A Higher-Order Polynomial Method for SPECT Reconstruction

Ying Jiang, Si Li, and Yuesheng Xu

***Abstract*— Existing single-photon emission computed tomography (SPECT) reconstruction methods are mostly based on discrete models that may be viewed as piecewise constant approximations of a continuous data acquisition process. Due to low accuracy order of piecewise constant approximations, a traditional discrete model introduces irreducible model errors which are a bottleneck of the quality improvement of reconstructed images in clinical applications. To overcome this drawback, we develop a higher-order polynomial method for SPECT reconstruction. Specifically, we represent the data acquisition of SPECT imaging by using an integral equation model, approximate the solution of the underlying integral equation by higher-order piecewise polynomials leading to a *new* discrete system and introduce two novel regularizers for the system, by exploring the *a priori* knowledge of the radiotracer distribution, suitable for the approximation. The proposed higher-order polynomial method outperforms significantly the cutting edge reconstruction method based on a traditional discrete model in terms of model error reduction, noise suppression, and artifact reduction. In particular, the coefficient of variation of images reconstructed by the piecewise linear polynomial method is reduced by a factor of 10 in comparison to that of a traditional discrete model-based method.**

***Index Terms*— Nuclear imaging, image reconstruction, noise and artifact reduction.**

I. Introduction

SINGLE-PHOTON emission computed tomography (SPECT) provides diagnostic information via estimates of radiotracer distribution through tomographic reconstruction of projection data. It has contributed enormous benefits to human health over the past decades. However, SPECT involves a fundamental trade-off between image quality and radiation dose to the patient. Increased radiation exposure carries a probability of health detriment to the patients, especially the increased risk of getting cancer later in life [1]. Hence, there is a real need to reduce the radiation dose in SPECT studies. Lower radiation dosage can be achieved by lowering the activity of administered radiotracer. It will inevitably lead to fewer gamma photons in the projection data and thus to increased noise in the reconstructed images. Such noisy images might not be clinically useful. Therefore, the challenge is to reduce the radiation dose in SPECT examinations without compromising the image quality. This can be accomplished by application of better radiotracers, improved imaging hardware, and/or superior reconstruction methods. In this work, we focus on image quality improvement in the low-dose SPECT studies by developing a higher-order polynomial reconstruction method.

The proposed higher-order polynomial image reconstruction method consists of four essential components. An integral equation model accurately describing the data acquisition process is the basis of the method. The ideal data acquisition is modeled as Radon transform [2]. However, real imaging is never ideal. An accurate model should take into account the degradations that distort the projection data. The effects of attenuation and spatial resolution are two major physical sources of image degradation. For this reason, we formulate a relatively comprehensive integral equation model for SPECT imaging by incorporating these two effects into the Radon transform using the Lambert-Beers law [3] and resolution modeling approaches [4].

The second component is approximation of the integral equation. Traditional reconstruction methods widely use discrete models that are piecewise constant approximations of the underlying integral equation model. They are consistent with usual sampling methods and convenient to implement. However, due to the low approximation accuracy order of piecewise constant approximations, the traditional discrete models impose bottleneck model errors that cannot be compensated by a reconstruction method, since such model errors do not have any prior knowledge of their distribution. The model errors will lead to a loss of accuracy in the reconstructed images. Moreover, the traditional discrete models may not be suitable for reconstruction from very low dose projection data. Very low radiation dosage leads to severe noise in the projection data, and such badly-contaminated data would in

Manuscript received September 22, 2018; revised November 10, 2018; accepted November 12, 2018. Date of publication November 23, 2018; date of current version May 1, 2019. This work was supported in part by the Special Project on High-performance Computing through the National Key R&D Program under Grant 2016YFB0200602, in part by the Natural Science Foundation of China under Grant 11771464, Grant 11601537, Grant 11571383, Grant 11471013, and Grant 61371126, in part by the U.S. National Science Foundation under Grant DMS-1522332, and in part by the Science and Technology Program of Guangzhou, China, under Grant 201804020053. *(Ying Jiang and Si Li are co-first authors.) (Corresponding author: Yuesheng Xu.)*

Y. Jiang is with the Guangdong Province Key Lab of Computational Science, School of Data and Computer Science, Sun Yat-sen University, Guangzhou 510275, China.

S. Li is with the School of Computer Science and Technology, Guangdong University of Technology, Guangzhou 510006, China (e-mail: reesiloveu@163.com).

Y. Xu is with the Department of Mathematics and Statistics, Old Dominion University, Norfolk, VA 23529 USA, and also with the Guangdong Province Key Lab of Computational Science, School of Data and Computer Science, Sun Yat-sen University, Guangzhou 510275, China (e-mail: y1xu@odu.edu; xuyuesh@mail.sysu.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMI.2018.2881919

turn lead to reconstructed images with very high noise and very low lesion detectability if reconstructed by a discrete model of low accuracy. We propose to adopt higher-order piecewise polynomials [5]–[10] to approximate the integral equation, to gain approximate solutions having a higher-order accuracy. The reconstruction method based on this approximation, referred to as the higher-order polynomial method throughout this work, can better recover the fluctuation pattern of the radioactivity distribution, and has great potential for reconstruction from very low dose projection data.

The third component is regularization of the resulting discrete system. Since the model is formulated as a Fredholm integral equation of the first kind with a compact integral operator, the system obtained from the approximation of the integral equation is ill-posed. Therefore, a proper regularization is required. In the context of piecewise constant approximation, different types of regularization functions have been proposed to address issues with various considerations, such as suppression of image noise and artifacts, preservation of image spatial resolution, and enhancement of lesion contrast. The total variation (TV) regularizer, first proposed by Rudin *et al.* [11] and introduced to the field of ECT reconstruction by Jonsson *et al.* [12] and Panin *et al.* [13], is particularly interesting because it preserves the high frequency components of the reconstructed radiopharmaceutical distribution, including large jumps and steep transitions. However, images resulting from the application of TV in the presence of noise often suffer from staircase artifacts (*i.e.*, appear as a collection of piecewise constant regions and are particularly severe for smooth transitions of activity). This issue may limit the clinical use of TV. By exploring the *a priori* knowledge of the radiotracer distribution, we design two novel regularizers suitable for piecewise polynomial solutions of the integral equation, with an aim of suppressing their high frequency oscillation and staircase artifacts. The two regularizers are formulated in terms of the ℓ^1-norm to impose sparsity of the approximate solution in an appropriate transform domain.

As the fourth component, a fast solver for the resulting non-smooth convex optimization problem is much needed. The optimization problem that results from the regularization involves a non-differentiable ℓ^1-norm and it is of large size. Standard numerical methods developed for solving smooth optimization problems do not generally apply to this problem. Solving such an optimization problem requires great care. We construct an efficient iteration scheme based on the fixed-point characterization [14]–[18] of the solution of this problem, with a preconditioning technique motivated by the classical EM algorithm to accelerate convergence of the proposed scheme.

We demonstrate that the higher-order polynomial method described above can be efficiently used for accurate SPECT reconstruction by providing numerical results. We perform numerical experiments on two types of phantoms: piecewise constant and smooth, because they contain the essential characteristics of activity distributions frequently encountered in SPECT imaging. We compare the proposed method with a cutting edge discrete method (the total variation regularization method based on a traditional discrete model) in terms of model error reduction, noise suppression and artifact reduction.

II. Method

A. Integral Equation Model

We first describe an integral equation model for SPECT imaging. SPECT imaging involves discrete measurements that correspond to certain integral transformation of a function of continuous variables. The associated data acquisition process can be naturally formulated as an integral equation. In particular, we denote by f the radioactivity distribution on the square image domain $\Omega \subset \mathbb{R}^2$ and by $\overline{g}$ a continuous distribution function on the projection domain $U \times \Theta$ with $U \subset \mathbb{R}$ being a closed interval on the detector face and $\Theta := [0, 2\pi]$ being the rotational range. The measurements acquired from a SPECT imaging system are discrete values of $\overline{g}$. The projection distribution function $\overline{g}$ relates to f through the system point response function h (to be given later) via the following integral equation:

$$\int_\Omega f(\mathbf{x}) h(u, \theta, \mathbf{x}) d\mathbf{x} = \overline{g}(u, \theta), \quad (u, \theta) \in U \times \Theta. \tag{1}$$

We now discuss the formulation of the system point response function h. We include the effects of photon attenuation and system spatial resolution in the integration kernel h and in later computer simulation, for they are two major physical sources of image degradation in SPECT imaging. For $\mathbf{x} := (x, y)$ and $\theta \in \Theta$, we let

$$u' := x \cos\theta + y \sin\theta \quad \text{and} \quad v := -x \sin\theta + y \cos\theta. \tag{2}$$

The attenuation factor with μ being the linear attenuation coefficient can be described by

$$R(u', \theta, \mathbf{x}) := \exp\left\{ - \int_{l(u', \theta, \mathbf{x})} \mu(\mathbf{x}') ds(\mathbf{x}') \right\},$$

where $l(u', \theta, \mathbf{x})$ denotes the attenuation path, and the Gaussian-type system point spread function with distance-dependent standard deviation σ_S can be represented by

$$K(u, u', v) := \exp\left\{ - \frac{(u - u')^2}{2\sigma_S^2(v)} \right\}.$$

The attenuation path $l(u', \theta, \mathbf{x})$ is a line segment from the point source $\mathbf{x} := (x, y)$ to the detector face, and it follows the direction $(-\sin\theta, \cos\theta)$. Thus, the integration kernel h has the form

$$h(u, \theta, \mathbf{x}) = R(u', \theta, \mathbf{x}) K\left(u, u', v\right). \tag{3}$$

We remark that the rotating $u'v$-coordinate system relates to the xy-coordinate system via the transform defined by (2).

B. Representation of Radioactivity Distribution

At the heart of SPECT reconstruction is a full discretization of model (1) whose right hand side is sampled at discrete points corresponding to measurements of a SPECT imaging system, which comes down to an approximation of the radioactivity distribution f.

The measurements in SPECT imaging are essentially realizations of certain random variables, and their expectation values are obtained by sampling the continuous function $\overline{g}$. Specifically, we assume that the imaging system consists of r evenly spaced projection views in a p-dimensional detector array. For $i := i_1 + (i_2 - 1)p$, with $i_1 = 1, 2, \ldots, p$, $i_2 = 1, 2, \ldots, r$, the measurement g_i denotes the number of gamma photons recorded by detector bin U_{i_1} at projection view θ_{i_2}. The corresponding expectation value is then given by the sampling $\overline{g}_i := \int_{U_{i_1}} \overline{g}(u, \theta_{i_2})du$.

In this work, we employ a higher-order piecewise polynomial to approximate f. Let $\mathbb{X}$ denote a finite dimensional piecewise polynomial subspace of $L^\infty(\Omega)$. Equation (1) is solved approximately by seeking $\widetilde{f} \in \mathbb{X}$ such that

$$\int_\Omega \widetilde{f}(\mathbf{x}) h_i(\mathbf{x}) d\mathbf{x} = \overline{g}_i, \quad i = 1, 2, \ldots, m, \tag{4}$$

where $h_i := \int_{U_{i_1}} h(u, \theta_{i_2}, \cdot)du$ is the radioactivity contribution to the ith measurement and $m := pr$ represents the number of the discrete measurements. When a basis of $\mathbb{X}$ is chosen, (4) becomes a fully discrete linear system. In particular, if one predefines a rectangular mesh of the image domain Ω and chooses a piecewise constant approximation for image representation [15], [17]–[20], (4) reduces to a traditional discrete model.

We next describe the bases for $\mathbb{X}$. We assume the dimension of the detector array is p and generate a uniform rectangular mesh for Ω, which consists of $p \times p$ square grids of equal areas (each of area h^2). Starting from the lowest left one, we index the matrix of grids columnwise. Indeed, the matrix of grids can be organized as a one-dimensional array in such a way that the (j_1, j_2)-th grid, with $j_1, j_2 = 1, 2, \ldots, p$, corresponds to the j-th element of the array with $j := j_1 + (j_2 - 1)p$. Given a specific grid ω_j, we choose the basis functions supported in such a grid as the tensor product polynomials $\varphi_{j,k}(x, y) := (x - x_j)^{k_1}(y - y_j)^{k_2}/h^{k_1+k_2}$, $(x, y) \in \omega_j$, where (x_j, y_j) denotes the coordinates of the lower left vertex of ω_j, and $k := k_1 + 1 + k_2 n$ with $k_1, k_2 = 0, 1, \ldots, n - 1$ and $n \in \mathbb{N}$. Outside ω_j, $\varphi_{j,k}$ are all zero. In Fig. 1, we show the graph of 4 piecewise linear basis functions $\varphi_{j,1}(x, y) = 1$, $\varphi_{j,2}(x, y) = x$, $\varphi_{j,3}(x, y) = y$, and $\varphi_{j,4}(x, y) = xy$ on the support $\omega_j := [0, 1] \times [0, 1]$. We choose $\mathbb{X} := \text{span}\{\varphi_{j,k} : j = 1, 2, \ldots, p^2; k = 1, 2, \ldots, n^2\}$ as the solution space, and thus have $d := \dim(\mathbb{X}) = n^2 p^2$. Clearly, $\mathbb{X}$ is a piecewise polynomial space.

We remark that the higher-order piecewise polynomial basis can better represent the basic elements, *e.g.* a triangle shape, as compared to the traditional piecewise constant basis (Fig. 2). In Fig. 2 (a), we show a triangle with vertices $(-10, -5)$, $(10, -5)$ and $(2, 7)$. We represented this triangle by the piecewise constant basis on a 80×80 mesh where the size of each grid is 0.5×0.5 [Fig. 2 (b)], and by the piecewise linear basis on the same 80×80 mesh [Fig. 2 (c)] and on a 40×40 *coarse* mesh where the size of each grid is 1×1 [Fig. 2 (d)]. Fig. 2 (b) shows that the piecewise constant basis can only recover the edge AB which is parallel to the x-axis, and the edges BC and AC are approximated by zigzag lines. Fig. 2 (c) shows that the piecewise linear basis, defined on the same mesh, can recover the edge AB, and the approximations of the edges BC and AC are better than the zigzag lines in Fig. 2 (b). Indeed, the piecewise linear basis model can better suppress the serration on the edges than the piecewise constant basis model. Thus, visually, the edge AC in Fig. 2 (c) is recovered more like a straight line, as compared to the zigzag line in Fig. 2 (b). The number of sawteeth on the edge BC is reduced by using the piecewise linear basis model. Moreover, the edge AC in Fig. 2 (d), which is recovered by using the piecewise linear basis on the coarser mesh, still better approximates the straight line than the zigzag line in Fig. 2 (b).

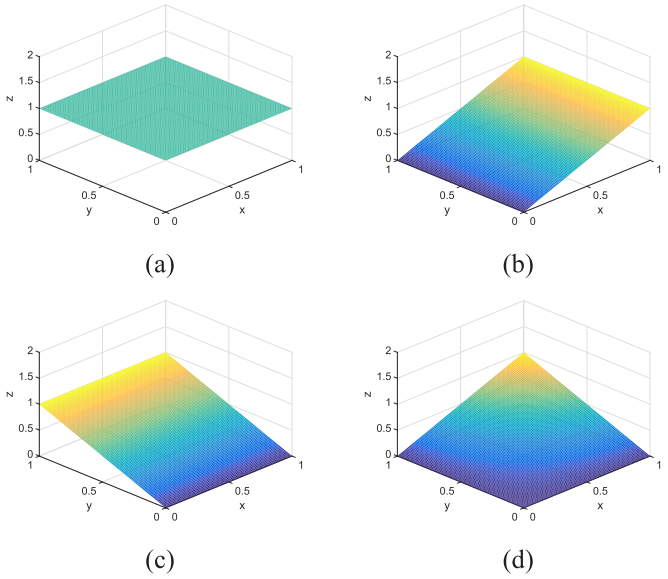
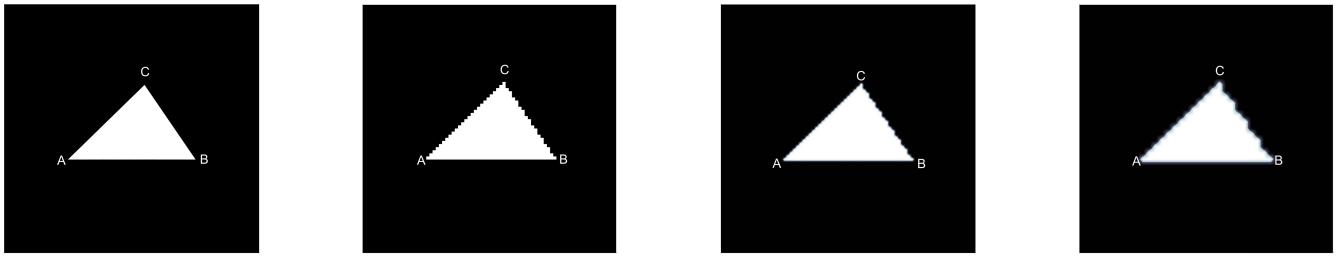


Fig. 1. Piecewise linear basis functions on the support $\omega_j := [0, 1] \times [0, 1]$: (a) $\varphi_{j,1}(x, y) = 1$, (b) $\varphi_{j,2}(x, y) = x$, (c) $\varphi_{j,3}(x, y) = y$, and (d) $\varphi_{j,4}(x, y) = xy$.

We prefer the piecewise polynomial space to other smooth function spaces, *e.g.* the spaces spanned by the blob bases [21]–[24], for two reasons. First, the system matrix after full discretization is less sparse if blobs are employed, because they are less localized than the piecewise polynomial bases and therefore contribute to more detector bins [22]. It requires more computational effort to implement the projection and back-projection operations during the iterative reconstruction using blobs. We remark that the rotational symmetry of blobs may not simplify the calculation of projections and back-projections in the case of model (1), where the depth-dependent attenuation and distance-dependent spatial resolution are taken into account. Second, blobs cannot represent a piecewise constant image accurately [23]. Indeed, the blob approximation consists of oscillations about the correct constant values, and these approximation errors have an adverse influence on the detectability of small lesions. Moreover, in order to suppress the approximation errors, as well as to meet the demand of required reconstruction characteristics, one needs to perform time-consuming pre-reconstruction experiments to determine the optimum blob parameters. Furthermore, piecewise polynomials are natural



(a) (b) (c) (d)

Fig. 2. (a) Benchmark triangle; (b) piecewise constant representation on the 80 × 80 mesh; (c) piecewise linear representation on the 80 × 80 mesh; (d) piecewise linear representation on the 40 × 40 mesh.

extensions of piecewise constants which lead to existing discrete models. In future studies, we will conduct comprehensive simulation comparison between the above competing basis models.

Next, we introduce an $m \times d$ SPECT system matrix

$$A := \begin{bmatrix} A_1 & A_2 & \cdots & A_{n^2} \end{bmatrix}, \tag{5}$$

where A_k is the $m \times p^2$ submatrix $[a_{ijk} : i = 1, 2, \ldots, m;\ j = 1, 2, \ldots, p^2]$, with $a_{ijk} := \int_{\omega_j} h_i(\mathbf{x})\varphi_{j,k}(\mathbf{x})d\mathbf{x}$. Defining two column vectors $\overline{\mathbf{g}} := \left[\overline{g}_1, \overline{g}_2, \ldots, \overline{g}_m\right]^\top$ and $\mathbf{f} := \left[\mathbf{f}_1^\top\ \mathbf{f}_2^\top\ \cdots\ \mathbf{f}_{n^2}^\top\right]^\top$ with $\mathbf{f}_k := \left[f_{1,k}, f_{2,k}, \ldots, f_{p^2,k}\right]^\top$, we write equation (4) as the linear system:

$$A\mathbf{f} = \overline{\mathbf{g}}. \tag{6}$$

Upon solving system (6) for vector $\mathbf{f}$, we obtain the solution $\widetilde{f}$ of equation (4) by calculating the linear combination $\widetilde{f} := \sum_{k=1}^{n^2}\sum_{j=1}^{p^2} f_{j,k}\varphi_{j,k}$.

C. Penalized Maximum Likelihood Estimation

In a SPECT system, both the gamma photon emission rate and the subsequent photon detection follow the temporal Poisson distribution. Therefore, the measurements $\mathbf{g} \in \mathbb{R}^m$ at m detector elements, which relate to the representation coefficients $\mathbf{f} \in \mathbb{R}^d$ through their expectations $\overline{\mathbf{g}}$ and the linear system (6), can be approximated by the Poisson model:

$$\mathbf{g} = \text{Poisson}\left(\overline{\mathbf{g}}\right) = \text{Poisson}\left(A\mathbf{f}\right). \tag{7}$$

In equation (7), Poisson(α) denotes a Poisson distributed random vector with mean α, and A is the $m \times d$ SPECT system matrix defined in (5).

Given a realization of $\mathbf{g}$, the goal of SPECT reconstruction is to estimate $\mathbf{f}$ so that the resulting approximation $\widetilde{f}$ preserves image features as much as possible, while suppressing Poisson noise. We formulate the reconstruction problem via the penalized maximum likelihood (ML) criterion. The penalized ML estimate is obtained by maximizing the sum of the log-likelihood function of $\mathbf{f}$ and the negative penalty term. Since the underlying likelihood function is assumed to be in a Poisson form described above, the proposed optimization model for SPECT reconstruction reads:

$$\mathbf{f}_\star = \arg\min\left\{\langle A\mathbf{f}, \mathbf{1}\rangle - \langle \ln(A\mathbf{f}), \mathbf{g}\rangle + \lambda R(\mathbf{f})\right\}. \tag{8}$$

The data fidelity term $F := \langle A\cdot, \mathbf{1}\rangle - \langle \ln(A\cdot), \mathbf{g}\rangle$ is the Kullback-Leibler (KL) divergence. The regularization term λR is introduced to enforce desired properties of the estimate with λ being a positive penalty-weight and R being a real-valued regularizer defined on $\mathbb{R}^d$.

1) The First-Order Discontinuity Penalization: We next motivate the formulation of regularization by desired image features and propose two types of the regularizer R. Since the space $\mathbb{X}$ is spanned by piecewise-defined bases, the underlying solution $\widetilde{f}$ is not necessarily smooth. This would yield low approximation accuracy when approximating the activity distributions in SPECT, for they are usually smooth images. On this account, the underlying regularizer should first penalize the first-order discontinuities, also referred to as jumps, at all the grid boundaries. Based on this motivation common to the both regularizers, we formulate the first one in an intuitive manner, via penalizing differences between the restrictions of the solution $\widetilde{f}$ on two neighboring grids at their common boundary. Indeed, the basic form of the regularizer for tomographic reconstruction, when using square grids, reads

$$R(\mathbf{f}) := \sum_{j=1}^{p^2}\sum_{l\in N_j} V_{jl}(\mathbf{f}). \tag{9}$$

That is, the total energy $R(\mathbf{f})$ is the sum, over the neighborhood index set N_j of each grid ω_j for all $j \in \{1, 2, \ldots, p^2\}$, of the potential functions V_{jl} related to the grid ω_j and grids ω_l within the neighborhood of ω_j. To avoid duplication of computation, we solely consider the jump penalty on the upper and the right boundaries of each grid. Under these circumstances, the neighborhood index set of ω_j is given by $N_j := \{j + 1, j + p\}$, and we propose a formulation of the potential function as

$$\begin{aligned}\sum_{l\in N_j} V_{jl}(\mathbf{f}) = &\int_{x_j}^{x_j+h} \left|\widetilde{f}_j(x, y_j + h) - \widetilde{f}_{j+1}(x, y_j + h)\right| dx \\ &+ \int_{y_j}^{y_j+h} \left|\widetilde{f}_j(x_j + h, y) - \widetilde{f}_{j+p}(x_j + h, y)\right| dy,\end{aligned} \tag{10}$$

where $\widetilde{f}_j$ is the restriction of the underlying solution $\widetilde{f}$ on ω_j, and thus, we have that $\widetilde{f}_j = \sum_{k=1}^{n^2} f_{j,k}\varphi_{j,k}$. The first term at the right hand side of (10) calculates the integral of

the difference between the restrictions on grids ω_j and ω_{j+1} at the common boundary $\{(x, y) : x_j \le x \le x_j + h, y = y_j + h\}$. The second term returns the integral of the difference between grids ω_j and ω_{j+p} at the boundary $\{(x, y) : x = x_j + h, y_j \le y \le y_j + h\}$. The reason we prefer the L^1-norm for the difference to quadratic or higher-order functions is that it increases linearly as the difference increases. Thus, it has a selective effect on smoothing small jumps in the difference due to noise or the use of the piecewise-defined bases while retaining large jumps in the difference which are often edges (if any) of the image. The regularizer defined by (9) and (10) (regularizer I) may allow preservation of flat surfaces and edges in the estimated activity distribution. By making changes of variables in the two integrals in (10) and rearranging the integrands as polynomials of the new variable u, we may rewrite regularizer I as

$$R_{\mathrm{I}}(\mathbf{f}) = h \sum_{j=1}^{p^2} \int_0^1 \left[\left| \sum_{k_1=0}^{n-1} s_{j,k_1} u^{k_1} \right| + \left| \sum_{k_2=0}^{n-1} t_{j,k_2} u^{k_2} \right| \right] du, \quad (11)$$

where variables $s_{j,k_1} := f_{j,k_1,0} - f_{j+1,k_1,0} + \sum_{k_2=1}^{n-1} f_{j,k_1,k_2}$ and $t_{j,k_2} := f_{j,0,k_2} - f_{j+p,0,k_2} + \sum_{k_1=1}^{n-1} f_{j,k_1,k_2}$ are introduced for notational simplicity, with $f_{j,k_1,k_2} := f_{j,k_1+1+k_2 n}$. The form of regularizer R_{I} makes the underlying optimization problem (8) difficult to solve. We remark that if the regularization term can be formulated as the composition of a convex function and a linear operator, one may resort to the fixed-point proximity algorithm, proposed originally in [14], to efficiently find the solution. To this end, we propose a surrogate for R_{I}, which has the aforementioned composite structure.

By applying the triangle inequality to all the absolute value terms in Eq. (11) and computing the integrals of the powers of u, we develop the following surrogate regularizer for R_{I}:

$$R_{\mathrm{surr}}(\mathbf{f}) := h \sum_{j=1}^{p^2} \left[\sum_{k_1=0}^{n-1} \frac{1}{k_1+1} |s_{j,k_1}| + \sum_{k_2=0}^{n-1} \frac{1}{k_2+1} |t_{j,k_2}| \right]. \quad (12)$$

We show in Appendix IV that R_{surr} is indeed equivalent to R_{I} in the sense that there exist positive numbers γ, η such that

$$\gamma R_{\mathrm{I}}(\mathbf{f}) \le R_{\mathrm{surr}}(\mathbf{f}) \le \eta R_{\mathrm{I}}(\mathbf{f}) \quad \text{for any } \mathbf{f} \in \mathbb{R}^d. \quad (13)$$

The above statement therefore justifies the proposed surrogate regularizer. We may further readily verify that R_{surr} is an upper bound of R_{I}, that is, $\gamma = 1$. In the remaining part of this paper, we shall also refer to the surrogate regularizer as regularizer I.

Next, we show that the surrogate regularizer defined by (12) can be formulated as a composition function $\Phi \circ B$ with Φ being the ℓ^1-norm and B being a linear operator. See [14] for more details on the composition formulation of regularization functions. Let I_α denote the $\alpha \times \alpha$ identity matrix, and D_α denote the $\alpha \times \alpha$ backward difference matrix with the Neumann boundary condition. That is, D_α is the $\alpha \times \alpha$ matrix having the diagonal entries $0, 1, \ldots, 1$, sub-diagonal entries all -1, and other entries all zero. In terms of the Kronecker tensor product notion $\otimes$, by B_x we denote the $np^2 \times np^2$ block diagonal matrix whose n diagonal blocks are $I_p \otimes D_p$, $\frac{1}{2} I_p \otimes D_p, \ldots, \frac{1}{n} I_p \otimes D_p$. Further, we denote by B_I the $np^2 \times np^2$ block diagonal matrix whose diagonal blocks are I_{p^2}, $\frac{1}{2} I_{p^2}, \ldots, \frac{1}{n} I_{p^2}$. For each $\kappa \in \{1, 2, \ldots, n\}$, we introduce an $np^2 \times np^2$ matrix B_{y_κ}. Indeed, if we view B_{y_κ} as a matrix with $n \times n$ blocks (each block is a $p^2 \times p^2$ submatrix), its κth row is given by an array of n blocks $\left[\frac{1}{\kappa} D_p \otimes I_p, \frac{1}{\kappa} I_{p^2}, \ldots, \frac{1}{\kappa} I_{p^2}\right]$, while the other blocks are all the $p^2 \times p^2$ zero submatrix. Using the above notation, we represent a $2np^2 \times n^2 p^2$ linear regularization operator B as

$$B := \begin{bmatrix} B_x & B_I & \cdots & B_I \\ B_{y_1} & B_{y_2} & \cdots & B_{y_n} \end{bmatrix}. \quad (14)$$

We can easily verify that if $\Phi : \mathbb{R}^{2np^2} \to \mathbb{R}$ is chosen as

$$\Phi(\mathbf{z}) := \|\mathbf{z}\|_1, \quad \mathbf{z} \in \mathbb{R}^{2np^2}, \quad (15)$$

then R_{surr} in (12) is equal to the composition $\Phi \circ B$.

We remark that when choosing the piecewise constant approximation, *i.e.* $n = 1$, for image representation, R_{surr} reduces to the conventional anisotropic TV regularization term [11], [14], [15].

2) The Higher-Order Derivative Penalization: Besides cartoon features, the radioactivity distribution also contains smooth structures. The first-order discontinuity penalization tends to reconstruct such structures as staircases. To ameliorate the staircase artifacts, we propose, in addition to penalizing the first-order discontinuity of the radioactivity distribution function, to include higher-order "derivative" penalties. To this end, we construct an approximation of the higher-order "derivatives" of f, and then apply the ℓ^1 regularization to this approximation to promote sparsity of the higher-order derivatives of f. The approximation of the higher-order "derivatives" of f is expressed in terms of the inner products of f with certain *multiscale* basis functions that annihilate polynomials of an appropriate order. We now describe the construction of the multiscale basis functions. We need two scales of grids for the image domain. We first introduce a set of p^2 overlapping coarse grids. For each $j \in \{1, 2, \ldots, p^2\}$, we define the index set $S_j := \{j, j+1, j+p, j+p+1\}$ and consider the subspace $\mathbb{X}_j := \mathrm{span}\{\varphi_{l,k} : l \in S_j; k = 1, 2, \ldots, n^2\}$ of $\mathbb{X}$. We then construct a coarse scale subspace of $\mathbb{X}_j$. Specifically, for each $j \in \{1, 2, \ldots, p^2\}$, we construct the jth coarse grid by taking the union of four neighboring original fine grids: $\Omega_j := \bigcup_{l \in S_j} \omega_l$. On each Ω_j, we introduce a coarse scale space $\mathbb{Y}_j := \mathrm{span}\{\upsilon_{j,k} : k = 1, 2, \ldots, n^2\}$ with tensor product basis $\upsilon_{j,k}(x, y) := (x - x_j)^{k_1} (y - y_j)^{k_2} / (2h)^{k_1+k_2}$ with $k = (k_1 - 1)n + k_2$, if $(x, y) \in \Omega_j$ and zero elsewhere. In fact, $\upsilon_{j,k}$ is a scale of $\varphi_{j,k}$. We naturally see that $\mathbb{Y}_j$ is a subspace of $\mathbb{X}_j$ and thus we can express $\mathbb{X}_j$ as the direct sum of $\mathbb{Y}_j$ and its orthogonal complement in $\mathbb{X}_j$. Let $\mathbb{W}_j$ be the orthogonal complement of $\mathbb{Y}_j$ in $\mathbb{X}_j$. That is, $\mathbb{X}_j = \mathbb{Y}_j \oplus^{\perp} \mathbb{W}_j$. It is easy to verify that $\dim(\mathbb{Y}_j) = n^2$ and hence $\dim(\mathbb{W}_j) = \dim(\mathbb{X}_j) - \dim(\mathbb{Y}_j) = 3n^2$. An orthonormal basis $\{\psi_{j,k} : k = 1, 2, \ldots, 3n^2\}$ for $\mathbb{W}_j$ can then be constructed from the basis of $\mathbb{X}_j$ by requiring every $\psi_{j,k}$ to be orthogonal to $\mathbb{Y}_j$. A specific construction of $\{\psi_{j,k}\}$ is presented in Appendix IV. For a function g sufficiently smooth on Ω_j, the inner product $\langle \psi_{j,k}, g \rangle$ returns an approximation of the nth order directional

derivative of g in Ω_j. This can be verified by using the Taylor expansion of g [5]–[7], [9].

We formulate the regularizer R (regularizer II) as the sum of $\left|\left\langle \psi_{j,k}, \widetilde{f}\right\rangle\right|$ for $j = 1, 2, \ldots, p^2$, $k = 1, 2, \ldots, 3n^2$ so as to penalize the nth order derivatives of $\widetilde{f}$ in respective coarse grids. We remark that regularizer II can promote the sparsity of the solution under the constructed system due to the use of linearly increasing measure (that is, the absolute sum). Specifically, the proposed regularizer reads

$$R_{\mathrm{II}}(\mathbf{f}) = \sum_{k=1}^{3n^2} \sum_{j=1}^{p^2} \left|\left\langle \psi_{j,k}, \widetilde{f}\right\rangle\right|. \tag{16}$$

For convenience of algorithmic development, we re-express the regularizer (16) as a composition of a convex function and a linear operator. Noting that $\widetilde{f} = \sum_{l=1}^{p^2} \sum_{\kappa=1}^{n^2} f_{l,\kappa} \varphi_{l,\kappa}$, $R_{\mathrm{II}}(\mathbf{f})$ can be rewritten as

$$R_{\mathrm{II}}(\mathbf{f}) = \sum_{k=1}^{3n^2} \sum_{j=1}^{p^2} \left| \sum_{\kappa=1}^{n^2} \sum_{l=1}^{p^2} \left\langle \psi_{j,k}, \varphi_{l,\kappa} \right\rangle f_{l,\kappa} \right|.$$

This leads us to form a matrix B of size $3n^2\ p^2 \times n^2\ p^2$ as

$$B = \left[\left\langle \psi_{j,k}, \varphi_{l,\kappa} \right\rangle : 1 \le l, j \le p^2, 1 \le k \le 3n^2, 1 \le \kappa \le n^2\right], \tag{17}$$

and choose the convex function Φ as the ℓ^1-norm on $\mathbb{R}^{3n^2\ p^2}$. Then the regularizer (16) is equal to the composition $\Phi \circ B$, *i.e.* $R_{\mathrm{II}}(\mathbf{f}) = \|B\mathbf{f}\|_1$. We remark that the entry at the $[3n^2(k-1)+j]$-th row and $[n^2(\kappa-1)+l]$-th column in matrix B is $\left\langle \psi_{j,k}, \varphi_{l,\kappa} \right\rangle$. Since $\Omega_j \cap \omega_l = \emptyset$ for $l \notin S_j$, $\left\langle \psi_{j,k}, \varphi_{l,\kappa} \right\rangle = 0$ for $l \notin S_j$. Thus, matrix B is a sparse matrix. More details on the generation of B can be found in Appendix IV.

D. Optimization Algorithm

Solving model (8) with regularizer I or II requires efficient iteration schemes for the corresponding non-smooth optimization problem. Traditional gradient-type and expectation maximization (EM)-type algorithms fail to solve the problem efficiently. We describe below a fixed-point proximity algorithm with preconditioning to solve the problem. By characterizing the solution of the optimization problem (8) as a fixed-point of a non-expansive map defined via the proximity operator of the conjugate function Φ^* of Φ, we can use the fixed-point proximity algorithm [14]–[18] to solve the underlying non-smooth optimization problem.

We employ the following iteration scheme proposed in [18] with preconditioning to solve the optimization problem (8)

$$\begin{cases} \mathbf{b}^{k+1} = \mathrm{prox}_{\Phi^*}^{Q}\left(\mathbf{b}^k + Q^{-1}B\mathbf{f}^k\right), \\ \mathbf{f}^{k+1} = \mathbf{f}^k - P^{-1}\left(\nabla F(\mathbf{f}^k) + B^\top\left(2\mathbf{b}^{k+1} - \mathbf{b}^k\right)\right), \end{cases} \tag{18}$$

where $\mathrm{prox}_{\Phi^*}^{Q}$ denotes the proximity operator of Φ^* with respect to Q (see, [28] for its definition). Here, $Q := \mu^{-1} I_q$ and $P := \beta^{-1} S^{-1}$ with μ, β being positive parameters and S being a $d \times d$ diagonal positive-definite preconditioning matrix.

Efficient implementation of (18) requires the availability of the closed form of $\mathrm{prox}_{\Phi^*}^{Q}$. Recalling the definition of the extended proximity operator [18], we may generalize Moreau's decomposition [28, Th. 14.3 (ii)] to the following form: $\mathcal{I} = \mathrm{prox}_{\Phi^*}^{Q} + Q^{-1} \circ \mathrm{prox}_{\Phi}^{Q^{-1}} \circ Q$ for any symmetric positive-definite matrix Q. We omit the proof of the generalization here. Using this formula, it suffices to obtain a closed form of $\mathrm{prox}_{\Phi}^{Q^{-1}}$. Indeed, the proposed regularizers (12) and (16) both choose the convex function Φ as the ℓ^1-norm on respective regularized transform domains. The proximity operator of Φ can then be calculated componentwise as

$$\mathrm{prox}_{\Phi}^{Q^{-1}}(\mathbf{x}) = \left[\mathrm{prox}_{\mu^{-1}|\cdot|}(x_1), \ldots, \mathrm{prox}_{\mu^{-1}|\cdot|}(x_q)\right]^\top,$$

for $\mathbf{x} \in \mathbb{R}^q$, where

$$\mathrm{prox}_{\mu^{-1}|\cdot|}(x_i) = \max\left\{|x_i| - \mu^{-1},\ 0\right\} \mathrm{sign}(x_i)$$

is the well-known soft thresholding operator with μ^{-1} being the threshold. The closed form of $\mathrm{prox}_{\Phi^*}^{Q}$ leads to a fast algorithm since we avoid solving many inner optimization problems by using the closed form formula above. We remark that in scheme (18), $\mathbf{b}^k \in \mathbb{R}^q$ is the dual iterate defined in the regularized transform domain. It is truncated by the soft thresholding operator at each iteration to remove noise of the image in the transform domain.

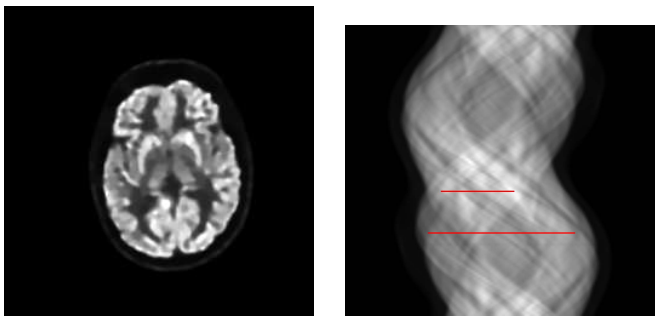
The choice of the preconditioning matrix S is crucial for the speed of convergence of the iteration (18). Motivated by the classical ML-EM algorithm [26], [27], we propose to choose S as the diagonal matrix $S^k := \mathrm{diag}(\mathbf{f}^k / A^\top \mathbf{1})$ at the kth iteration. In this way, S is updated once a new primal iterate $\mathbf{f}^k$ is available. The diagonal matrix S^k, which determines the direction of the next iteration in the classical ML-EM algorithm, helps us guide the iteration direction in the second step of (18). By employing the preconditioning technique, we accelerate significantly the convergence speed of the iteration.

III. Simulations and Numerical Results

We performed numerical simulations on piecewise constant and smooth phantoms frequently appearing in SPECT imaging. Such simple phantoms can facilitate direct comparative studies. We employed the piecewise linear polynomial (PLP) basis model for discretization of integral equation (1), and compared it with a traditional discrete model DM (*i.e.*, using the piecewise constant basis for discretization of (1)) in terms of projection accuracy, image noise performance, lesion detectability, feasibility of dose reduction and artifact visualization. In particular, we employed TV regularization in the traditional DM-based image reconstruction, because TV can preserve high frequency components of the reconstructed image, including large jumps and steep transitions, and is well suited for reconstruction of piecewise constant phantom. Throughout the entire simulation section, we applied iteration scheme (18) to solve all the regularized optimization problems.

A. Comparison of Projection Accuracy

The first validation verified model error reduction of the PLP model by using a brain phantom [Fig. 3 (a)]. The brain phantom was taken from a slice of a SPECT brain



(a) (b)

Fig. 3. Brain phantom (a) and the associated benchmark parallel-beam SPECT data (b) generated by the traditional DM on a 256 × 256 mesh. The red lines are used for line profile illustration.

scan and was subsequently denoised using the high-order TV regularization [17]. A SIEMENS E.CAM gamma camera with low energy high resolution (LEHR) parallel-beam collimator was simulated. The detector orbit was circular covering 360°, and the radius of rotation was set to 15 cm. The parallel-collimated SPECT projection data for this validation consisted of 120 projection views in a 128-dimensional detector array and were generated using three competing analytical system matrices. We set the detector bin size to 2.2 mm, then the active detector size is 28.16 cm. We also simulated the effect of uniform attenuation in the analytical projection process. In particular, we used the accurate system matrix (discretized using the traditional DM defined on a 256 × 256 rectangular mesh) to obtain the benchmark SPECT data [Fig. 3 (b)] for the brain phantom. We qualitatively and quantitatively compared the benchmark data with two other sets of analytic data, which were generated by the traditional DM on the 128 × 128 mesh and by the PLP model on the 64 × 64 mesh.

The simulation results show that the PLP model on the 64 × 64 coarse mesh outperforms the traditional DM on the 128 × 128 fine mesh. In Fig. 4 the horizontal line profiles (marked in Fig. 3 (b)) confirm that the PLP model produced a projection with better contrast and spatial resolution, as compared to that produced by the traditional DM. Moreover, we used the root mean square error (RMSE) to assess accuracy of projections. The RMSE is a global quality metric defined by

$$\text{RMSE} := \left\{ \frac{1}{N} \sum_{i=1}^{N} (g_i - g_{Oi})^2 \right\}^{1/2}, \tag{19}$$

where g_i and g_{Oi} are bin detections in the competing and benchmark projection sets, respectively, and N is the total number of detector bins in the projection space. Table I shows that the RMSE value produced by the PLP model is significantly smaller than that produced by the traditional DM.

B. Reconstruction Comparison of Piecewise Constant Image and Feasibility of Dose Reduction

The second validation assessed performance of the PLP model on reconstructing the piecewise constant phantom.

TABLE I
RMSE VALUES OF THE PLP MODEL AND THE TRADITIONAL DM

Method	RMSE
PLP model	0.056
DM	0.26

To this end, we created a digital circular emission phantom with uniform mean background activity distribution and a set of seven uniform hot spheres embedded in the circle. The hot spheres simulate hyperperfused defects, which are of interest in nuclear medicine. One of the main tasks of SPECT is detection of such defects. The mean activity ratio of hot:background areas is 4 : 1. The grid size used for this digital phantom is 1.375×10^{-1} mm. The radius of the circular phantom is 92.4 mm, and the radii of the spheres in the phantom are 3.3, 4.4, 5.5, 6.6, 7.7, 9.9 and 15.4 mm. The locations of the spheres in transaxial plane are shown in Fig. 5 (a). A SIEMENS E.CAM gamma camera with LEHR parallel-beam collimator was simulated. The detector orbit was circular covering 360°, and the radius of rotation was set to 15 cm. The parallel-collimated SPECT projection data in this validation consisted of 120 projection views in a 256-dimensional detector array with detector bin size 1.1 mm and were generated using an analytical system matrix that results from the piecewise constant discretization. The generated data were multiplied by appropriate constants to reach four count levels in 2D acquisition mode (a total of 2.8×10^5, 1.12×10^5, 0.84×10^5 and 0.56×10^5 counts in 120 views) for four different scenarios, respectively. We note that these values roughly correspond to the ACR SPECT phantom QA scans (32 million total counts in 3D acquisition mode) with 545, 218, 164 and 109 MBq of Tc-99m, respectively [29]. Based on these noise-free projection data, we used a Poisson random number generator to create 100 different noise realizations for each count level (400 data sets in total). We then employed three reconstruction methods: the traditional discrete model with TV regularizer (DM-TV) (the reconstruction method proposed in [15]) and the PLP model with regularizer I (PLP-I) and with regularizer II (PLP-II) to reconstruct the noisy projection data of the highest (2.8×10^5, denoted by $280k$) and the lowest (0.56×10^5, denoted by $56k$) count levels. Remark that in each competing reconstruction method, we simulated the effect of LEHR parallel-beam collimation in the system matrix. Moreover, to verify the feasibility of radiation-dose reduction by use of the proposed PLP model, we applied the PLP-I method to the reconstruction of all four count levels, and compared its lesion detectability with DM-TV and the clinically used EM algorithm with Gaussian post filter (DM-GPF-EM).

The competing reconstruction methods required estimation of penalty-weight λ. The optimal value of λ was obtained by performing a set of trial reconstructions with λ ranging from 10^{-4} to 1. Five observers qualitatively considered the dependence of spatial resolution, contrast and image noise on λ. Indeed, each observer independently and subjectively balanced the tradeoff among the above three metrics, and selected the best λ for each reconstruction method and each count level.

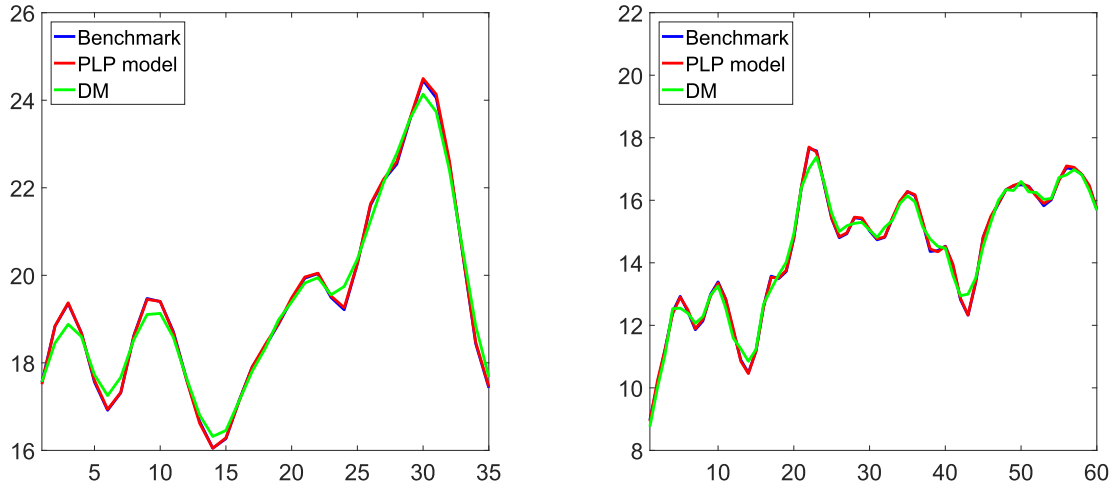


Fig. 4. Horizontal line profiles of the three investigated projection data sets: subfigure (a) corresponds to the upper red line in Fig. 3, while subfigure (b) corresponds to the bottom red line.

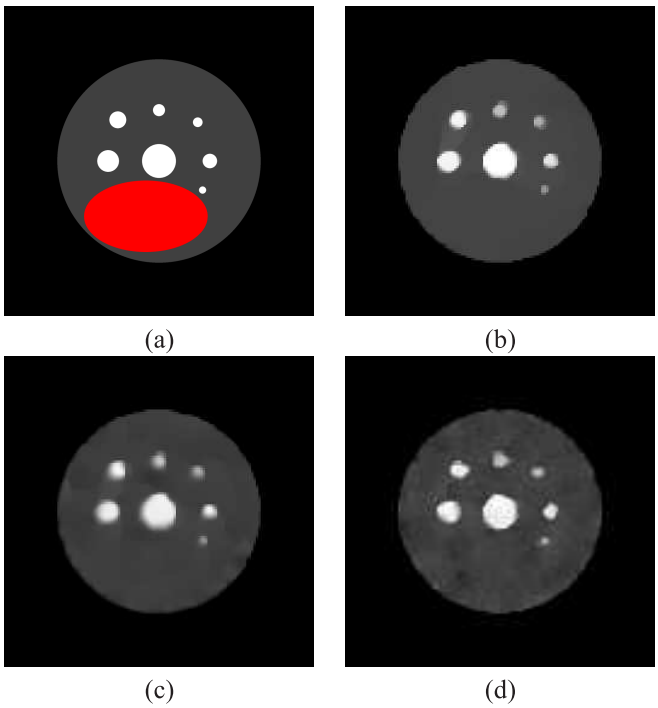


Fig. 5. Piecewise constant phantom containing seven uniform hot spheres (a) and the images reconstructed by (b) PLP-I, (c) PLP-II and (d) DM-TV from analytical noisy projection data at the 280k count level. The red ellipse represents the selected background ROI.

The penalty-weights selected by five observers were averaged and the average values were used as the optimal parameters. For construction of the 280k count level, the resulting optimal parameters for DM-TV, PLP-I and PLP-II are, respectively, $\lambda = 0.11$, 0.10 and 0.20. As for reconstruction of the 56k count level, we followed the same procedure and obtained the optimal parameters $\lambda = 0.28$, 0.18 and 0.40 for DM-TV, PLP-I and PLP-II, respectively.

We show in Fig. 5 the hot spheres reconstructed using the three competing methods with optimal penalty-weights for the same noise realization of the 280k count level. We observe much higher background noise in the images reconstructed by DM-TV, as compared to those by PLP-I or PLP-II. This is further confirmed by Table II, which presents the coefficients of variation (CoV) in a selected background region-of-interest (ROI) for the 280k and 56k count levels reconstructed by the three competing methods. The definition of CoV may be referred to [30]. We show in Fig. 5 (a) the location of the background ROI (the red ellipse), which was placed away from the edges of the hot spheres and the phantom. In the case of the 280k count level, the CoV for PLP-I is reduced by a factor of **10** in comparison to that for DM-TV. Even at the very low 56k count level, the CoV for PLP-I is still reduced by a factor of **3.7**. These numerical results show that PLP-I is more appropriate for the reconstruction of piecewise constant phantom, in comparison with PLP-II. Moreover, in terms of CoV, the images reconstructed by PLP-I and PLP-II even from the very low 56k count level outperform significantly that by DM-TV from the 280k count level.

TABLE II
COV IN THE SELECTED BACKGROUND ROI FOR THE 280k AND 56k COUNT LEVELS RECONSTRUCTED BY THE THREE COMPETING METHODS

Method	280k Counts	56k Counts
DM-TV	7.85%	9.90%
PLP-I	0.74%	2.67%
PLP-II	2.05%	6.45%

In addition, we applied the channelized Hotelling observer (CHO) [31] to quantify the detectabilities of PLP-I and PLP-II on the uniform hot spheres. Following the approaches described in [32] and [33], we find that in the case of the 280k count level, PLP-I produced much higher CHO detectability indices than PLP-II (Fig. 6). The DM-TV method performed worst in this category. As shown in Fig. 7, the CHO detectability indices obtained for the uniform hot spheres at the 1.12×10^5 (112k) count level using the PLP-I method are higher than or comparable to that obtained for the hot spheres

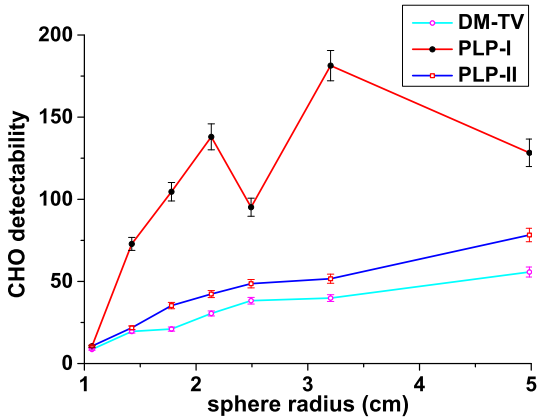


Fig. 6. CHO detectability indices of uniform hot spheres at the 280*k* count level. The solid lines connecting calculated data points were provided as visual aids.

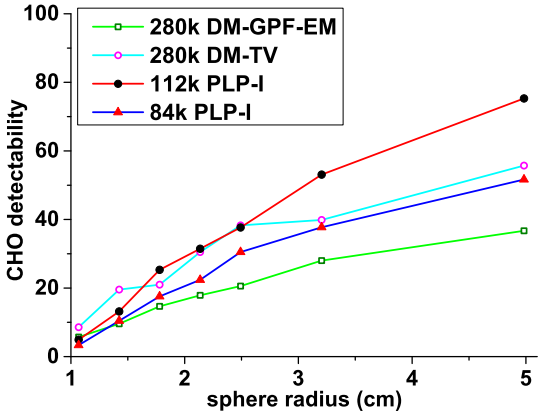


Fig. 7. CHO detectability indices of hot spheres vs. number of counts in the analytical SPECT data. The PLP-I method for 112*k* count level data: $\lambda = 0.14$; the PLP-I method for 84*k* count level data: $\lambda = 0.15$; the DM-TV method for 280*k* count level data: $\lambda = 0.11$; and the DM-GPF-EM method for 280*k* count level data: the standard deviation of the Gaussian filter is 4.53 mm. The reconstruction for DM-GPF-EM was stopped at 50 iterations, while the other reconstructions were stopped at 300 iterations. The solid lines connecting the data points were provided as visual aids only.

at the 280*k* count level using the DM-TV method. Moreover, the PLP-I reconstructed hot spheres from the projection data of the 0.84×10^5 (84*k*) count level exhibit even higher CHO detectability indices than the hot spheres reconstructed by the DM-GPF-EM method from the 280*k* count level data. The above numerical results show that even with only 30% of the number of counts used in the DM-GPF-EM reconstruction, the CHO detectability indices of hot spheres in PLP-I reconstructed images still surpassed that in DM-GPF-EM reconstructed images, indicating that a 70% radiation dose reduction might be possible. When compared with the DM-TV method, PLP-I might still allow a 60% dose reduction.

C. Reconstruction Comparison With the Traditional DM With Very High Grid Dimension

Using the hot spheres phantom, we further compare the proposed PLP model to the traditional DM with very high grid dimension. Specifically, we employed the traditional DM

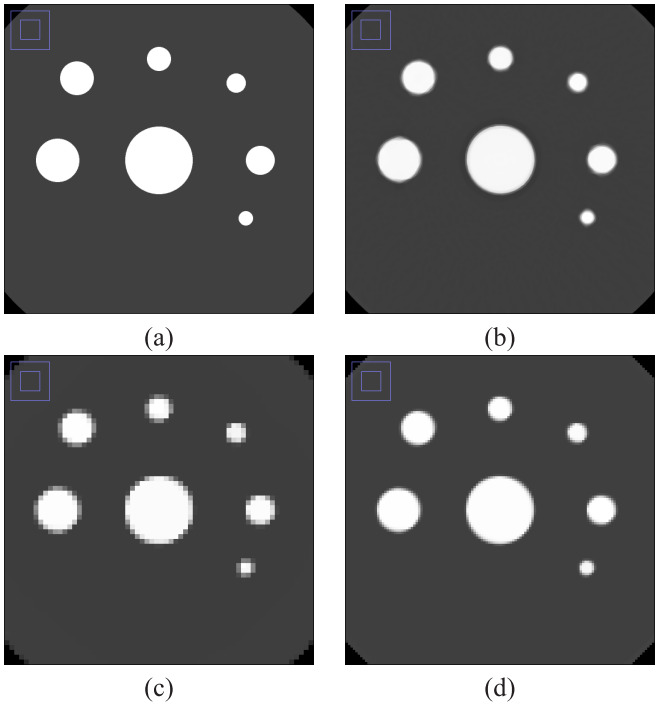


Fig. 8. Zoom-in portions of the enlarged hot spheres images. (a) Enlarged phantom; (b) enlarged reconstructed image by using the traditional DM on 2048×2048 mesh; (c) enlarged reconstructed image by using the traditional DM on 128×128 mesh; (d) enlarged reconstructed image by using the PLP model on 128×128 mesh.

on the 2048×2048 mesh to discretize the integral equation (1). The resulting system matrix is considered as an accurate projector. We next show that the reconstruction method based on such an accurate projector would generate accurate images. To avoid a potential bias caused by random perturbation and to better evaluate reconstruction accuracy, we focused on the reconstruction of noise-free projection data of the uniform spheres phantom. In this case, no regularization is needed and we applied the EM algorithm to solve the underlying optimization problems. We compared the above accurate results with two other sets of images, which were reconstructed, respectively, by the traditional DM and the PLP model on the 128×128 mesh. As we can see in Fig. 8 (b), the accurate projector can well recover the circular boundaries of the hot spheres and preserve very high spatial resolution. Fig. 8 (c) shows that the traditional DM approximates the circular boundaries by zigzag lines and suffers from poor spatial resolution. The PLP model on the same mesh [Fig. 8 (d)] significantly outperforms the traditional DM in terms of both boundary approximation and resolution preservation. We remark that although the accurate projector-based reconstruction method generates accurate images of high resolution, it is indeed very time-consuming and requires huge storage space for the system matrix involved. We show in Table III the CPU time expended and the computer memory used by the three competing models. We can see that both models on the 128×128 mesh need very little computer memory, and spend comparable computational time to run 300 iterations, which is much less than the time spent by the accurate projector.

TABLE III
COMPARISON OF THE PERFORMANCE OF THREE COMPETING MODELS FOR RECONSTRUCTION OF THE NOISE-FREE ANALYTICAL DATA OF HOT SPHERES PHANTOM. THE CPU TIME DENOTES THE TIME (UNITS: second) NEEDED TO RUN 300 ITERATIONS, AND MEM. DENOTES THE COMPUTER MEMORY (UNITS: GigaByte) REQUIRED TO STORE THE SYSTEM MATRIX

Model	CPU time	Mem.
DM on 2048×2048 mesh	4251.4	6.16
DM on 128×128 mesh	48.5	0.042
PLP model on 128×128 mesh	52.0	0.16

D. Reconstruction Comparison of Smooth Image

We conducted the third validation to assess the performance of PLP-I and PLP-II on artifact reduction. A lumpy phantom with Gaussian spheres was designed to accomplish the required assessment. A cylinder with base radius 10.4 cm and length 14.1 cm was used as the body of the phantom. The cylindrical phantom has lumpy background, obtained by superimposing randomly distributed Gaussian blobs onto a uniform background. Specifically, a lumpy distribution was modeled by Lumpy $= \sum_{k=1}^{K} \text{Gauss}(h, \sigma, p_k)$, where $\text{Gauss}(h, \sigma, p_k)$ represents a 3D Gaussian blob with standard deviation σ and maximum activity h centered at the kth random point source p_k. We added the lumpy distribution to a uniform cylinder to generate the lumpy background. We chose $K = 200$, $h = 1$ MBq/cm^3 (which is 2 times of the averaged background activity), and $\sigma = 30$ mm. A transaxial layer of six 3D Gaussian spheres (whose σ varied over the range $4 - 9$ mm) was placed within the lumpy background. The spheres, all with the same starting maximum activity, were Gaussian-blurred. The ratio of starting maximum activity to mean background is 4 : 1 [Fig. 9 (a)].

We then used the Monte Carlo simulation software package SIMIND [34] to obtain parallel-beam Tc-99m SPECT projection data for the lumpy phantom. A SIEMENS E.CAM gamma camera with LEHR collimator was simulated. The simulation consisted of 120 projection views in a 128×128-dimensional detector matrix with detector bin size 2.2^2 mm^2, and the reconstruction space grid size is 2.2^3 mm^3. During the data acquisition, the radius of rotation for the circular detector orbit was set to 15 cm. We used an 18% main energy window centered at 141 keV. The gamma photons within this energy window ($128 - 154$ keV) were considered as primary or first-order scattered photons. Moreover, we set the scatter energy window as $123 - 128$ keV. We simulated a total of 2.18×10^{10} photon histories to suppress the photon-flux fluctuation. The Monte Carlo simulated photopeak-window data were multiplied by an appropriate constant to reach a total of 8.4×10^6 counts in 120 views. Based on the rescaled noise-free simulation data, we used a Poisson random number generator to create 100 noise realizations for the photopeak-window data. Scatter compensation was implemented by adding the estimated scatter counts in forward projection at each iteration of the reconstruction. The scatter counts were estimated using the scatter-window data [35].

The total volume of the simulated phantom is approximately 4791 cm^3, whereas the volume of human brain usually does not exceed 1500 cm^3. Therefore, we expect that the above count level would correspond to 2.6×10^6 counts in 120 views for a brain with similar mean activity. This is approximately half of a typical administered dose, since the recommended total number of counts for brain perfusion SPECT is 5.0×10^6 or more [36]. Moreover, we repeated the aforementioned parameter determination procedure by multiple observers, and determined the optimal penalty-weights to be 0.5, 1.3 and 2.1 for DM-TV, PLP-I and PLP-II, respectively.

In the image quality assessment of this validation, we focused on the peak signal-to-noise ratio (PSNR) and background ensemble variance. The PSNR is a global image quality metric, which calculates the ratio between the maximum possible power of an image and the power of corrupting noise that affects the fidelity of its approximation:

$$\text{PSNR} := 20 \log_{10} \left(\frac{\text{MAX}_{f_O}}{\|f - f_O\|_2} \right). \tag{20}$$

In equation (20), f and f_O denote the reconstructed image and phantom, respectively, and MAX_{f_O} is the maximum possible pixel value of the phantom.

The ensemble variance provides a useful measure of image noise across multiple noise realizations. For the kth region-of-interest ROI_k ($k = 1, 2, \ldots, K$ and K is the total number of ROIs), the metric is defined as the variance of ROI mean activities $m_{r,k}$ over multiple independent noise realizations $r = 1, 2, \ldots, R$ [37]:

$$\sigma^2_{\text{ensemble},k} := \frac{1}{R-1} \sum_{r=1}^{R} (m_{r,k} - \overline{m}_k)^2, \tag{21}$$

where $\overline{m}_k := \frac{1}{R} \sum_{r=1}^{R} m_{r,k}$ is the average of mean activities in ROI_k over R realizations. We then averaged this metric over the K background ROIs and obtained the final background ensemble variance $\sigma^2_{\text{ensemble}} := \frac{1}{K} \sum_{k=1}^{K} \sigma^2_{\text{ensemble},k}$. In this validation, we used 100 noise realizations and a total of 60 circular ROIs centered at slices 21 to 25, where slice 23 is the central cross-section of Gaussian spheres. The radii of the ROIs were set to 10 grids.

Fig. 9 shows examples of Gaussian spheres reconstructed using the three competing methods with optimal penalty-weights for the same noise realization of the Monte Carlo simulated data. Radial line profiles passing through the centers of the 11 o'clock and 5 o'clock Gaussian spheres are shown in Fig. 10. We observe that staircase artifacts are evident in the reconstruction of DM-TV. PLP-I and PLP-II, though, greatly reduced such artifacts and preserved the image fluctuation patterns. Moreover, the PLP model based methods produced higher PSNR values and much lower background ensemble variances than the DM-TV did (Table IV).

We further show in Table V the number of complete iterations used and the CPU time (units: second) consumed by the three competing methods. The relative error we used in calculation is defined as

$$tol := \|\mathbf{f}^k - \mathbf{f}^{k+1}\|_2 / \|\mathbf{f}^{k+1}\|_2. \tag{22}$$

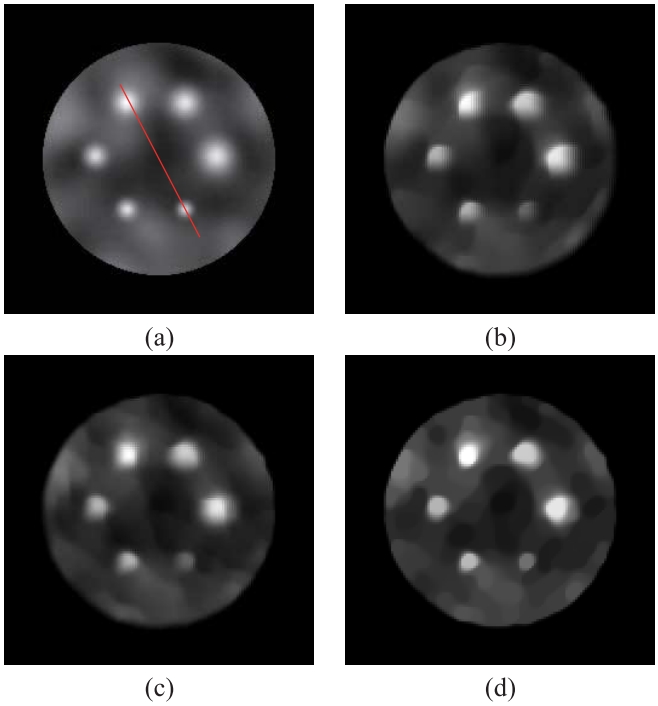


Fig. 9. (a) Lumpy phantom containing Gaussian hot spheres; and images reconstructed by (b) PLP-I, (c) PLP-II and (d) DM-TV from simulated noisy projection data. The red line is used for line profile illustration.

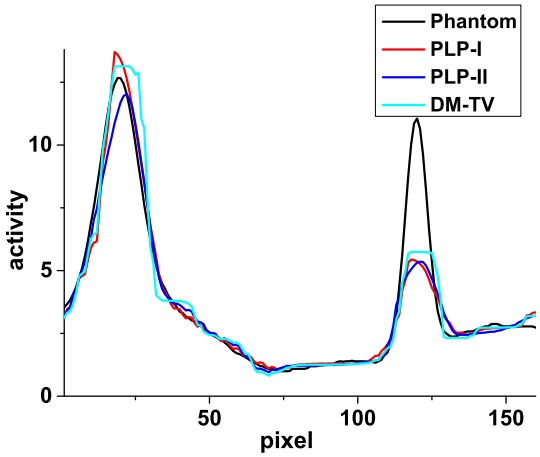


Fig. 10. Radial line profiles passing through the centers of the 11 o'clock and 5 o'clock Gaussian spheres in Fig. 9.

TABLE IV
PSNR VALUES AND BACKGROUND ENSEMBLE VARIANCES FOR LUMPY PHANTOM RECONSTRUCTED BY THE THREE COMPETING METHODS

Method	PSNR (dB)	Ensemble variance
DM-TV	29.22	0.94%
PLP-I	31.55	0.16%
PLP-II	32.20	0.18%

From Table V we conclude that, under the same stopping criteria, PLP-II performed better than DM-TV in terms of both convergence speed and computational time. When the required accuracy is higher than 10^{-2}, PLP-I also began outperforming DM-TV. The relative error *tol* vs. iteration number is shown in Fig. 11, which further confirms the above conclusions. This also validates numerically the convergence of iteration scheme (18). The rigourous theoretical convergence analysis for the iteration scheme may be referred to [18].

TABLE V
COMPARISON OF THE PERFORMANCE OF DM-TV, PLP-I AND PLP-II FOR RECONSTRUCTION OF THE MONTE CARLO SIMULATED DATA OF LUMPY PHANTOM. THE PAIR $(\cdot,\cdot)$ REPRESENTS THE NUMBER OF COMPLETE ITERATIONS AND THE CPU TIME (UNITS: SECOND) NEEDED TO ACHIEVE VARIOUS RELATIVE ERROR LEVELS

Recon. Method \ tol	10^{-1}	10^{-2}	10^{-3}
DM-TV	$(5, 18.7)$	$(46, 69.6)$	$(212, 282.8)$
PLP-I	$(5, 18.8)$	$(53, 79.8)$	$(179, 233.5)$
PLP-II	$(5, 19.3)$	$(19, 36.6)$	$(82, 117.4)$

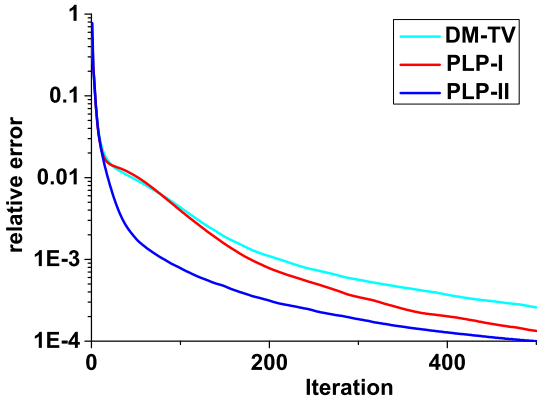


Fig. 11. Relative error vs. iteration number for the three competing reconstruction methods.

IV. CONCLUSION

We have established a novel approach for SPECT reconstruction based on the piecewise polynomial approximation discretizing the integral equation model of data acquisition and the regularization penalizing undesired discontinuity of the reconstructed image. The numerical experiments which we have conducted demonstrate that the proposed reconstruction methods outperform significantly the existing method based on a traditional discrete model in model error reduction, noise suppression, dose reduction and artifact reduction.

APPENDIX A
EQUIVALENCE BETWEEN REGULARIZER I AND ITS SURROGATE

We provide in this appendix a proof for the equivalence between the surrogate regularizer defined by (12) and regularizer I defined by (11). It suffices to show that the integral $G(\mathbf{x}) := \int_0^1 \left|\sum_{l=0}^{n-1} x_l u^l\right| du$ is equivalent to the summation $G_{\text{surr}}(\mathbf{x}) := \sum_{l=0}^{n-1} \frac{1}{l+1} |x_l|$ for any vector $\mathbf{x} := [x_0, x_1, \ldots, x_{n-1}]^\top \in \mathbb{R}^n$. We accomplish this task by proving that both G and G_{surr} are norms on $\mathbb{R}^n$. The argument (13) then follows immediately from the fact that all norms on a finite-dimensional vector space are equivalent.

We first show that G is a norm. We can readily see that G is non-negative over $\mathbb{R}^n$, and that G is definite in the sense that if $G(\mathbf{x}) = 0$, then the integrand vanishes and thus $\mathbf{x} = \mathbf{0}$ is the zero vector. Moreover, for any $c \in \mathbb{R}$ and $\mathbf{x} \in \mathbb{R}^n$, we have that

$$G(c\mathbf{x}) = \int_0^1 \left|\sum_{l=0}^{n-1} c x_l u^l\right| du = |c| G(\mathbf{x}).$$

This ensures that G is absolutely homogeneous. Next, for any $\mathbf{y} := [y_0, y_1, \ldots, y_{n-1}]^\top \in \mathbb{R}^n$, we can check that

$$\begin{aligned} G(\mathbf{x}+\mathbf{y}) &= \int_0^1 \left| \sum_{l=0}^{n-1} (x_l + y_l) u^l \right| du \\ &\le \int_0^1 \left(\left| \sum_{l=0}^{n-1} x_l u^l \right| + \left| \sum_{l=0}^{n-1} y_l u^l \right| \right) du \\ &= G(\mathbf{x}) + G(\mathbf{y}). \end{aligned}$$

In other words, G satisfies the triangle inequality. In summary, the function G is a norm on the vector space $\mathbb{R}^n$.

Following a similar proof, we can verify that G_{surr} is also a norm on $\mathbb{R}^n$. This completes the current appendix.

Appendix B
Construction of an Orthonormal Basis for $\mathbb{W}_j$

In this appendix, we present a construction of an orthonormal basis $\{\psi_{j,k} : k = 1, 2, \ldots, 3n^2\}$ of $\mathbb{W}_j$, the orthogonal complement of $\mathbb{Y}_j$ in $\mathbb{X}_j$. The orthogonality of $\mathbb{W}_j$ and $\mathbb{Y}_j$ requires that $\langle \psi_{j,k}, \upsilon_{j,\widetilde{k}} \rangle = 0$ for $\widetilde{k} = 1, 2, \ldots, n^2$. Since $\psi_{j,k}$ can be represented by the basis of $\mathbb{X}_j$, *i.e.*, $\psi_{j,k} = \sum_{l \in S_j} \sum_{\kappa=1}^{n^2} w_{l,\kappa}^{(j,k)} \varphi_{l,\kappa}$, the underlying construction amounts to solving the following underdetermined system of linear equations

$$\left\langle \sum_{l \in S_j} \sum_{\kappa=1}^{n^2} w_{l,\kappa}^{(j,k)} \varphi_{l,\kappa}, \upsilon_{j,\widetilde{k}} \right\rangle = 0, \quad \widetilde{k} = 1, 2, \ldots, n^2, \tag{23}$$

where $w_{l,\kappa}^{(j,k)}$ are unknown. The definition of $\upsilon_{j,\widetilde{k}}$ shows that it can be re-expressed as a translation of $\upsilon_{1,\widetilde{k}}$, i.e., $\upsilon_{j,\widetilde{k}}(x, y) = \upsilon_{1,\widetilde{k}}(x - x_j + x_1, y - y_j + y_1)$. Thus, when $\psi_{j,k}(x, y) = \psi_{1,k}(x - x_j + x_1, y - y_j + y_1)$ and $\langle \psi_{1,k}, \upsilon_{1,\widetilde{k}} \rangle = 0$, there holds that $\langle \psi_{j,k}, \upsilon_{j,\widetilde{k}} \rangle = 0$. This implies that $\psi_{j,k}$ can be constructed by solving (23) with $j = 1$. More specifically, $\psi_{j,k} = \sum_{l \in S_j} \sum_{\kappa=1}^{n^2} w_{l-j+1,\kappa}^{(1,k)} \varphi_{l,\kappa}$. In sum, a set of $3n^2$ linearly independent special solutions of (23) with $j = 1$ yields a basis for $\mathbb{W}_j$, $j \in \{1, 2, \ldots, p^2\}$. We may further follow the Gram-Schmidt process to orthonormalize the resulting basis functions, thereby obtaining the orthonormal basis $\{\psi_{j,k}\}$.

Appendix C
Generation of Matrix B for the Composition Model of Regularizer II

We provide details for the generation of matrix B and present the composition form of regularizer R_{II}. In fact, the constructed multiscale basis $\{\psi_{j,k} : k = 1, 2, \ldots, n^2\}$ supported in each square coarse grid Ω_j is a repetition of one another and has the tensor product structure. Therefore, each multiscale basis function is independent of index j and can be factorized into the product of two univariate functions: $\psi_{j,k}(x, y) = \phi_{k_1}(x)\phi_{k_2}(y)$ with ϕ_l, $l \in \{1, 2, \ldots, 2n\}$, being a univariate basis function supported in a 1D coarse grid. Below, we give the relation of k with k_1 and k_2. Note that the tensor product of two 1D coarse grids gives a square coarse grid. Specifically, we introduce the index sets $\mathbb{I}_1 := \{(k_1, k_2) : k_1 = 1, 2, \ldots, n; k_2 = n+1, n+2, \ldots, 2n\}$, $\mathbb{I}_2 := \{(k_1, k_2) : k_1 = n+1, n+2, \ldots, 2n; k_2 = 1, 2, \ldots, n\}$, and $\mathbb{I}_3 := \{(k_1, k_2) : k_1 = n+1, n+2, \ldots, 2n; k_2 = n+1, n+2, \ldots, 2n\}$, so that $\{\phi_{k_1}(\bullet)\phi_{k_2}(\star) : (k_1, k_2) \in \mathbb{I}_1 \cup \mathbb{I}_2 \cup \mathbb{I}_3\}$ forms the desired tensor product orthonormal basis for $\mathbb{W}_j$. Moreover, the subscript k relates to the index pair (k_1, k_2) via the following piecewise expression

$$k = \begin{cases} k_1 + (k_2 - n - 1)n, & \text{if } (k_1, k_2) \in \mathbb{I}_1, \\ k_1 - n + (k_2 - 1)n + n^2, & \text{if } (k_1, k_2) \in \mathbb{I}_2, \\ k_1 - n + (k_2 - n - 1)n + 2n^2, & \text{if } (k_1, k_2) \in \mathbb{I}_3. \end{cases}$$

We readily see that k runs from 1 to $3n^2$. We introduce two commands to anchor a matrix to its vectorization: the command Vec unwraps the matrix into a vector by column stacking, as described in Section II-B, whereas the command Fold operates in the opposite way. In this way, we may rewrite the regularizer (16) as

$$\sum_{(k_1,k_2) \in \mathbb{I}_1 \cup \mathbb{I}_2 \cup \mathbb{I}_3} \left\| \sum_{\widetilde{k}_1, \widetilde{k}_2 = 0}^{n-1} \text{Vec}\left((\tau_{k_1,\widetilde{k}_1} \otimes \tau_{k_2,\widetilde{k}_2}^\top) * \text{Fold}(\mathbf{f}_{\widetilde{k}}) \right) \right\|_1$$

with $\tau_{l,\widetilde{l}}$, $l \in \{1, 2, \ldots, 2n\}$, $\widetilde{l} \in \{0, 1, \ldots, n-1\}$, being 1D filters associated with polynomials of degree $\widetilde{l}$ in the univariate basis function ϕ_l and $\widetilde{k} = \widetilde{k}_1 + 1 + \widetilde{k}_2\, n$. For the purpose of algorithmic development, we express the above two-dimensional convolutions in matrix forms. Specifically, under the symmetric boundary condition, we let $H_{l,\widetilde{l}}$ denote a $p \times p$ circulant matrix specified by the filter $\tau_{l,\widetilde{l}}$. Then we have that

$$\text{Vec}\left((\tau_{k_1,\widetilde{k}_1} \otimes \tau_{k_2,\widetilde{k}_2}^\top) * \text{Fold}(\mathbf{f}_{\widetilde{k}}) \right) = (H_{k_1,\widetilde{k}_1} \otimes H_{k_2,\widetilde{k}_2})\mathbf{f}_{\widetilde{k}}.$$

We further let $D_{k_1,k_2,\widetilde{k}} := H_{k_1,\widetilde{k}_1} \otimes H_{k_2,\widetilde{k}_2} \in \mathbb{R}^{p^2 \times p^2}$ and $W_k := \left[D_{k_1,k_2,1} \;\; D_{k_1,k_2,2} \;\; \cdots \;\; D_{k_1,k_2,n^2} \right] \in \mathbb{R}^{p^2 \times d}$. With this notation, we have that

$$B = \left[W_1^\top \quad W_2^\top \quad \cdots \quad W_{3n^2}^\top \right]^\top, \tag{24}$$

and choose Φ as the ℓ^1-norm on $\mathbb{R}^{3n^2\ p^2}$. Then we can readily verify that composition $\Phi \circ B$ is equal to the regularizer (16).

Finally, we describe the 1D filters associated with the univariate multiscale piecewise linear basis functions used in the numerical simulation presented in Section III. The tensor products of these 1D filters may form a two-dimensional orthogonal piecewise linear filter bank. In particular, we set the highest coordinate order n to be 2 and thus have a tensor product multiscale piecewise linear basis for each space $\mathbb{W}_j$. Clearly, we see that $\dim(\mathbb{W}_j) = 12$. Each tensor product piecewise linear basis function can be factorized into the product of two univariate piecewise linear basis functions. Recalling the notation in Section II-C, we denote by $\{\phi_l : l = 1, 2, 3, 4\}$ the set of univariate basis functions. These univariate functions are all piecewise linear polynomials supported in the 1D coarse grid. With the above preparation, the 1D filters associated with polynomials of degree 0 in the univariate multiscale piecewise linear basis functions are given by

$$\tau_{1,0} = [1, 1], \quad \tau_{2,0} = [-1, 1], \quad \tau_{3,0} = [-1, 1], \quad \tau_{4,0} = [1, -1].$$

The 1D filters associated with polynomials of degree 1 in the above univariate basis functions are

$$\tau_{1,1} = \left[\frac{1}{2}, \frac{1}{2}\right], \quad \tau_{2,1} = \left[-\frac{1}{3}, \frac{2}{3}\right], \quad \tau_{3,1} = \left[-\frac{3}{2}, \frac{1}{2}\right],$$
$$\tau_{4,1} = \left[\frac{1}{3}, \frac{2}{3}\right].$$

Under the symmetric boundary condition, the $p \times p$ circulant matrix representations of the above series of 1D filters are given by

$$H_{1,0} = \begin{bmatrix} 2 & & & \\ 1 & 1 & & \\ & \ddots & \ddots & \\ & & 1 & 1 \end{bmatrix}, \quad H_{2,0} = \begin{bmatrix} 0 & & & \\ -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{bmatrix},$$

and $H_{3,0} = -H_{4,0} = H_{2,0}$,

$$H_{1,1} = \begin{bmatrix} 1 & & & \\ \frac{1}{2} & \frac{1}{2} & & \\ & \ddots & \ddots & \\ & & \frac{1}{2} & \frac{1}{2} \end{bmatrix}, \quad H_{2,1} = \begin{bmatrix} \frac{1}{3} & & & \\ -\frac{1}{3} & \frac{2}{3} & & \\ & \ddots & \ddots & \\ & & -\frac{1}{3} & \frac{2}{3} \end{bmatrix},$$

$$H_{3,1} = \begin{bmatrix} -1 & & & \\ -\frac{3}{2} & \frac{1}{2} & & \\ & \ddots & \ddots & \\ & & -\frac{3}{2} & \frac{1}{2} \end{bmatrix}, \quad H_{4,1} = \begin{bmatrix} 1 & & & \\ \frac{1}{3} & \frac{2}{3} & & \\ & \ddots & \ddots & \\ & & \frac{1}{3} & \frac{2}{3} \end{bmatrix}.$$

REFERENCES

[1] D. J. Brenner and R. K. Sachs, "Estimating radiation-induced cancer risks at very low doses: Rationale for using a linear no-threshold approach," *Radiat. Environ. Biophys.*, vol. 44, no. 4, pp. 253–256, 2006.

[2] M. A. King *et al.*, "Attenuation, scatter, and spatial resolution compensation in SPECT," in *Emission Tomography: The Fundamentals of PET and SPECT*, M. N. Wernick and J. N. Aarsvold, Eds. Cambridge, MA, USA: Academic, 2004, pp. 473–498.

[3] V. Y. Panin, F. Kehren, C. Michel, and M. Casey, "Fully 3-D PET reconstruction with system matrix derived from point source measurements," *IEEE Trans. Med. Imag.*, vol. 25, no. 7, pp. 907–921, Jul. 2006.

[4] F. A. Kotasidis *et al.*, "Single scan parameterization of space-variant point spread functions in image space via a printed array: The impact for two PET/CT scanners," *Phys. Med. Biol.*, vol. 56, no. 10, pp. 2917–2942, 2011.

[5] C. A. Micchelli and Y. Xu, "Using the matrix refinement equation for the construction of wavelets on invariant sets," *Appl. Comput. Harmon. Anal.*, vol. 1, no. 4, pp. 391–401, 1994.

[6] C. A. Micchelli and Y. Xu, "Reconstruction and decomposition algorithms for biorthogonal multiwavelets," *Multidimensional Syst. Signal Process.*, vol. 8, no. 1, pp. 31–69, 1997.

[7] Z. Chen, C. A. Micchelli, and Y. Xu, "A construction of interpolating wavelets on invariant sets," *Math. Comput.*, vol. 68, no. 22, pp. 1569–1587, 1999.

[8] Y. Lu, L. Shen, and Y. Xu, "Integral equation models for image restoration: High accuracy methods and fast algorithms," *Inverse Problems*, vol. 26, no. 4, p. 045006, 2010.

[9] Z. Chen, C. A. Micchelli, and Y. Xu, *Multiscale Methods for Fredholm Integral Equations*. Cambridge, U.K.: Cambridge Univ. Press, 2015.

[10] Y. Liu, L. Shen, Y. Xu, and H. Yang, "A collocation method solving integral equation models for image restoration," *J. Integral Equ. Appl.*, vol. 28, no. 2, pp. 263–307, 2016.

[11] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D, Nonlinear Phenomena*, vol. 60, nos. 1–4, pp. 259–268, 1992.

[12] E. Jonsson, S. C. Huang, and T. Chan, "Total variation regularization in positron emission tomography," Univ. California, Los Angeles, Los Angeles, CA, USA, UCLA CAM Rep. 98-48, 1998.

[13] V. Y. Panin, G. L. Zeng, and G. T. Gullberg, "Total variation regulated EM algorithm [SPECT reconstruction]," *IEEE Trans. Nucl. Sci.*, vol. 46, no. 6, pp. 2202–2210, Dec. 1999.

[14] C. A. Micchelli, L. Shen, and Y. Xu, "Proximity algorithms for image models: Denoising," *Inverse Problems*, vol. 27, no. 4, p. 045009, 2011.

[15] A. Krol, S. Li, L. Shen, and Y. Xu, "Preconditioned alternating projection algorithms for maximum *a posteriori* ECT reconstruction," *Inverse Problems*, vol. 28, no. 11, p. 115005, 2012.

[16] Q. Li, L. Shen, Y. Xu, and N. Zhang, "Multi-step fixed-point proximity algorithms for solving a class of optimization problems arising from image processing," *Adv. Comput. Math.*, vol. 41, no. 2, pp. 387–422, 2015.

[17] S. Li *et al.*, "Effective noise-suppressed and artifact-reduced reconstruction of SPECT data using a preconditioned alternating projection algorithm," *Med. Phys.*, vol. 42, no. 8, pp. 4872–4887, 2015.

[18] Z. Wu, S. Li, X. Zeng, Y. Xu, and A. Krol, "Reducing staircasing artifacts in SPECT reconstruction by an infimal convolution regularization," *J. Comput. Math.*, vol. 34, no. 6, pp. 624–645, 2016.

[19] R. M. Lewitt and S. Matej, "Overview of methods for image reconstruction from projections in emission computed tomography," *Proc. IEEE*, vol. 91, no. 10, pp. 1588–1611, Oct. 2003.

[20] A. J. Reader and J. Verhaeghe, "4D image reconstruction for emission tomography," *Phys. Med. Biol.*, vol. 59, no. 22, pp. 371–418, 2014.

[21] R. M. Lewitt, "Multidimensional digital image representations using generalized Kaiser–Bessel window functions," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 7, no. 10, pp. 1834–1846, 1990.

[22] R. M. Lewitt, "Alternatives to voxels for image representation in iterative reconstruction algorithms," *Phys. Med. Biol.*, vol. 37, no. 3, pp. 705–716, 1992.

[23] S. Matej and R. M. Lewitt, "Practical considerations for 3-D image reconstruction using spherically symmetric volume elements, " *IEEE Trans. Med. Imag.*, vol. 15, no. 1, pp. 68–78, Feb. 1996.

[24] M. E. Daube-Witherspoon, S. Matej, J. S. Karp, and R. M. Lewitt, "Application of the row action maximum likelihood algorithm with spherical basis functions to clinical PET imaging," *IEEE Trans. Nucl. Sci.*, vol. 48, no. 1, pp. 24–30, Feb. 2001.

[25] P. L. Combettes and V. R. Wajs, "Signal recovery by proximal forward-backward splitting," *Multiscale Model. Simul.*, vol. 4, no. 4, pp. 1168–1200, 2005.

[26] K. Lange, M. Bahn, and R. Little, "A theoretical study of some maximum likelihood algorithms for emission and transmission tomography," *IEEE Trans. Med. Imag.*, vol. 6, no. 2, pp. 106–114, Jun. 1987.

[27] N. H. Clinthorne, T.-S. Pan, P.-C. Chiao, W. L. Rogers, and J. A. Stamos, "Preconditioning methods for improved convergence rates in iterative reconstructions," *IEEE Trans. Med. Imag.*, vol. 12, no. 1, pp. 78–83, Mar. 1993.

[28] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. New York, NY, USA: Springer, 2011, pp. 197–206.

[29] American College of Radiology, Reston, VA, USA. *American College of Radiology Nuclear Medicine Accreditation Program: Testing Instructions*. Accessed: Oct. 19, 2018. [Online]. Available: https://www.acraccreditation.org/-/media/ACRAccreditation/Documents/NucMed-PET/Nuclear-Medicine-Forms/NucMed-Testing-Instructions.pdf?la=en

[30] B. S. Everitt, *The Cambridge Dictionary of Statistics*. Cambridge, U.K.: Cambridge Univ. Press, 1998.

[31] H. H. Barrett, J. Yao, J. P. Rolland, and K. J. Myers, "Model observers for assessment of image quality," *Proc. Nat. Acad. Sci. USA*, vol. 90, no. 21, pp. 9758–9765, 1993.

[32] E. C. Frey, K. L. Gilland, and B. M. W. Tsui, "Application of task-based measures of image quality to optimization and evaluation of three-dimensional reconstruction-based compensation methods in myocardial perfusion SPECT," *IEEE Trans. Med. Imag.*, vol. 21, no. 9, pp. 1040–1050, Sep. 2002.

[33] X. Niu, Y. Yang, M. A. King, and M. N. Wernick, "Detectability of perfusion defect in five-dimensional gated-dynamic cardiac SPECT images," *Med. Phys.*, vol. 37, no. 9, pp. 5102–5112, 2010.

[34] M. Ljungberg, *Monte Carlo Calculations in Nuclear Medicine: Applications in Diagnostic Imaging*, M. Ljungberg, S. E. Strand, and M. A. King, Eds., 2nd ed. Boca Raton, FL, USA: CRC Press, 2012.

[35] K. Ogawa, Y. Harata, T. Ichihara, A. Kubo, and S. Hashimoto, "A practical method for position-dependent Compton-scatter correction in single photon emission CT," *IEEE Trans. Med. Imag.*, vol. 10, no. 3, pp. 408–412, Sep. 1991.

[36] J. E. Juni *et al.* Procedure Guideline for Brain Perfusion SPECT Using 99 mTc Radiopharmaceuticals 3.0. Society of Nuclear Medicine and Molecular Imaging, VA, USA. Accessed: Jul. 1, 2009. [Online]. Available: http://snmmi.files.cms-plus.com/docs/Brain_SPECT_Guideline_2003.pdf

[37] A. Rahmim and J. Tang, "Noise propagation in resolution modeled PET imaging and its impact on detectability," *Phys. Med. Biol.*, vol. 58, no. 19, pp. 6945–6968, 2013.