)$ that an emission from j-th pixel is detected by i-th pair of detectors of PET ring [1–5, 26]. In Equation (2), $\mathbf{y}_t \in R^M$ is a M-dimensional vector representing $y(s, \theta, t)$ at the moment of t. $\mathbf{H}_t\mathbf{f}$ accounts for the convolution of Gaussian kernel and object, where $\mathbf{H}_t$ is a Topletiz matrix generated from h. Note that without otherwise specific claim, the notation of $\mathbf{f}$ is a N-length vector, which is stacked from the two-dimensional image $\mathbf{f}$ with size of $n \times n$.

The reconstruction of $\mathbf{f}$ can be addressed through the following least square problem, in particular,

$$\hat{\mathbf{f}} = \arg\min_{\mathbf{f}} \sum_{t=-T}^{T} \|\mathbf{y}_t - \mathbf{A}_t\mathbf{f}\|_2^2. \tag{3}$$

As pointed out in previous section, almost all the traditional methods fail to produce meaningful solution to (3) for undersampling TOF-PET scenarios in most cases, especially for the case of noisy observation. To circumvent this problem, two priors of piecewise constant and structure-correlated sparse are incorporated to the objective function of Equation (3), accordingly,

$$\hat{\mathbf{f}} = \arg\min_{\mathbf{f}} \left\{ \sum_{t=-T}^{T} \|\mathbf{y}_t - \mathbf{A}_t\mathbf{f}\|_2^2 + \gamma_1 \|\mathbf{f}\|_{p-TV} + \gamma_2 \|\Phi\mathbf{f}\|_1 \right\}. \quad (4)$$

In Equation (4), $\|\mathbf{f}\|_{p-TV}$ is defined as

$$\|\mathbf{f}\|_{p-TV} = \sum_{i,j} \left((\Delta^x\mathbf{f}_{i,j})^2 + (\Delta^y\mathbf{f}_{i,j})^2 \right)^{p/2} \quad (5)$$

and

$$\Delta^x\mathbf{f}_{i,j} = \mathbf{f}_{i+1,j} - \mathbf{f}_{i,j}, \quad \Delta^y\mathbf{f}_{i,j} = \mathbf{f}_{i,j+1} - \mathbf{f}_{i,j}, \quad (6)$$

where $\mathbf{f}$ in Equations (5) and (6) is the $n \times n$ matrix representation of discretized object. The regularization factors of γ_1 and γ_2 allow the tradeoff between the data fidelity and the priors of $\mathbf{f}$, which should be carefully determined. Φ is a suitable sparse transformed basis, and specified as discrete cosine transformation (DCT) here.

It is noted that the objective function of Equation (4) is analogous to one proposed by Lustig et al. [17]; however, compared to [17] and other relevant works this work presents two specific differences: (a) instead of implementing a conjugate gradient algorithm to the whole objective function, we provide an alternative approach to solve Equation (4), i.e., separately dealing with p-TV regularization term with iteratively reweighted method, while applying a shrinkage-thresholding operation on l_1-norm constraint; (b) the generalized TV, i.e., p-TV, instead of standard TV regularization is adopted. As a matter of fact, $\|\mathbf{f}\|_{p-TV}$ is the straightforward extension of standard TV-regularization, which was firstly mentioned by Rodriguez and Wohlberg [14]. However, in [14], only the 1-TV and 2-TV are numerically studied for imaging painting and de-noise. In the subsequent section we will discuss the performances of different p-TV adapted for our TOF-PET imaging problem, which show that the 0.5-TV is optimal.

To efficiently solve the Equation (4), we introduce an auxiliary variable $\mathbf{d} = \Phi\mathbf{f}$; consequently, Equation (4) can be equivalently expressed by the following two-fold optimization problem

$$\min_{\mathbf{f},\mathbf{d}} \left\{ \sum_{t=-T}^{T} \|\mathbf{A}_t\mathbf{f} - \mathbf{y}_t\|_2^2 + \gamma_1 \|\mathbf{f}\|_{p-TV} + \gamma_2 \|\mathbf{d}\|_1 \right\}, \quad \text{s.t. } \mathbf{d} = \Phi\mathbf{f}. \quad (7)$$

Furthermore, the Lagrange form of the above constrained optimization problem can be readily derived as

$$\min_{\mathbf{f},\mathbf{d}} \left\{ \sum_{t=-T}^{T} \|\mathbf{A}_t\mathbf{f} - \mathbf{y}_t\|_2^2 + \gamma_1 \|\mathbf{f}\|_{p-TV} + \gamma_2 \|\mathbf{d}\|_1 + \gamma_3 \|\mathbf{d} - \Phi\mathbf{f}\|_2^2 \right\}. \quad (8)$$

Intuitively, the classical alternatively iterative strategy can be exploited to treat Equation (8), the working procedure consists of alternating between performing the update of $\mathbf{f}$ using the current estimate for $\mathbf{d}$, and computing $\mathbf{d}$ based on the update of $\mathbf{f}$. In this work, the maximum iteration number and minimum tolerance error are employed as criterions to stop the iterative process, and set to be 60 and 10^{-5}, respectively. The tolerance error is defined as

$$Err = \frac{\displaystyle\sum_{t=-T}^{T} \|\mathbf{y}_t - \mathbf{A}_t\mathbf{f}\|_2}{\displaystyle\sum_{t=-T}^{T} \|\mathbf{y}_t\|_2}. \quad (9)$$

Assuming we have the solutions of $\mathbf{f}$ and $\mathbf{d}$ at the $(n-1)$-th iteration, then the updates for $\mathbf{f}$ and $\mathbf{d}$ at the n-th iteration are made by the following strategy, specifically, alternatively solving the following two sub-optimization problems represented by Equation (10) and Equation (11), namely,

$$\mathbf{f}^{(n)} = \arg\min_{\mathbf{f}} \left\{ \sum_{t=-T}^{T} \|\mathbf{A}_t\mathbf{f} - \mathbf{y}_t\|_2^2 + \gamma_1 \|\mathbf{f}\|_{p-TV} + \gamma_3 \left\|\mathbf{d}^{(n-1)} - \Phi\mathbf{f}\right\|_2^2 \right\} \quad (10)$$

and

$$\mathbf{d}^{(n)} = \arg\min_{\mathbf{d}} \left\{ \gamma_2 \|\mathbf{d}\|_1 + \gamma_3 \left\|\mathbf{d} - \Phi\mathbf{f}^{(n)}\right\|_2^2 \right\}, \quad (11)$$

respectively.

2.1. Step I: Update of f

In here we would like to solve Equation (10) by using the strategy of iteratively reweighted approach [32]. The term of $\|\mathbf{f}\|_{p-TV}$ can be

equivalently expressed as

$$\|\mathbf{f}\|_{p-TV} = \sum_{i,j} \left((\Delta^x \mathbf{f}_{i,j})^2 + (\Delta^y \mathbf{f}_{i,j})^2 \right)^{p/2}$$

$$= \sum_{i,j} \frac{(\Delta^x \mathbf{f}_{i,j})^2}{\left((\Delta^x \mathbf{f}_{i,j})^2 + (\Delta^y \mathbf{f}_{i,j})^2 \right)^{1-p/2}} + \sum_{i,j} \frac{(\Delta^x \mathbf{f}_{i,j})^2}{\left((\Delta^x \mathbf{f}_{i,j})^2 + (\Delta^y \mathbf{f}_{i,j})^2 \right)^{1-p/2}}$$

$$= \|\mathbf{D}_x \mathbf{f}\|_{\mathbf{W}}^2 + \|\mathbf{D}_y \mathbf{f}\|_{\mathbf{W}}^2 \approx \|\mathbf{D}_x \mathbf{f}\|_{\mathbf{W}^{(n-1)}}^2 + \|\mathbf{D}_y \mathbf{f}\|_{\mathbf{W}^{(n-1)}}^2 , \tag{12}$$

where

$$(\mathbf{D}_x \mathbf{f})_{i,j} = \Delta^x \mathbf{f}_{i,j}, \quad (\mathbf{D}_y \mathbf{f})_{i,j} = \Delta^y \mathbf{f}_{i,j} \tag{13}$$

and

$$\mathbf{W}^{(n-1)} = \mathrm{diag}\left(\left(\left(\Delta^x \mathbf{f}_{i,j}^{(n-1)} \right)^2 + \left(\Delta^y \mathbf{f}_{i,j}^{(n-1)} \right)^2 \right)^{p/2-1} \right). \tag{14}$$

In order to avoid singularity of $((\Delta^x \mathbf{f}_{i,j}^{(n-1)})^2 + (\Delta^y \mathbf{f}_{i,j}^{(n-1)})^2)$, a small positive real ε is introduced to modify $\mathbf{W}$, i.e.,

$$\mathbf{W}^{(n-1)} = \mathrm{diag}\left(\left(\left(\Delta^x \mathbf{f}_{i,j}^{(n-1)} \right)^2 + \left(\Delta^y \mathbf{f}_{i,j}^{(n-1)} \right)^2 + \varepsilon \right)^{p/2-1} \right). \tag{15}$$

Now, we can readily derive the iterative solution to (10) as

$$\mathbf{f}^{(n)} = \arg\min_{\mathbf{f}} \left\{ \sum_{t=-T}^{T} \|\mathbf{A}_t \mathbf{f} - \mathbf{y}_t\|_2^2 + \gamma_3 \left\| \mathbf{d}^{(n-1)} - \Phi \mathbf{f} \right\|_2^2 \right.$$

$$\left. + \gamma_1 \left(\|\mathbf{D}_x \mathbf{f}\|_{\mathbf{W}^{(n-1)}}^2 + \|\mathbf{D}_y \mathbf{f}\|_{\mathbf{W}^{(n-1)}}^2 \right) \right\}. \tag{16}$$

Explicitly, the closed-form estimation of $\mathbf{f}^{(n)}$ can be achieved as

$$\mathbf{f}^{(n)} = \left(\sum_{t=-T}^{T} \mathbf{A}'_t \mathbf{A}_t + \gamma_1 \left(\mathbf{D}'_x \mathbf{W}^{(n-1)} \mathbf{D}_x + \mathbf{D}'_y \mathbf{W}^{(n-1)} \mathbf{D}_y \right) + \gamma_3 \mathbf{I} \right)^{-1}$$

$$\left(\gamma_3 \Phi' \mathbf{d}^{(n-1)} + \sum_{t=-T}^{T} \mathbf{A}'_t \mathbf{y}_t \right). \tag{17}$$

To avoid the overwhelming computation of calculating matrix inverse involved in Equation (17), the classical preconditioned conjugate gradient (PCG) method was explored to solve Equation (16) in this work.

2.2. Step II: Update of d

Through the standard implementation widely made in the literature of compressed sampling, the shrinkage-thresholding solution $\mathbf{d}^{(n)}$ to problem (11) can be obtained as [27]

$$\mathbf{d}^{(n)} = \arg\min_{\mathbf{d}} \left\{ \gamma_2 \|\mathbf{d}\|_1 + \gamma_3 \left\| \mathbf{d} - \Phi\mathbf{f}^{(n)} \right\|_2^2 \right\} = Soft\_Thr\left(\Phi\mathbf{f}^{(n)}, \frac{\gamma_2}{\gamma_3} \right). \quad (18)$$

Finally, the whole procedure of the proposed algorithm for solving sparsity-promoted TOF-PET reconstruction (i.e., Equation (4)) has been summarized in Table 1.

Regarding the choices of regularization parameters γ_1, γ_2, and γ_3, they should be carefully chosen. Specifically, if the regularization parameters are large, the resulting solution will seriously stray from the true solution; on the other hand, small regularization parameters will cause the objective function a highly ill-conditioned problem, and extremely slow convergence of PCG. Though some efforts have been made such as L-curve [33], the generalized cross validation method [34], Stein's Unbiased Risk Estimator (SURE) [35], and so on, this problem is still an open challenging question mainly due to the computational consideration. In this paper the choices of γ_1, γ_2, and γ_3 are empirically determined though a large number of numerical experiments.

Table 1. Procedure for solving Equation (4) using the sparsity-promoted algorithm.

Initial setup: Initial setup: $\mathbf{d}^{(0)} = \mathbf{0}$; $\mathbf{W}^{(0)} = \mathbf{I}$; $\mathbf{f}^{(0)} = \mathbf{0}$; $n = 1$

While $Err \geq 10^{-5}$ ＆ $n \leq 50$

$$\mathbf{f}^{(n)} = \left(\sum_{t=-T}^{T} \mathbf{A}'_t\mathbf{A}_t + \gamma_1 \left(\mathbf{D}'_x\mathbf{W}^{(n-1)}\mathbf{D}_x + \mathbf{D}'_y\mathbf{W}^{(n-1)}\mathbf{D}_y \right) + \gamma_3\mathbf{I} \right)^{-1}$$

$$\left(\gamma_3\Phi'\mathbf{d}^{(n-1)} + \sum_{t=-T}^{T} \mathbf{A}'_t\mathbf{y}_t \right)$$

$$\mathbf{d}^{(n)} = Soft\_Thr\left(\Phi\mathbf{f}^{(n)}, \frac{\gamma_2}{\gamma_3} \right)$$

$$\mathbf{W}^{(n)} = \text{diag}\left(\left(\left(\Delta^x\mathbf{f}^{(n)}_{i,j} \right)^2 + \left(\Delta^y\mathbf{f}^{(n)}_{i,j} \right)^2 + \varepsilon \right)^{p/2-1} \right)$$

$$Err = \frac{\sum\limits_{t=-T}^{T} \|\mathbf{A}_t\mathbf{f}^{(n)} - \mathbf{y}_t\|_2^2}{\sum\limits_{t=-T}^{T} \|\mathbf{y}_t\|_2^2}$$

$\gamma_1 = 0.8 \times \gamma_1$

$n = n + 1$

End While

Furthermore, the popular continuation technique was applied on γ_1 by initially setting a large γ_1 and gradually decreasing its value as the iteration is preceded [28]. Therefore, the optimized solution with previous value of γ_1 provides a warm start for the next iteration. In our implementation, we adopted the continuation technique and empirically chose $\gamma_1 = 0.8 \times \gamma_1$.

3. NUMERICAL SIMULATION AND DISCUSSION

In this section, numerical experiments were carried out to investigate the benefits of the sparsity-promoted reconstruction in TOF-PET imaging, where the tradeoff between time resolution and number of detectors (or range of angle) is particularly of concern. The phantom used in the simulation is the well-known Shepp-Logan phantom, as shown in Figure 2. We considered two different undersampling measurement configurations. The first one is parallel to the current traditional PET system, except fewer detectors are sparsely and uniformly distributed over the complete PET ring, as sketched in Figure 3(a). The second one is the limited-angle PET system represented in [5], where the detectors are furnished with the density of $384/360°$ over two opposite partial PET rings, as shown in Figure 3(b). For convenience, the first configuration is referred as configuration A while the second one is called configuration B. In this work, the radius of PET ring is set to be 35 cm, while the scale of ROI into which the phantom is embedded is 30 cm by 30 cm. The phantom used in the study is the well-known Shepp-Logan simulated using MATLAB tool (i.e., phantom), which is discretized into 128 by 128 pixels.
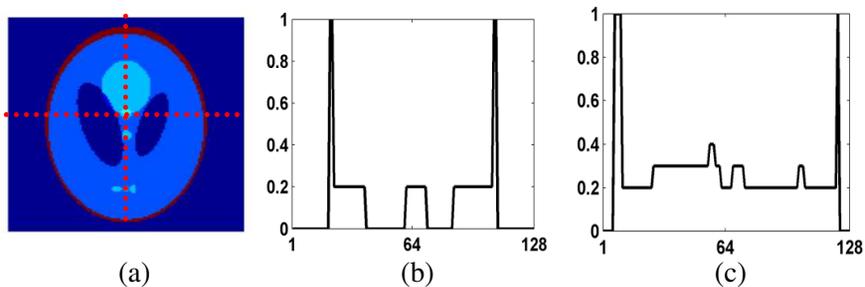

Figure 2. (a) Ground truth of 128×128 Shepp–Logan phantom used in numerical simulations; (b) horizontal profile that corresponds to the slice indicated by horizontal dotted line in panel (a); (c) vertical profile which corresponds to the slice indicated by vertical dotted line in panel (a).
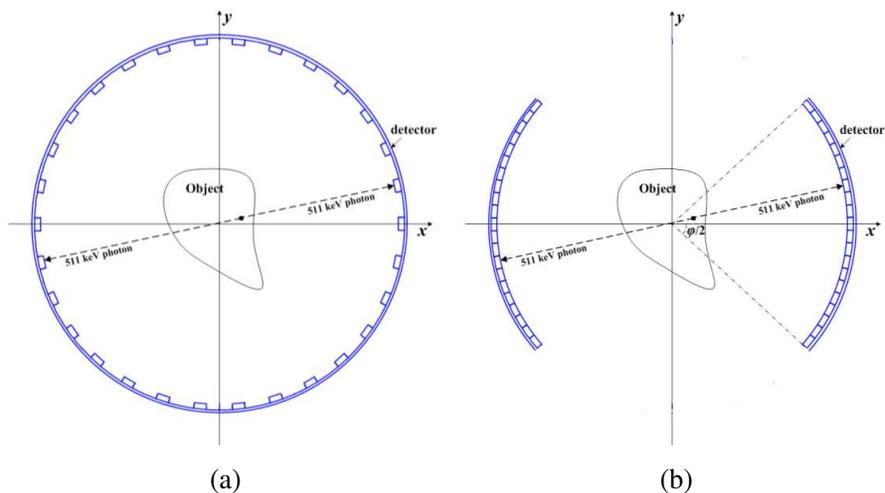
(a)   (b)

Figure 3. Sketch maps of two undersampling TOF-PET measurement configurations, (a) configuration A with detectors distributed sparsely and uniformly over the full PET ring; (b) configuration B with detectors distributed compactly over two opposite partial rings, where φ denotes the range of angle.

This section is arranged as follows. First, we show the superiority of sparsity-promoted approach in dealing with undersampling observation mentioned previously. Second, the solution through the combined regularizations of p-TV and l_1-norm is examined, which empirically shows that $p = 0.5$ is the best candidate to achieve acceptable solution for our problem. Third, the reconstruction performances dependent on system time resolution t, and the numbers of detectors N for configuration A while the range of angle for configuration B are investigated in Subsections 3.3 and 3.4, respectively.

In this paper, RMSE (Root of Mean Square Error) and SSIM (Structural Similarity) [29] are used as criterions to assess the reconstruction quality, whose definitions are

$$\text{RMSE} = \frac{\|\mathbf{f}_{rec} - \mathbf{f}_{true}\|}{\|\mathbf{f}_{true}\|} \tag{19}$$

and

$$\text{SSIM} = \frac{\left(2\mu_{rec}\mu_{true} + c_1\right)\left(2\sigma_{cross} + c_2\right)}{\left(\mu_{rec}^2 + \mu_{true}^2 + c_1\right)\left(\sigma_{rec}^2 + \sigma_{true}^2 + c_2\right)} \tag{20}$$

respectively. Here, μ_{rec} and μ_{true} are the average of the reconstructed

image ($\mathbf{f}_{rec}$) and true image ($\mathbf{f}_{true}$), respectively; σ^2_{rec} and σ^2_{true} are the associated variances of them, σ_{cross} is the covariance of $\mathbf{f}_{rec}$ and $\mathbf{f}_{true}$, c_1 and c_2 are two variables to stabilize the division with weak dominator.

3.1. Numerical Test 1

In this subsection, a set of numerical simulations under configuration A is carried out to demonstrate the superiority of the proposed sparsity-promoted solver over traditional EM algorithm, where PET with and without TOF information cases are considered. In here, the time resolution and TOF bin are set as 500 ps and 67 ps, respectively [31].

Firstly we consider the PET imaging without TOF. With 50, 70, 110, 170 and 190 detectors used, Figure 4(a) shows the reconstructed images and their associated horizontal profiles obtained through the proposed sparsity-promoted algorithm. For comparison, the corresponding reconstructed results using traditional EM algorithms are reported in Figure 4(b). It shows intuitively that the sparsity-promoted method is able to provide much better results than the traditional EM algorithm. Moreover, we also report the RMSEs and (1-SSIM)s of the reconstructed PET images in Table 2.

Secondly, a similar analysis on TOF-PET imaging is also carried out. The reconstructed images and horizontal profiles through proposed sparsity-promoted and EM algorithms are shown in Figures 5(a) and 5(b), respectively. Also the corresponding RMSEs and (1-SSIM)s are included in Table 2. It is verified again that the sparsity-promoted method gave much better results than the traditional EM algorithm.

Table 2. RMSEs and (1-SSIM)s of PET and TOF-PET images reconstructed through sprasity-promoted and EM algorithms for different detectors used, where measurement configuration A is assumed.

| Number of detector		50	70	110	190	270
Non-TOF	RMSE	0.4563	0.3994	0.2364	0.1797	0.0338
EM	1-SSIM	0.7547	0.3144	0.2142	0.1681	0.0470
Non-TOF	RMSE	0.3976	0.3126	4.44×10^{-4}	1.92×10^{-4}	1.8×10^{-4}
Sparse	1-SSIM	0.1119	0.1251	6.08×10^{-6}	5.40×10^{-6}	5.33×10^{-6}
TOF	RMSE	0.3095	0.2515	0.2156	0.2177	0.0320
EM	1-SSIM	0.2361	0.1706	0.1445	0.0914	0.0351
TOF	RMSE	0.2599	0.1046	3.24×10^{-4}	1.79×10^{-4}	1.66×10^{-4}
Sparse	1-SSIM	0.0407	0.0024	5.54×10^{-6}	5.33×10^{-6}	5.31×10^{-6}
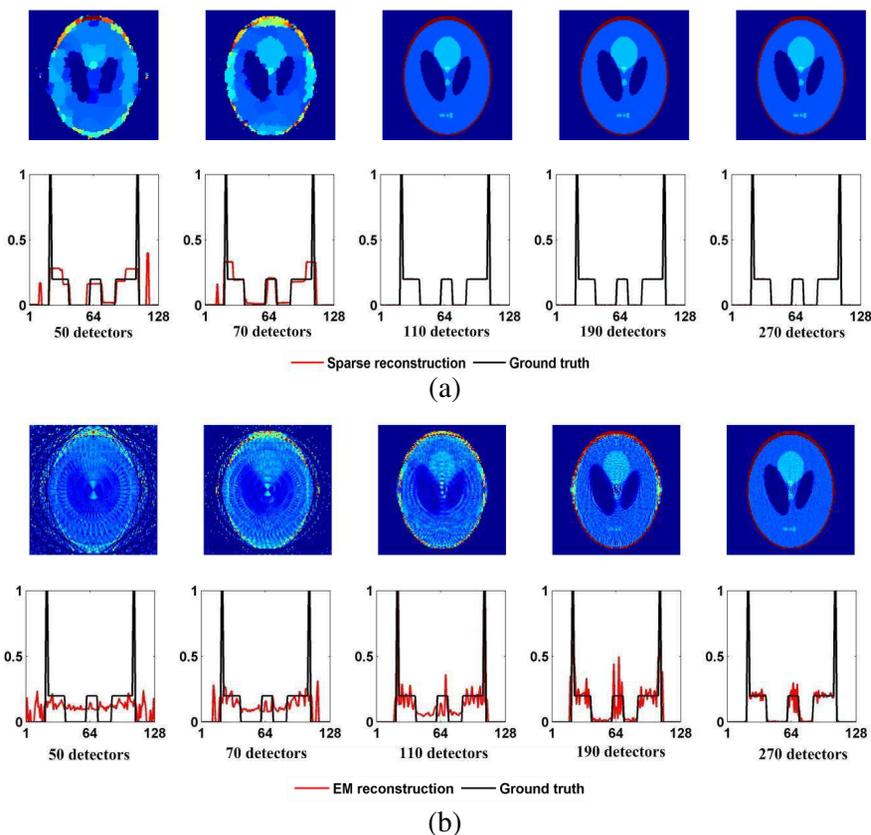
Figure 4. PET images reconstructed using (a) the proposed sparsity-promoted and (b) the traditional EM algorithms for different number of detectors used. The corresponding horizontal profiles are also illustrated via red solid line.

From the above results, we can verify that TOF information is favorable to enhance image quality. More importantly, we notice that the sparsity-promoted approach to PET imaging performs significantly better than EM algorithm, in the sense of whether visual observation or RMSE and 1-SSIM. Specifically, for TOF-PET imaging the sparse solver can provide a visually acceptable solution when 70 detectors are used; in contrast, 190 detectors are required by traditional EM algorithm. Even in the absence of TOF information, the sparsity-promoted approach can provide almost exact reconstruction with only 110 detectors; however, the reconstruction images obtained from EM algorithm suffer from serious salt and pepper noise, even with 270
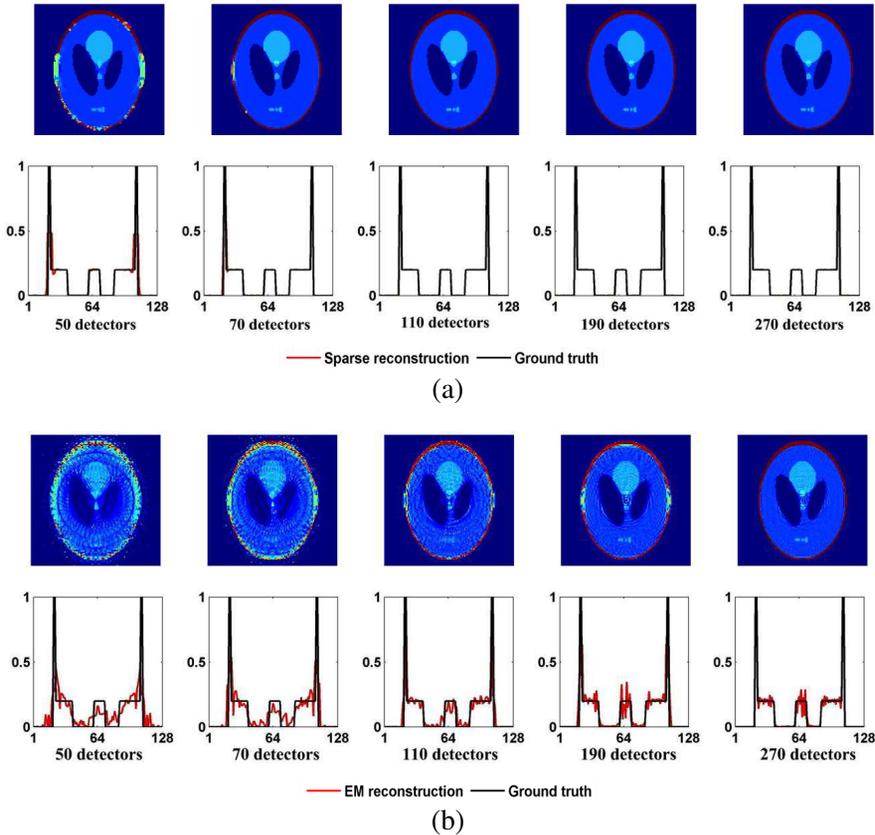
(a)

(b)

Figure 5. TOF-PET images reconstructed using (a) the proposed sparsity-promoted and (b) the traditional EM algorithms for different number of detectors used. The corresponding horizontal profiles are also illustrated via red solid line.

detectors.

From Table 2, we can quantitatively observe the benefits rendered by the sparsity prior and TOF information. Overall, the RMSE and 1-SSIM levels of the reconstructed images using the sparsity-promoted algorithm are much lower than those obtained using the EM algorithm. For instance, with 110 detectors used, RMSEs of images reconstructed by EM algorithm are 0.2156 and 0.2364 for the modalities with and without TOF information. However, they can be enhanced to 3.24×10^{-4} and 4.44×10^{-4}, respectively, with the sparsity-promoted algorithm implemented. For the cases of much fewer detectors, e.g., 50 detectors and 70 detectors, the sparsity-promoted

algorithm still significantly outperforms EM algorithm.

Regarding the horizontal profiles shown in Figures 4(a), 5(a) and 4(b), 5(b) corresponding to the reconstructed images obtained from the sparsity-promoted and EM algorithms, we can see that the profiles obtained from sparsity-promoted algorithm are almost indistinguishable from the ground truth, and they are more exact and smooth than those obtained from EM algorithm. Furthermore, they demonstrate the capability of the sparsity-promoted algorithm in reducing the artifacts produced in reconstruction process.

From the above results, we can conclude that a substantial improvement in reconstructed image quality is expected with the combination of TOF information and sparsity-promoted algorithm.

3.2. Numerical Test 2: p-TV

As we known, the different choices of variable p involved in p-TV will produce different effects on the TOF-PET reconstruction quality. Roughly, the sparse solution in the sense of gradient counterpart can be promoted when $0 \leq p \leq 1$; in contrast, no sparsity is enforced otherwise. Furthermore, it is well known that $p = 0$, $p = 1$ and $p = 2$ are three candidates most widely used in dealing with inverse problems. Specifically, $p = 0$ corresponds to the so-called support detector [30]. It usually presents computational challenge due to the intrinsic non-convexity. $p = 1$ is known as the standard TV regularizer [11], and $p = 2$ is the well-known Tikhnov regularization. So far, it hasn't been clear yet what the optimal value of p is for efficiently dealing with practical problem. In here, we intend to obtain a suitable value of p for an optimal reconstruction. Same as in numerical test 1, this set of experiments is proceeded under configuration A, and the time resolution and TOF bin are set to be 500 ps and 67 ps, respectively. Figure 6 shows a series of reconstructed TOF-PET images corresponding to $p = 0$, 0.5, 1, 1.5 and 2 and Figure 7 shows their associated convergences. Table 3 presents their associated RMSEs and (1-SSIM)s. From Figures 6 and 7, we can observe that $p = 0.5$ is capable of achieving a solution matching the observed data best; therefore, it has the strongest ability of avoiding trapping into the local minimum. At the same time, Figures 8(a) and 8(b) give the dependences of RMSE and 1-SSIM on the values of p varying from 0 through 2 with an interval of 0.1, which support the standpoint that $p = 0.5$ or so is the best candidate to achieve sparsity-promoted solution for our problem. In summary, we can conclude that $p = 0.5$ provides best TOF-PET reconstruction from Figure 6 through Figure 8.
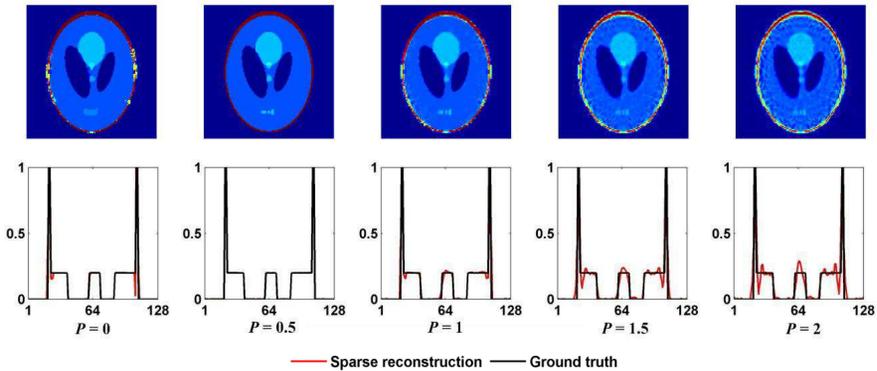
Figure 6. Reconstructed TOF-PET images and their associated horizontal profiles corresponding to different values of $p = 0$, 0.5, 1, 1.5 and 2 for p-TV.
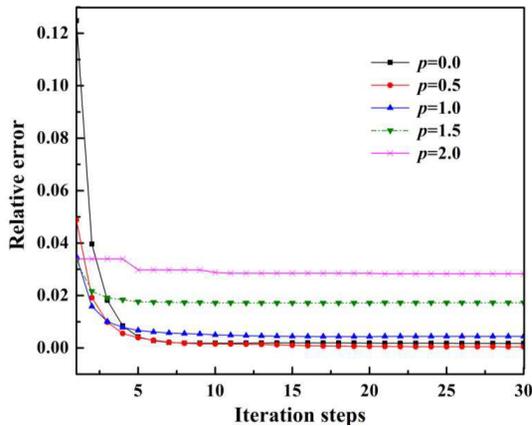

Figure 7. Convergence curves of the sparsity-promoted algorithm for different values of $p = 0$, 0.5, 1, 1.5 and 2 for p-TV.

Table 3. RMSEs and (1-SSIM)s of TOF-PET images reconstructed through sparsity-promoted algorithm with different choices of p-TV, where configuration A is used.

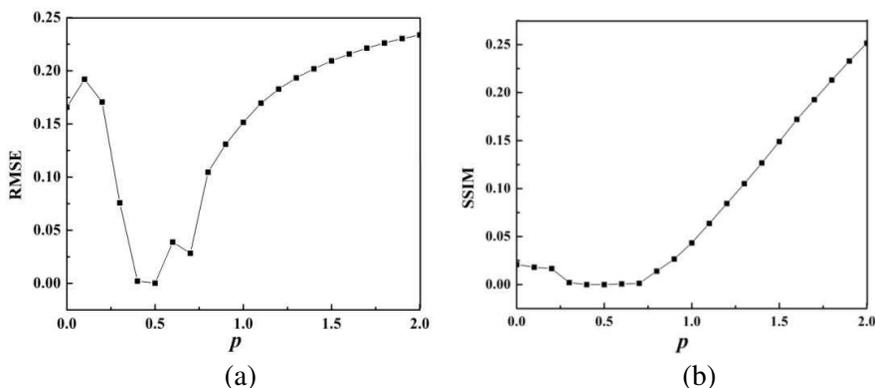| P	0.0	0.5	1.0	1.5	2.0
RMSE	0.1658	0.0001	0.1515	0.2094	0.2303
1-SSIM	0.0209	5.72×10^{-6}	0.0434	0.1490	0.2329

Figure 8. Reconstruction performance with the different choices of p value for p-TV under configuration A with 70 detectors. (a) RMSE versus p, and (b) 1-SSIM versus p.

3.3. Numerical Test 3: Investigation on Configuration A

Through a set of numerical experiments associated with configuration A, we studied the relationship between the RMSE and 1-SSIM of images reconstructed through sparsity-promoted algorithm, the time resolution t of TOF-PET system, and the numbers of detectors N. The basic purpose of this investigation is to find the sufficient condition of getting acceptable TOF-PET reconstructions for limited detectors uniformly distributed over a complete PET ring. We would like to mention that normally the parameters of γ_1, γ_2 and γ_3 in our sparsity-promoted algorithm are empirically tuned and set to be 1×10^{-2}, 1×10^{-4}, and 1×10^{-8}, respectively.

Reconstructed images with different time resolutions of 100 ps, 700 ps, 1300 ps, 1900 ps and 2500 ps obtained through the sparsity-promoted method are shown in Figure 9, where 70 detectors were used. Their corresponding horizontal profiles are also provided. From this set of figures one can observe that the reconstruction quality gets worse as the time resolution increases. Their associated RMSEs and (1-SSIM)s are presented in Table 4.

The generic dependences of RMSEs and (1-SSIM)s on the number of detectors N and the TOF time resolution t are shown in Figures 10(a) and 10(b), respectively, where x-axis denotes the time resolution t in ps while y-axis is for number of detectors N. In this set of simulations, the TOF time resolution t ranges from 100 ps to 3000 ps, with interval of 100 ps, while the number of detectors N varies from 50 through 110 with interval of 10. From Figure 10, we can see a tradeoff
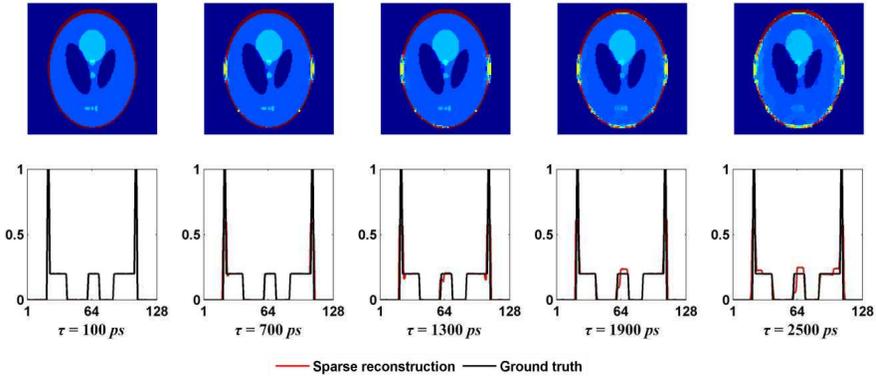
Figure 9. TOF-PET images reconstructed through sparsity-promoted algorithm and their corresponding horizontal profiles for different TOF time resolutions under configuration A using 70 detectors.
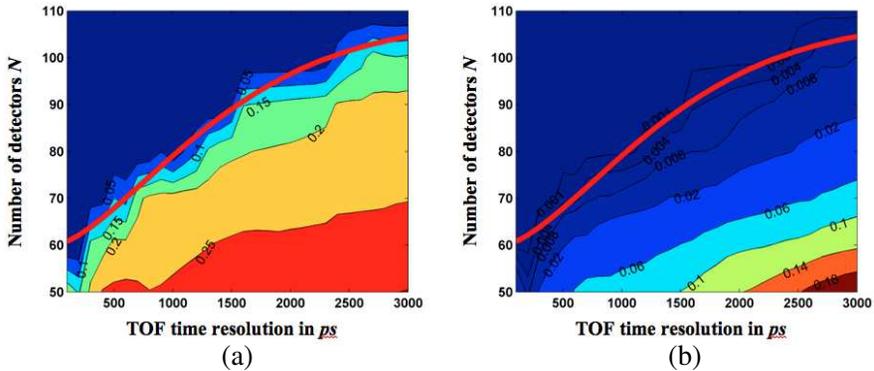

Figure 10. Contours of (a) RMSE and (b) 1-SSIM of TOF-PET reconstructed images as a function of numbers of detectors N and TOF resolution t, for configuration A. The red solid lines represent the phase transition curves with a RMSE threshold of 0.1.

Table 4. RMSEs and (1-SSIM)s of TOF-PET images reconstructed through sparsity-promoted algorithm with different TOF-PET time resolutions, where configuration A with 70-detectors is used.

Time resolution (ps)	100	700	1300	1900	2500
RMSE	1.99×10^{-4}	0.1887	0.2263	0.2274	0.2422
1-SSIM	5.04×10^{-6}	0.0067	0.0148	0.0284	0.0463

Table 5. RMSEs and (1-SSIM)s of TOF-PET images reconstructed through sparsity-promoted algorithm with different TOF-PET time resolutions, where configuration B with range of angle being 60 degree is used.

Time resolution (ps)	100	700	1300	1900	2500
RMSE	1.79×10^{-4}	0.0763	0.0527	0.1832	0.2584
1-SSIM	4.93×10^{-6}	6.84×10^{-4}	3.58×10^{-4}	0.0082	0.0164

between the number of detectors and time resolution, in particular: (a) the more the detectors, the better the reconstruction; and (b) the higher the TOF time resolution, the better the reconstruction. We can choose suitable critical values for RMSE and 1-SSIM as a threshold according to the requirements of reconstruction quality; consequently, we can see from Figure 10 that for each t there is a critical value N_c above which the reconstructed TOF-PET images can be considered acceptable. Appealingly, it is desirable to get the explicit dependence of N_c on t for given threshold values of RMSE or 1-SSIM. For instance, if the threshold of RMSE is chosen as 0.1, after carrying out the standard least square method we can derive $N_c(t)$ as

$$N_c(t) = 108 - 48 \exp\left(-0.016 t^{1.5}\right). \tag{21}$$

For convenience, we would like to refer this curve represented by Equation (21) as a transition phase curve, which has been superposed in Figure 10(a) by a red line. In addition, this curve is very close to that obtained by a similar analysis on 1-SSIM with the threshold of 0.004, as shown by red line in Figure 10(b).

3.4. Numerical Test 4: Investigation on Configuration B

With almost the same computational setup as those used in numerical test 3, we conducted a parallel numerical investigation on configuration B to explore the relationship between the reconstruction performance represented by RMSE and 1-SSIM, TOF time resolution t, and the range of angle φ. Different from previous subsection, γ_1, γ_2 and γ_3 are set to be 1×10^{-3}, 1×10^{-4}, and 1×10^{-8}, respectively.

Figure 11 shows the TOF-PET reconstructed images through sparsity-promoted algorithm with the range of angle fixed at $60°$ and their corresponding horizontal profiles for TOF time resolutions of 100 ps, 700 ps, 1300 ps, 1900 ps and 2500 ps. Their associated RMSEs and (1-SSIM)s are presented in Table 5. From these figures and tables, we can obtain the following two conclusions:
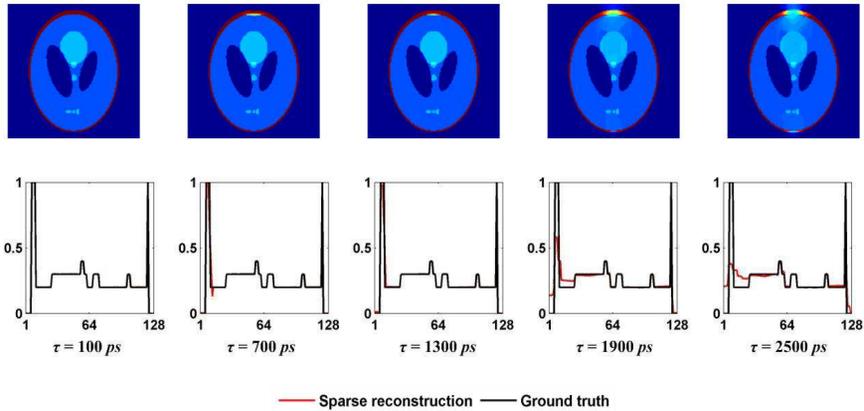
Figure 11. TOF-PET images and their corresponding vertical profiles for different time resolutions, under configuration B with range of angle being $60°$.

(a) for a fixed range of angle, the reconstruction quality goes down as the time resolution increases, similar as one drawn through numerical test 3.

(b) the reconstructions close to the upper and bottom edges of phantom are distorted, mainly resulting from deficient rays traveling through these two regions.

The relations between RMSE and 1-SSIM of reconstructed images, the range of angle φ and the TOF resolution t have been illustrated in Figures 12(a) and 12(b), respectively. Here, the TOF time resolution t ranges from 100 ps to 3000 ps, with interval of 100 ps, while the range of angle varies between $40°$ and $110°$ with interval of $10°$. In Figure 12, the x-axis denotes the time resolution t in ps while the range of angle φ is for y-axis. From Figure 12, we see that for each t there is a critical value φ_c above which the reconstructed TOF-PET images can be acceptable. Similarly, if the threshold of RMSE is specified as 0.05, the phase transition curve of $\varphi_c(t)$ can be fitted into

$$\varphi_c(t) = -510 - 3284a \tan(9t + 300). \tag{22}$$

This curve is close to that fitted for 1-SSIM with a threshold of 0.0001. These curves have been superposed in Figures 12(a) and 12(b) by a solid red line, respectively.

Finally, we would like to emphasize that besides the dependence on specific choice of threshold involved in Equations (21) and (22) to meet
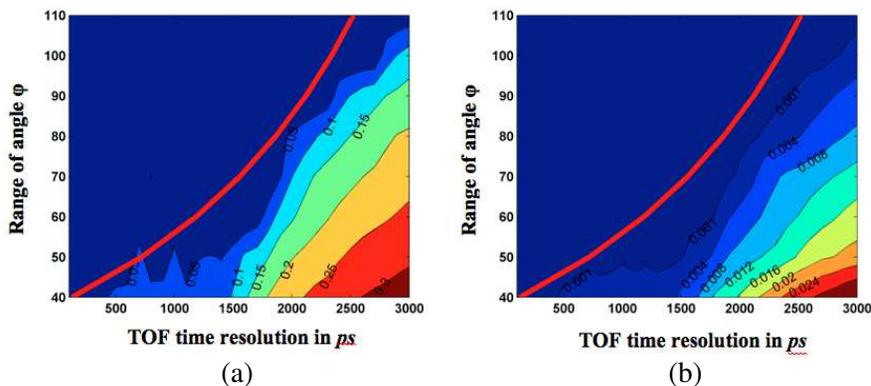
Figure 12. Contours of (a) RMSE and (b) 1-SSIM of TOF-PET reconstructed images as a function of the range of angle φ, and the TOF resolution t, for configuration B. The red solid lines represent the phase transition curves with a RMSE threshold of 0.05.

the specific requirement on imaging quality, these phase transition curves are also highly dependent on the structural complexity (or information content) of the phantom used. Using the CS terminology, we would like to use the "sparsity degree" to measure it. Therefore, it is appealing and instructive to carry out the further analysis of $N_c(t)$ and $\varphi_c(t)$ on more phantoms. We will leave it for future work.

4. CONCLUSIONS

Owing to the consideration of reducing hardware cost and improving reconstruction quality, this paper first studied the benefits of sparsity-promoted reconstruction for two different undersampling TOF-PET configurations. Two major contributions of this paper are as follows: (1) we developed an efficient algorithm for the combined sparse reconstruction of p-TV and l_1-norm to deal with TOF-PET imaging, which renders us an important conclusion that 0.5-TV regularization is optimal for our problem; (2) we built the relationship between the number of detectors (or the range of angle) and TOF time resolution, which provides an empirical guide of designing a novel low-cost TOF-PET system while ensuring acceptable reconstruction. Several representative numerical experiments have been provided to validate the proposed methodology, which demonstrates the promising applicability of undersampling TOF-PET imaging.

REFERENCES

1. Jamieson, D. G. and J. H. Greenberg, "Positron emission tomography of the brain," *Computerized Medical Imaging and Graphics*, Vol. 13, No. 1, 61–79, 1989.

2. Ollinger, J. M. and J. A. Fessler, "Positron-emission tomography," *IEEE Signal Processing Magazine*, Vol. 14, No. 1, 43–55, 1997.

3. Conti, M., et al., "First experimental results of time-of-flight reconstruction on an LSO PET scanner," *Physics in Medicine and Biology*, Vol. 50, 4507–4526, 2005.

4. Muehllehner, G. and J. S. Karp, "Positron emission tomography," *Physics in Medicine and Biology*, Vol. 51, R117–R137, 2006.

5. Surti, S. and J. S. Karp, "Design considerations for a limited angle, dedicated breast, TOF PET scanner," *Physics in Medicine and Biology*, Vol. 53, 2911–2921, 2008.

6. Mallat, S., *A Wavelet Tour of Signal Processing*, Academic-Press, 1998.

7. Aharon, M., M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Prcoessing*, Vol. 54, 4311, Nov. 2006.

8. Lee, K., S. Tak, and J. Ye, "A data-driven sparse GLM for fMRI analysis using sparse dictionary learning with MDL criterion," *IEEE Trans. Medical Imaging*, Vol. 30, 1076–1089, May 2011.

9. Ravishankar, S. and Y. Bresler, "MR image reconsruction from highly undersampled K-space data by dictionary learning," *IEEE Trans. Medical Imaging*, Vol. 30, 1028–1041, 2011.

10. Bouman, C. and K. Sauer, "A generalized Gaussian image model for edge-perserving map estimation," *IEEE Trans. Signal Processing*, Vol. 2, 296–310, Jul. 1993.

11. Rudin, L., S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physics D*, Vol. 60, 259–268, Jul. 1992.

12. Unser, M. and P. Tafti, "Stochastic models for sparse and piecewise-smooth processing," *IEEE Trans. Signal Processing*, Vol. 59, 989–1006, Mar. 2011.

13. Karahanoglu, F., I. Bayram, and D. van de Ville, "A signal processing approach to generalized 1-D total variation," *IEEE Trans. Signal Processing*, Vol. 59, 5265–5274, Nov. 2011.

14. Rodriguez, P. and B. Wohlberg, "Efficient minimization method for a generalized total variation functional," *IEEE Trans. Image Processing*, Vol. 18, 322–332, Feb. 2009.

15. Candes, E., J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Commun. Pure Appl. Math.*, Vol. 59, 1027–1223, 2006.

16. Candes, E., J. Romberg, and T. Tao, "Robust uncertainty principle: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, Vol. 52, 489–509, Feb. 2006.

17. Lustig, M., D. Donoho, and J. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magnetic Reson. Med.*, Vol. 58, 1182–1195, Apr. 2007.

18. Bian, J., J. Siewedsen, X. Han, et al., "Evaluation of sparse-view reconstruction from flat-panel-detector cone-beam CT," *Phys. Med. Biol.*, Vol. 55, 6575, 2010.

19. Han, X., J. Bian, D. Eaker, et al., "Algorithm-enabled low-dose micro-CT imaging," *IEEE Trans. Medical Imaging*, Vol. 30, 606–620, Mar. 2011.

20. Harmany, Z. T., R. F. Marcia, and R. M. Willett, "Sparsity-regularized photon-limited imaging," *IEEE International Symposium on Biomedical Imaging from Nano to Macro*, 772–775, 2010.

21. Wang, G. and J. Qi, "Direct reconstruction of dynamic PET parameteric images using sparse spectral representation," *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, 867–870, 2009.

22. Ahthoine, S., J. F. Aujol, Y. Broursier, and C. Melot, "On the efficiency of proximal methods for CBCT and PET reconstruction with sparsity constraint," *4th Workshop on Signal Processing with Adaptive Sparse Structured Representations*, 25, 2011.

23. Meinshausen, N. and B. Yu, "LASSO-type recovery of sparse representations for high-dimensional data," *Annals of Statistics*, Vol. 37, 246–270, 2009.

24. Zhu, C., "Stable recovery of sparse signals via regularized minimization," *IEEE Trans. Information Theory*, Vol. 54, 3364–3367, Jul. 2008.

25. Mallon, A. and P. Grangeat, "Three-dimensional PET reconstruction with time-of-flight measurement," *Phys. Med. Biol.*, Vol. 37, 717–729, 1992.

26. Cho, S., S. Ahn, Q. Li, and R. Leahy, "Analytical properties of time-of-flight PET data," *Phys. Medi. Biol.*, Vol. 53, 2809–2821, 2008.

27. Beck, A. and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imaging Sciences*,

Vol. 2, No. 1, 183–202, 2009.

28. Richter, S. and R. De Carlo, "Continuation methods: Theory and applications," *IEEE Transactions on Automatic Control*, Vol. 28, No. 6, 660–665, 1983.

29. Wang, Z. and A. C. Bovik, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Processing*, Vol. 13, No. 4, 1–14, 2004.

30. Zhdanov, M. and E. Tolstaya, "Minimum support nonlinear parameterization in the solution of a 3D magnetotelluric inverse problem," *Inverse Problems*, Vol. 20, 937–952, 2004.

31. Lois, C., et al., "An assessment of the impact of incorporating time-of-flight information into clinical PET/CT imaging," *The Journal of Nuclear Medicine*, Vol. 51, No. 2, 237–245, 2010.

32. Daubechies, I., et al., "Iteratively re-weighted least squares minimization for sparse reconvery," *Communications of Pure and Applied Mathematics*, Vol. 63, No. 1, 1–38, 2010.

33. Vogel, C. R., "Non-convergence of the L-curve regularization parameter selection method," *Inverse Problems*, Vol. 12, 535–547, 1996.

34. Golub, G. H., M. Heath, and G. Wahba, "Generalized cross-validation as a method for choosing a good ridge parameter," *Technometrics*, Vol. 21, No. 2, 215–223, 1979.

35. Li, K.-C., "From STEIN's unbiased risk estimates to the method of generalized cross validation," *The Annals of Statistics*, Vol. 13, No. 4, 1362–1377, 1985.