INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 18, Number 6, Pages 834–864

AN INTEGRAL EQUATION MODEL FOR PET IMAGING

XINHUANG TANG[1,2], CHARLES ROSS SCHMIDTLEIN[2], SI LI[3], AND YUESHENG XU[4,*]

Abstract. Positron emission tomography (PET) is traditionally modeled as discrete systems. Such models may be viewed as piecewise constant approximations of the underlying continuous model for the physical processes and geometry of the PET imaging. Due to the low accuracy of piecewise constant approximations, discrete models introduce an irreducible modeling error which fundamentally limits the quality of reconstructed images. To address this bottleneck, we propose an integral equation model for the PET imaging based on the physical and geometrical considerations, which describes accurately the true coincidences. We show that the proposed integral equation model is equivalent to the existing idealized model in terms of line integrals which is accurate but not suitable for numerical approximation. The proposed model allows us to discretize it using higher accuracy approximation methods. In particular, we discretize the integral equation by using the collocation principle with piecewise linear polynomials. The discretization leads to new ill-conditioned discrete systems for the PET reconstruction, which are further regularized by a novel wavelet-based regularizer. The resulting non-smooth optimization problem is then solved by a preconditioned proximity fixed-point algorithm. Convergence of the algorithm is established for a range of parameters involved in the algorithm. The proposed integral equation model combined with the discretization, regularization, and optimization algorithm provides a new PET image reconstruction method. Numerical results reveal that the proposed model substantially outperforms the conventional discrete model in terms of the consistency to simulated projection data and reconstructed image quality. This indicates that the proposed integral equation model with appropriate discretization and regularizer can significantly reduce modeling errors and suppress noise, which leads to improved image quality and projection data estimation.

Key words. Positron emission tomography, integral equation model, image reconstruction.

1. Introduction

Positron emission tomography (PET) is a critical tool for the *in vivo* detection of cancer, due to its exquisite sensitivity to positron emitting radio-labeled molecules. In PET, patients are administered a small amounts of a radio-labeled tracer that has affinities for a particular molecular target (*e.g.*, ^{18}F-fluorodeoxyglucose (FDG) for glucose metabolism). After sufficient time for the tracer to circulate and bind (typically 1-hour for ^{18}F-FDG), the patient is then placed in a PET scanner where the photons emitted from positron annihilation are counted. These photons are emitted (nearly) back-to-back, and are counted in coincidence since their emission is correlated in time. The coincidence count data are then reconstructed via a deblurring/denoising approach based on a physical model into images that represents the tracer's biodistribution. In PET, the physical model usually considers three types of detected coincident counts, that is, true coincidences, scatter coincidences and random coincidences. True coincidences come from the same decay events that have not been scattered before hitting the detectors, whereas scatter coincidences

Received by the editors July 5, 2021.
2000 *Mathematics Subject Classification.* 65R20, 65R32.
*Corresponding author: Yuesheng Xu; E-mail: y1xu@odu.edu.

are the events that have been scattered. Random coincidences involve two separate decay events occurring within a physically allowable timing window (coincidence window, typically 2-5 ns). In current practice, a reduced model, based on the true coincidences portion of the physical model, is generally used and discussed in the literature. In this reduced model, the scatter and random counts are externally estimated as additive count corrections. Discretization based on a piecewise constant approximation of the reduced model leads to the commonly used discrete system model. However, the resulting discrete model introduces irreducible modeling error due to its use of the piecewise constant approximation. Moreover, even though a great deal of work has been done in investigating the reduced model, the explicit development of higher accuracy representations of the physical model and its subsequent regularization is less explored.

As mentioned in a comprehensive review [10], substantial focus has been placed on improving the accuracy of the physical models. Accurately modeling the physical effects can improve the fidelity of the model, and further increase the quality (such as spatial resolution, lesion contrast) of the reconstructed image. If the discrete nature of the projection data is ignored, the attenuated Radon transform model (ART) [7, 33] represents an idealized and accurate model for true coincidences. However, in a real PET data acquisition, a new form of the ART model that accounts for the discreteness of data is required. Existing models for true coincidences, mainly factorization approaches, can be considered as the discretization of the ART model and still have the non-negligible modeling errors. Factorization approaches [4, 24, 26, 32] formulate the system matrix as the product of independent sparse matrices (*e.g.*, attenuation, positron range and detector efficiency). This leads to a sparse system matrix, which substantially reduces storage space and can help reduce reconstruction times. However, these factorizations are based on the decoupling of the integral kernel that can introduce modeling errors that fundamentally limit image quality. These models also fail to consider the continuous nature of the problem and can be regarded as piecewise constant approximations of the integral equation model. The data/model mismatch mentioned above results in irreducible modeling error and thus imposes a fundamental bottleneck in the improvement of image quality. Therefore, it is the goal of this study to reduce the modeling error through a holistic approach via establishing a continuous physical model and its discretization of higher order accuracy.

In this study, we propose an integral equation model combined with the discretization, regularization, and optimization algorithm, leading to a new PET image reconstruction method. In contrast to existing discrete models (DM), the true coincidences model proposed in this paper improves consistency with the underlying physics and geometry of the PET imaging. We consider the attenuation and effective detection angle of each point source to determine the contribution weights of the true coincidences. The model is then discretized by using piecewise polynomials, resulting in a novel discrete system for PET reconstruction. Piecewise polynomials can more precisely express the tracer distribution than a piecewise constant function, and result in a higher order of accuracy than a piecewise constant function. Based on the smoothness of the tracer distribution function, we designed a suitable sparse wavelet-based regularizer to penalize the approximation error of the discretization and to treat the ill-posedness of the proposed piecewise

polynomial discrete system. The resulting non-smooth minimization problem is solved by employing a preconditioned proximity fixed-point algorithm.

We validate numerically the accuracy of the proposed true coincidences model for PET imaging by comparing the accuracy of the projection data of the proposed model with that of the conventional discrete implementation of the ART model. In addition, we compare the performance of the proposed reconstruction method with the conventional discrete method. Numerical results that we have obtained reveal that the proposed integral equation model with the regularization can reduce significantly the modeling and approximation errors, suppress noise, and improve the reconstructed image's quality, demonstrating its remarkable improvement over conventional discrete models.

This paper is organized in 8 sections. In Section 2, we introduce an integral equation model for true coincidences in PET imaging. The equivalence of the proposed integral equation model with an existing idealized continuous-to-discrete ART model is established in section 3. In Section 4, we discretize the proposed integral equation by using a piecewise linear basis. In Section 5, we apply the wavelet-based regularized maximum likelihood estimation for the resulting discrete system and describe a fixed-point iterative algorithm for solving the minimization problem. Section 6 is devoted to deriving ranges for the parameters appearing in the fixed-point iterative algorithm which ensure its convergence. We present in Section 7 numerical results that demonstrate the performance of the proposed integral equation model for PET image reconstruction. Finally, a conclusion is drawn in Section 8.

2. Integral Equation Model for PET Imaging

In this section, we introduce an integral equation model for PET imaging. The proposed model is constructed based on the physical effects and geometry of a PET system.

The PET imaging process is to reconstruct a tracer distribution function from data recorded by a PET scan. In the PET imaging system, the instrument records the coincidence events, that is, effectively detected pairs of gamma photon emitted from the radioactive tracer distributed inside the patient's body. The radiotracer emits a positively charged positron when it undergoes positron emission decay. After a random trip, the positron loses its kinetic energy, and interacts with an electron. A pair of gamma photons with 511 keV energy is then produced and emitted in approximately the opposite direction. The photon pair is detected by a pair of detectors after the attenuation, where the number of photons, detection time and energy information are recorded. There are three types of coincidence events recorded on the detectors: true coincidences, scatter coincidences and random coincidences. Among them, the true coincidences are the major events. Particularly, the two photons of a true coincidence come from the same decay event, and neither photon has been scattered before entering the detector. In a scatter coincidence, two photons come from the same decay event, but one or both of them undergo the Compton scattering. The random coincidence is an event in which two photons from different decay events are recorded.

There are many physical effects and geometry involved in the PET imaging process. These factors are mainly divided into three phases. The first phase occurs

before the photon pair is produced and consists mainly of the positron range, the positron annihilation and the non-collinearity. In the second phase, the photon pairs are generated and undergo attenuation and Compton scattering. The third phase occurs within the detector and includes the photoelectric effect and inter-crystal scatter. The physical effects and geometry of the PET imaging process have a significant impact on the data collection. Therefore, accurately modeling of these factors is crucial for PET imaging. In general, a PET imaging model is mainly based on the modeling of the true coincidences with the addition of several physical effects and geometric related factors. In this work, we focus on developing an integral equation model for the true coincidences, by combining the attenuation of photons and geometric structure of the system, but ignoring the positron range and the non-collinearity. A more comprehensive model will be given in a future project.

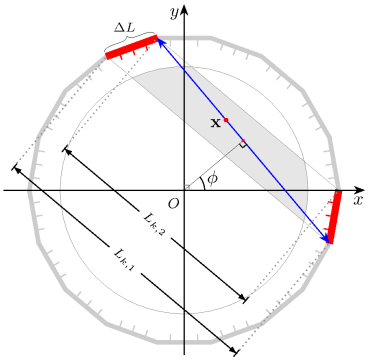
In an idealized PET environment, true coincidences can be accurately described by the ART model. Existing studies based on the ART model consist of two categories. One uses the inverse attenuated Radon transform (IART) [7, 29, 33] to give an approximate analytical form of the tracer distribution function. The other employs the collected counts on a line-of-response (LOR) to approximate the collected counts on a detector pair [4, 24, 26, 32, 33]. Both of these approaches fail to incorporate the properties of the detector's discrete distribution to give a specific expression of the relationship between the tracer distribution function and the collected data. The ART model is applicable only for the idealized PET environment where detectors are assumed to be continuously distributed. In fact, by considering the properties of the discrete distribution of the detectors, one can obtain a continuous-to-discrete ART model. However, the continuous-to-discrete ART model suffers from its limitations. It is difficult to expand the model by adding other physical consideration. Moreover, the line integral involved in the ART model may result in large approximation errors and high computational complexity in implementing it. The model we introduce in this section for the true coincidences takes a different viewpoint by considering the contribution weight of photons at each point source to a detector pair, avoiding using the line integral. We will prove in the next section that the proposed model is mathematically equivalent to the ART model. This equivalence ensures that the proposed model is a feasible replacement for the ART model, and more importantly, it overcomes the shortcoming of the latter.

We now briefly describe the geometry of a PET system considered in this work. As shown in Figure 1A, we use a cylindrical PET system with a circular full ring scanner. The detectors with width ΔL are distributed on a detector ring of radius R. We assume that there are N_d effective detector pairs that can detect the true coincidence events, and denote by g_k, $k = 1, 2, \ldots, N_d$, the true coincidence counts detected by the kth detector pair. The field-of-view (FOV) or the tracer distribution area is a disk with radius r, denoted by $\Omega \subset \mathbb{R}^2$. For the true coincidences without the positron range and non-collinearity, the gamma photon pairs emitted isotropically from each point source $\boldsymbol{x}$ in the FOV having the tracer distribution function $f(\boldsymbol{x})$ will be detected by the detectors, after interacting with the tissues inside the patient's body having the attenuation coefficient distribution function $\mu(\boldsymbol{x})$. Considering the contribution of the tracer distribution function f to the detected count vector $[g_1, g_2, \ldots, g_{N_d}]$ leads to our model.

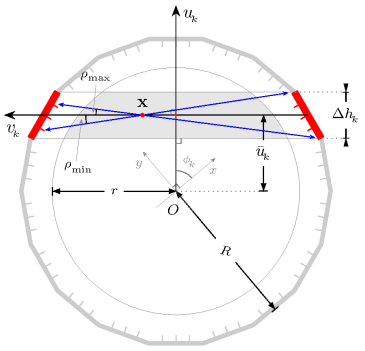
We now introduce the integral equation model for the true coincidences. For each detector pair, its recorded true coincidence counts come from a small region within the FOV, which is called an effective contribution region. Each point source in the region will have a weighted contribution to the total counts recorded by the detector pair. Specifically, for the kth detector pair, $k = 1, 2, \ldots, N_d$, we denote by $\Omega_k \subset \Omega$ the effective contribution region, and by $\mathcal{W}_k$ the contribution weight of the point source. The true coincidence counts g_k are the total contribution of the function f with the contribution weight $\mathcal{W}_k$ over the region Ω_k. That is,

$$g_k = \int_{\Omega_k} \mathcal{W}_k(\boldsymbol{x}) f(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}, \quad \text{for} \quad k = 1, 2, \ldots, N_d. \tag{1}$$

The contribution weight $\mathcal{W}_k$ is determined by the material attenuation, isotropic multi-angle projection and the geometry of the PET system, which will be detailed later.



(A)



(B)

FIGURE 1. In each figure, the outer circle is the detector ring, the inner circular area is the FOV. The thick red lines on the outer circle are detectors, and the blue lines with two arrows are the LORs. The strip between the two detectors is the SOR.

The effective contribution region Ω_k is determined only by the system geometry. As shown in Figure 1A, we introduce a fixed xy-coordinate system whose origin is the center of the FOV. A line-of-response, the line connecting the two detectors of a detector pair, represents the allowable path of an unscattered photon coincidence pair. We next describe the LOR passing through a point source $\boldsymbol{x} \coloneqq (x, y) \in \Omega$ and having a normal forming an angle $\phi \in [0, \pi)$ with the positive direction of the x-axis, as shown in Figure 1A. The LOR can be described by

$$l(\boldsymbol{x}, \phi) \coloneqq \{\bar{\boldsymbol{x}} \coloneqq (\bar{x}, \bar{y})\colon \bar{x} \coloneqq x - t\sin\phi, \bar{y} \coloneqq y + t\cos\phi, t \in \mathbb{R}\}. \tag{2}$$

For a particular detector pair, all the effective LORs form a strip-of-response (SOR). The intersection of the SOR and FOV is the effective contribution region Ω_k for the kth detector pair. It is the gray area shown in Figure 1.

The contribution weight $\mathcal{W}_k(\boldsymbol{x})$ is the probability that the photon pairs emitted from a point source $\boldsymbol{x} \in \Omega_k$ and detected by the kth detector pair. For $\boldsymbol{x} \in \Omega_k$, only the photons that emitting toward the kth detector pair and have not been attenuated are counted in $g_{_k}$. Thus, the detection probability by a given detector pair is the product of the probability of a photon pair emitting towards the detector pair and the probability of the photon pair avoiding attenuation. For an isotropic source in the 2D plane, the emission probability of a photon at an angle between ρ and $\rho + \mathrm{d}\rho$ is given by $\mathrm{d}\rho/(2\pi)$, whereas for a photon coincidence pair, it is $\mathrm{d}\rho/\pi$. The attenuation of a photon beam is the amount of the light intensity (or the photon number) lost as the beam passes through a medium. According to the Beer-Lambert law [2], the amount of attenuated photons is proportional to the length of the medium through which the photon beam penetrates and the attenuation coefficient of the medium. The attenuation kernel may be written as an exponential function of the total attenuation on the path through the medium. By the definition (2) of the LOR, the attenuation distribution function μ at a point $\bar{\boldsymbol{x}} \in l(\boldsymbol{x}, \phi)$ can be also represented by $\mu(x - t\sin\phi, y + t\cos\phi)$. In addition, since the attenuation distribution function is defined on the FOV, the effective attenuation region of a LOR $l(\boldsymbol{x}, \phi)$ satisfies $t_1 \le t \le t_2$ with

$$\begin{cases} t_1 = x\sin\phi - y\cos\phi - \sqrt{r^2 - (x\cos\phi + y\sin\phi)^2}, \\ t_2 = x\sin\phi - y\cos\phi + \sqrt{r^2 - (x\cos\phi + y\sin\phi)^2}. \end{cases}$$

Therefore, the attenuation on the LOR $l(\boldsymbol{x}, \phi)$ is given by

$$\mathcal{K}_{\mathrm{a}}(\boldsymbol{x}, \phi) \coloneqq \exp\left\{-\int_{t_1}^{t_2} \mu(x - t\sin\phi, y + t\cos\phi)\,\mathrm{d}t\right\}. \tag{3}$$

The detection probability on this LOR is then given by $\mathcal{K}_{\mathrm{a}}(\boldsymbol{x}, \phi)\,\mathrm{d}\rho/\pi$. Next, we account for the detection probability on all possible LORs through the point source $\boldsymbol{x}$. According to the geometry of the PET system, shown in Figure 1B, the effective detection angle of the point $\boldsymbol{x}$ and the normal angle of the LOR $l(\boldsymbol{x}, \phi)$ have the relation $\phi = \phi_k - \rho$. Here ϕ_k is the angle between the positive direction of the x-axis and a line from the origin perpendicular to the SOR. For a given detector pair, the effective LORs fall in its SOR, thus, the angle ρ is constrained by the minimum effective detection angle $\rho_{_{\min,k}}(\boldsymbol{x})$ and the maximum effective detection angle $\rho_{_{\max,k}}(\boldsymbol{x})$. Therefore, the total contribution weight $\mathcal{W}_k(\boldsymbol{x})$ for the kth detector

pair is an integral of the detection probability over all possible LORs. That is,

$$\mathcal{W}_k(\boldsymbol{x}) \coloneqq \frac{1}{\pi}\int_{\rho_{\min,k}(\boldsymbol{x})}^{\rho_{\max,k}(\boldsymbol{x})} \mathcal{K}_{\mathrm{a}}(\boldsymbol{x}, \phi_k - \rho)\,\mathrm{d}\rho. \tag{4}$$

We next specify $\rho_{\min,k}$ and $\rho_{\max,k}$ that appear in (4). To this end, we introduce a new u_kv_k-coordinate system for the kth detector pair by rotating the xy-coordinate system counter-clockwise by angle ϕ_k, and translating the system by $\bar{u}_k$ along the u_k-axis, where $\bar{u}_k$ is the number illustrated in Figure 1B. We let Δh_k denote the width of the SOR in Figure 1B, $L_{k,1}$ and $L_{k,2}$ denote the lengths of the two edges of the SOR in Figure 1A, respectively. The coordinates (u_k, v_k) of a point source $\boldsymbol{x} \coloneqq (x, y) \in \Omega_k$ in the u_kv_k-coordinate system are given by

$$u_k \coloneqq x\cos\phi_k + y\sin\phi_k - \bar{u}_k \quad \text{and} \quad v_k \coloneqq -x\sin\phi_k + y\cos\phi_k.$$

Finally, the minimum and maximum effective detection angle of the point source $\boldsymbol{x}$ are given by

$$\begin{cases} \rho_{\min,k}(\boldsymbol{x}) \;\coloneqq -\min\left\{\arctan\dfrac{\Delta h_k - 2u_k}{L_{k,1} + 2v_k}, \arctan\dfrac{\Delta h_k + 2u_k}{L_{k,2} - 2v_k}\right\}, \\ \rho_{\max,k}(\boldsymbol{x}) \;\coloneqq \quad \min\left\{\arctan\dfrac{\Delta h_k - 2u_k}{L_{k,1} - 2v_k}, \arctan\dfrac{\Delta h_k + 2u_k}{L_{k,2} + 2v_k}\right\}. \end{cases}$$

With the above discussion, a complete integral equation model for the true coincidence counts may be described as

$$g_{_k} = \frac{1}{\pi}\int_{\Omega_k}\int_{\rho_{\min,k}(\boldsymbol{x})}^{\rho_{\max,k}(\boldsymbol{x})} \mathcal{K}_{\mathrm{a}}(\boldsymbol{x}, \phi_k - \rho) f(\boldsymbol{x})\,\mathrm{d}\rho\,\mathrm{d}\boldsymbol{x}, \quad k = 1, 2, \ldots, N_d. \tag{5}$$

This is the model on which our reconstruction method is based.

3. Equivalence of the Proposed Model with the Continuous-to-Discrete ART Model

The existing ART model [1, 3, 7, 29, 33] is the accurate model for true coincidences in an idealized PET data acquisition. Hence, the continuous-to-discrete (CtoD) ART model that considers the discrete distribution of detectors can give the accurate true coincidences for each detector pair. However, the CtoD ART model is not computationally efficient due to the involvement of line integrals. We establish in this section the mathematical equivalence between the proposed model and the CtoD ART model. This ensures that the proposed model which is more computationally efficient can be used as a feasible surrogate of the idealized model and can overcome the computational shortcoming of the CtoD ART model. A comparison of the performance of the proposed model with that of a conventional discretization of the CtoD ART model will be presented in the numerical experiment section.

Before the introduction of the ART model, we first describe a new form for the LOR and then consider the attenuation for this new form. This alternative formulation describes the line representing the LOR using the normal angle and the distance from the origin. We note that the new form of LOR and the form in (2) can be converted to each other through the transformation between their variables. Let $\mathcal{U} \coloneqq [-r, r]$, $\Theta \coloneqq [0, \pi)$ and $u \coloneqq x\cos\phi + y\sin\phi$. Clearly, $|u|$ is the distance from the origin to the line-of-response. In PET, for the two parallel

lines with same distance from the origin, if the angles between the perpendicular lines from the origin to the line and the positive direction of the x-axis are ϕ and $\phi + \pi$, respectively, these two lines are considered to have the same normal angle ϕ. To distinguish these two lines, we use $|u|$ and $-|u|$ to label these two lines, respectively. Thus, each line-of-response has a corresponding pair $(|u|, \phi)$, if the foot of the perpendicular line is above the x-axis, and $(-|u|, \phi)$, otherwise. It can be verified that in the former case, $u \geq 0$; while in the latter case, $u \leq 0$. In summary, a line-of-response has its corresponding pair (u, ϕ), with $u \in \mathcal{U}$ and angle $\phi \in \Theta$. By denoting $v \coloneqq t - x\sin\phi + y\cos\phi \in \mathbb{R}$, a new form of the line-of-response is given by

$$\tilde{l}(u,\phi) \coloneqq \{\bar{\boldsymbol{x}} \coloneqq (\bar{x},\bar{y})\colon \bar{x} \coloneqq u\cos\phi - v\sin\phi, \bar{y} \coloneqq u\sin\phi + v\cos\phi, v \in \mathbb{R}\}. \tag{6}$$

We then consider the attenuation on the LOR $\tilde{l}(u,\phi)$. Since the attenuation distribution function μ is defined on Ω, we only consider the attenuation of the interval on the line-of-response $\tilde{l}(u,\phi)$ within the FOV. Combining the geometry in PET, we have the following constraint for the variable v

$$v \in [v_1(u), v_2(u)] \ \text{ with } \ v_1(u) \coloneqq -\sqrt{r^2 - u^2} \ \text{ and } \ v_2(u) \coloneqq \sqrt{r^2 - u^2} \ \text{ for } \ u \in \mathcal{U}.$$

For any $\bar{\boldsymbol{x}} \in \tilde{l}(u,\phi)$, the attenuation distribution function $\mu(\bar{\boldsymbol{x}})$ can be expressed by $\mu(u\cos\phi - \bar{v}\sin\phi, u\sin\phi + \bar{v}\cos\phi)$. Similar to (3), the attenuation on the LOR $\tilde{l}(u,\phi)$ is given by

$$\widetilde{\mathcal{K}}_{\mathrm{a}}(u,\phi) \coloneqq \exp\left\{-\int_{v_1(u)}^{v_2(u)} \mu(u\cos\phi - \bar{v}\sin\phi, u\sin\phi + \bar{v}\cos\phi)\,\mathrm{d}\bar{v}\right\}. \tag{7}$$

We now describe the existing ART model for the true coincidences. In an idealized PET data acquisition, the detectors are assumed to be continuously distributed on the detector ring. The collected data at each point on the detectors can be precisely located so that the projected data are continuously distributed. Let g denote the continuous projection function on $\mathcal{U} \times \Theta$. The projection function g is an integral of the tracer distribution function f on the LOR $\tilde{l}(u,\phi)$ with the attenuation defined by (7). Again, the tracer distribution function f is defined on Ω, we only consider the photon pairs located at the interval on the line-of-response $\tilde{l}(u,\phi)$ within the FOV. For any point source $\boldsymbol{x} = (u\cos\phi - v\sin\phi, u\sin\phi + v\cos\phi)$, the ART model is given by

$$g(u,\phi) = \int_{v_1(u)}^{v_2(u)} \widetilde{\mathcal{K}}_{\mathrm{a}}(u,\phi) f(\boldsymbol{x})\,\mathrm{d}v, \quad \text{for all} \ \ (u,\phi) \in \mathcal{U} \times \Theta. \tag{8}$$

The above model (8) is only applicable for the idealized PET environment with continuous distribution of detectors. By considering the properties of the discrete distribution of the detectors, we can further obtain a CtoD ART model. We divide the range $\mathcal{U} \times \Theta$ of g into N_d ranges according to the N_d detector pairs. Specifically, we let

$$\tilde{u}_{k,1} \coloneqq \bar{u}_k - \Delta h_k/2, \quad \tilde{u}_{k,2} \coloneqq \bar{u}_k + \Delta h_k/2$$
$$\text{and} \ \ \eta_k(u) \coloneqq \min\{\zeta(u, \tilde{u}_{k,1}, L_{k,1}), \zeta(u, \tilde{u}_{k,2}, L_{k,2})\},$$

where

$$\zeta(u, u_*, L_*) \coloneqq \arccos\left(\frac{4uu_* - L_*\sqrt{L_*^2 - 4u^2 + 4u_*^2}}{4u^2 + L_*^2}\right).$$

For the kth detector pair, we define a set of points (u, ϕ) by

$$\mathcal{S}_k \coloneqq \{(u, \phi)\colon \tilde{u}_{k,1} \le u \le \tilde{u}_{k,2}, \phi_k - \eta_k(u) \le \phi \le \phi_k + \eta_k(u)\}.$$

Then, we have that a partition for $\mathcal{U} \times \Theta$, that is,

$$\mathcal{U} \times \Theta = \bigcup_{k=1}^{N_d} \mathcal{S}_k.$$

By considering the collected counts on all the LORs within the SOR of a given detector pair, we obtain the relationship between the collected data of the detector pair and the tracer distribution function. In particular, accumulating all $g(u, \phi)$ with $(u, \phi) \in \mathcal{S}_k$, we have the total collected data g_k for the kth detector pair. In addition, since the photon pairs emitted by the point sources are isotropic, a normalization factor $1/\pi$ needs to be considered in the modeling. The CtoD ART model is then given by

$$g_k = \frac{1}{\pi} \iint_{\mathcal{S}_k} \int_{v_1(u)}^{v_2(u)} \widetilde{\mathcal{K}}_{\mathrm{a}}(u, \phi) f(\boldsymbol{x}) \, \mathrm{d}v \, \mathrm{d}u \, \mathrm{d}\phi, \text{ for } k = 1, 2, \ldots, N_d. \tag{9}$$

The above CtoD ART model is considered as an accurate model for the true coincidences. We now establish the main result of this section that the proposed model is equivalent to the CtoD ART model. We first construct a bijective mapping between the two integration regions of the two models. For $k = 1, 2, \ldots, N_d$, we denote

$$\mathcal{D}_k \coloneqq \Omega_k \times [\rho_{\min,k}, \rho_{\max,k}]$$

by the integral region for the integration variables $(\boldsymbol{x}, \rho)$ of the proposed model, and

$$\mathcal{E}_k \coloneqq \mathcal{S}_k \times [v_1(u), v_2(u)]$$

by the integral region for the integration variables (u, ϕ, v) of the CtoD ART model. We then define the mapping

$$\mathcal{J}\colon (u, \phi, v) \in \mathcal{E}_k \to (\boldsymbol{x}, \rho) \in \mathcal{D}_k \tag{10}$$

by

$$\mathcal{J}(u, \phi, v) \coloneqq ((x(u, \phi, v), y(u, \phi, v)), \rho(u, \phi, v)),$$

where

$$\begin{cases} x(u, \phi, v) \coloneqq u\cos\phi - v\sin\phi, \\ y(u, \phi, v) \coloneqq u\sin\phi + v\cos\phi, \\ \rho(u, \phi, v) \coloneqq \phi_k - \phi. \end{cases}$$

Lemma 3.1. *The mapping $\mathcal{J}\colon (u, \phi, v) \to (\boldsymbol{x}, \rho)$ defined in* (10) *is bijective.*

Proof. We prove this lemma by showing that $\mathcal{J}$ is injective and surjective.

We first show that $\mathcal{J}$ is injective. Suppose that for (u', ϕ', v'), $(u'', \phi'', v'') \in \mathcal{E}_k$, we have that

$$\mathcal{J}(u', \phi', v') = \mathcal{J}(u'', \phi'', v'').$$

By the definition of $\mathcal{J}$, we observe that

$$\begin{cases} u' \cos \phi' - v' \sin \phi' = u'' \cos \phi'' - v'' \sin \phi'', \\ u' \sin \phi' + v' \cos \phi' = u'' \sin \phi'' + v'' \cos \phi'', \\ \phi_k - \phi' \qquad\qquad\quad = \phi_k - \phi''. \end{cases}$$

From the last equality, we get that $\phi' = \phi''$. Substituting this into the first two equalities, we obtain that $u' = u''$ and $v' = v''$. Therefore, $\mathcal{J}$ is injective.

It remains to prove that $\mathcal{J}$ is surjective. For any $(\boldsymbol{x}, \rho) \in \mathcal{D}_k$, we have that

$$\begin{cases} u(\boldsymbol{x}, \rho) \coloneqq \quad x \cos(\phi_k - \rho) + y \sin(\phi_k - \rho), \\ v(\boldsymbol{x}, \rho) \coloneqq -x \sin(\phi_k - \rho) + y \cos(\phi_k - \rho), \\ \phi(\boldsymbol{x}, \rho) \coloneqq \phi_k - \rho. \end{cases}$$

We now verify that

$$(u(\boldsymbol{x}, \rho), \phi(\boldsymbol{x}, \rho), v(\boldsymbol{x}, \rho)) \in \mathcal{E}_k.$$

Considering the geometry in Figure 1B, for $(\boldsymbol{x}, \rho) \in \mathcal{D}_k$, the minimum and maximum of $u(\boldsymbol{x}, \rho)$ can be taken by $\bar{u}_k - \Delta h_k/2$ and $\bar{u}_k + \Delta h_k/2$, respectively. Hence,

$$u(\boldsymbol{x}, \rho) \in \left[\bar{u}_k - \frac{\Delta h_k}{2}, \bar{u}_k + \frac{\Delta h_k}{2} \right].$$

Moreover, by the definition of $u(\boldsymbol{x}, \rho)$ and $v(\boldsymbol{x}, \rho)$, we obtain that

$$u^2(\boldsymbol{x}, \rho) + v^2(\boldsymbol{x}, \rho) = x^2 + y^2 \leq r^2.$$

This implies that

$$v(\boldsymbol{x}, \rho) \in \left[-\sqrt{r^2 - u^2(\boldsymbol{x}, \rho)}, \sqrt{r^2 - u^2(\boldsymbol{x}, \rho)} \right].$$

Substituting the minimum $\rho_{\min,k}$ and maximum $\rho_{\max,k}$ of ρ into the definition of $\phi(\boldsymbol{x}, \rho)$, the minimum and the maximum of $\phi(\boldsymbol{x}, \rho)$ can be given by $\phi_k - \eta_k(u)$ and $\phi_k + \eta_k(u)$, respectively. Thus,

$$\phi(\boldsymbol{x}, \rho) \in [\phi_k - \eta_k(u), \phi_k + \eta_k(u)].$$

Based on the value range of $u(\boldsymbol{x}, \rho), \phi(\boldsymbol{x}, \rho)$, and $v(\boldsymbol{x}, \rho)$, we obtain that $(u(\boldsymbol{x}, \rho), \phi(\boldsymbol{x}, \rho), v(\boldsymbol{x}, \rho)) \in \mathcal{E}_k$. Further, according to the definition of the mapping $\mathcal{J}$ in (10), we have

$$\mathcal{J}(u(\boldsymbol{x}, \rho), \phi(\boldsymbol{x}, \rho), v(\boldsymbol{x}, \rho)) = (\boldsymbol{x}, \rho).$$

Hence, $\mathcal{J}$ is surjective. □

With the above Lemma 3.1, we are now ready to present the relationship between the proposed model and the CtoD ART model.

Theorem 3.2. *The integral equation model* (5) *and the CtoD ART model* (9) *are equivalent.*

Proof. We prove the equivalence between the proposed integral equation model and the CtoD ART model. This is done by proving that the integrals of multiple variables for the two models are equal. To do so, we construct a bijective differentiable mapping between the two integration regions and show that with the mapping the two integrals are equal.

According to Lemma 3.1, the mapping $\mathcal{J}$ from the integration region $\mathcal{E}_k$ to the integration region $\mathcal{D}_k$ in (10) is bijective. It's easy to verify that $\mathcal{J}\colon \mathcal{E}_k \to \mathcal{D}_k$ is differentiable with

$$\mathcal{J}'(u,\phi,v) \coloneqq \frac{\partial(\boldsymbol{x},\rho)}{\partial(u,\phi,v)}.$$

Thus, the Jacobian determinant of $\mathcal{J}$ is given by

$$\det \mathcal{J}'(u,\phi,v) = \begin{vmatrix} \cos\phi & -u\sin\phi - v\cos\phi & -\sin\phi \\ \sin\phi & u\cos\phi - v\sin\phi & \cos\phi \\ 0 & -1 & 0 \end{vmatrix} = 1.$$

Therefore, we have

$$\mathrm{d}\rho\,\mathrm{d}\boldsymbol{x} = \det \mathcal{J}'(u,\phi,v)\,\mathrm{d}u\,\mathrm{d}\phi\,\mathrm{d}v = \mathrm{d}u\,\mathrm{d}\phi\,\mathrm{d}v.$$

Finally, we consider the integrands of two models. By the definitions of the attenuation kernel in (3) and (7), $\mathcal{K}_{\mathrm{a}}(\boldsymbol{x},\phi_k-\rho)$ and $\widetilde{\mathcal{K}}_{\mathrm{a}}(u,\phi)$ represent the attenuation on the lines $l(\boldsymbol{x},\phi)$ and $\tilde{l}(u,\phi)$, respectively. Since the lines $l(\boldsymbol{x},\phi)$ and $\tilde{l}(u,\phi)$ are the same line-of-response, thus, the attenuation kernels satisfy

$$\mathcal{K}_{\mathrm{a}}(\boldsymbol{x},\phi_k-\rho) = \widetilde{\mathcal{K}}_{\mathrm{a}}(u,\phi).$$

Further, the integrand $\mathcal{K}_{\mathrm{a}}(\boldsymbol{x},\phi_k-\rho)f(\boldsymbol{x})$ of the proposed model is equal to the integrand $\widetilde{\mathcal{K}}_{\mathrm{a}}(u,\phi)f(\boldsymbol{x})$ of the CtoD ART model. Thus, we have

$$\begin{aligned} g_k &= \frac{1}{\pi}\int_{\Omega_k}\int_{\rho_{\min,k}}^{\rho_{\max,k}} \mathcal{K}_{\mathrm{a}}(\boldsymbol{x},\phi_k-\rho)f(\boldsymbol{x})\,\mathrm{d}\rho\,\mathrm{d}\boldsymbol{x} \\ &= \frac{1}{\pi}\iiint_{\mathcal{S}_k}\int_{v_1(u)}^{v_2(u)} \widetilde{\mathcal{K}}_{\mathrm{a}}(u,\phi)f(\boldsymbol{x})\,\det \mathcal{J}'(u,\phi,v)\,\mathrm{d}v\,\mathrm{d}u\,\mathrm{d}\phi \\ &= \frac{1}{\pi}\iiint_{\mathcal{S}_k}\int_{v_1(u)}^{v_2(u)} \widetilde{\mathcal{K}}_{\mathrm{a}}(u,\phi)f(\boldsymbol{x})\,\mathrm{d}v\,\mathrm{d}u\,\mathrm{d}\phi. \end{aligned}$$

Hence, the proposed integral equation model and the CtoD ART model are equivalent. □

We conclude this section with a brief discussion of the shortcomings of the CtoD ART model that are addressed by the proposed model. The CtoD ART model itself is an accurate model for true coincidences. However, its discrete implementation introduces some limitations. The implementation of the CtoD ART model requires calculating the line integrals over a finite number of intersections of the LORs with the grids. In some parts of the image domain, the contribution weight of point sources to the photon counts on the detector pair may have large approximation errors. This is because the finite sampling of the LORs may introduce error to the regions of the grid that contribute to the photon counts. In the worst cases, some

regions of the grid may even be incorrectly treated as zero. For example, this can occur when point sources are at the smaller intersections of some grids with the SOR, or for counts in the grid near the edges of the SOR. To solve these problems, a sampling method that can adapt to all intersection cases between the LORs and the grids is necessary but at present does not exist. Increased sampling can reduce the approximation error, but this is computationally expensive. In addition, the entries of the system matrix cannot be given by a concise format, especially when the non-constant bases are used for the discretization. Finally, it is difficult to extend this model to include other physical effects, such as non-collinearity and Compton scatter. On the other hand, the proposed model (5) avoids the above-mentioned problems, allowing it to be implemented in a computationally efficient way. Importantly, the discretization of the proposed modeling framework can be accurately represented by basis functions, and models of the other physical effects can be easily added. The equivalence of the two models is shown via Theorem 3.2, which ensures that the proposed model is an equivalent and feasible surrogate to the CtoD ART model. To evaluate the performance of the proposed model, we compare the proposed integral equation model with an implementation using the CtoD ART model. We add that the implementation of the CtoD ART model is used to represent conventional discrete models. The details of the implementations are given in the numerical experiments section.

4. Discretization of Integral Equation Model

In this section, we propose a discretization for integral equation (5) by employing the piecewise linear polynomial collocation method [5, 6, 19, 22].

We start with choosing a basis that favors the expectation maximization (EM) algorithm, since the coefficients of the approximation of f in the basis will be a solution of a regularization problem to be solved by an EM-type algorithm. Generally, the EM-type algorithm will produce non-negative solutions at each iteration. Because the distribution function f is non-negative, it is desirable to choose non-negative basis functions. To this end, we propose a Lagrange interpolatory basis (LIB) to discretize integral equation (5). Since the digital images are often displayed in a rectangular shape, we need to extend the FOV to a rectangular image domain. For the sake of simplicity, we use the tangential square of the FOV as the extended image domain, denoted by Ω again, abusing the notation. The function value of f is zero in the complement of the FOV. We define a $p \times p$ square mesh for the extended domain Ω, with mesh-size h. For the ith grid $\omega_i \coloneqq [x_i, x_i + h] \times [y_i, y_i + h]$, we then choose four interpolation points $(x_i + mh/4, y_i + nh/4)$ for $m, n = 1, 3$. Then the corresponding LIB is given by

$$\begin{cases} \varphi_{i,1}(x,y) \coloneqq \quad (4x - 4x_i - 3h)(4y - 4y_i - 3h)/(4h^2), \\ \varphi_{i,2}(x,y) \coloneqq -\,(4x - 4x_i - 3h)(4y - 4y_i - \ \ h)/(4h^2), \\ \varphi_{i,3}(x,y) \coloneqq -\,(4x - 4x_i - \ \ h)(4y - 4y_i - 3h)/(4h^2), \\ \varphi_{i,4}(x,y) \coloneqq \quad (4x - 4x_i - \ \ h)(4y - 4y_i - \ \ h)/(4h^2), \end{cases} \quad \text{for} \quad (x,y) \in \omega_i,$$

and they all vanish on the outside of ω_i. Note that the value of each of these basis functions is 1 at one interpolation point and 0 at the other three interpolation points, and these functions are discontinuous with support ω_i. Using all basis

functions associated with the p^2 grids, we construct the space of trial functions as

$$\mathbb{X} \coloneqq \operatorname{span}\{\varphi_{i,j} \colon i = 1, 2, \ldots, p^2,\ j = 1, 2, 3, 4\}. \tag{11}$$

The interpolation of the tracer distribution function f from the space $\mathbb{X}$ is given by

$$\tilde{f} \coloneqq \sum_{i=1}^{p^2} \sum_{j=1}^{4} f_{i,j} \varphi_{i,j}, \tag{12}$$

where $f_{i,j}$ is the function value of f at the jth interpolation point in ω_i.

We discretize integral equation (1) by replacing the tracer distribution function f in the equation by its interpolation function $\tilde{f}$ from $\mathbb{X}$. We then obtain the discrete system of the proposed model:

$$g_k = \sum_{i=1}^{p^2} \sum_{j=1}^{4} f_{i,j} \int_{\omega_i \cap \Omega_k} \mathcal{W}_k(\boldsymbol{x}) \varphi_{i,j}(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}, \quad k = 1, 2, \ldots, N_d. \tag{13}$$

Equation (13) may be written in a matrix form. To this end, we define the column vectors $\bar{\mathbf{g}} \in \mathbb{R}_+^{N_d}$ by $\bar{\mathbf{g}} \coloneqq [g_k \colon k = 1, 2, \ldots, N_d]$, and $\mathbf{f} \in \mathbb{R}_+^{4p^2}$ with components $f_q \coloneqq f_{i,j}$, for $q = (j-1)p^2 + i$, $i = 1, 2, \ldots, p^2$, and $j = 1, 2, 3, 4$. The system matrix $\mathbf{A} \in \mathbb{R}^{N_d \times 4p^2}$ for the proposed model has the entries

$$a_{k,q} \coloneqq \int_{\omega_i \cap \Omega_k} \mathcal{W}_k(\boldsymbol{x}) \varphi_{i,j}(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}, \quad k = 1, 2, \ldots, N_d,\ q = 1, 2, \ldots, 4p^2. \tag{14}$$

With the notation defined above, equation (13) with positive background counts γ can be written in a compact form as

$$\bar{\mathbf{g}} = \mathbf{A}\mathbf{f} + \gamma. \tag{15}$$

Upon solving system (15) for $\mathbf{f}$ and substituting it into the right hand of (12), we obtain an approximation $\tilde{f}$ of the tracer distribution function f.

5. Regularized Maximum Likelihood Estimation and Reconstruction Algorithm

In this section, we describe a regularized maximum likelihood estimation for solving system (15) and a fixed-point proximity algorithm.

We first introduce the regularized maximum likelihood estimation for PET imaging derived from the linear system (15). In PET system, the detected coincidence counts vector is a random vector that follows the Poisson distribution. Hence, the measurements $\mathbf{g} \coloneqq [\mathbf{g}_k \colon k = 1, 2, \ldots, N_d]$ recorded at the N_d detector pairs is a random vector which relates to the vector $\bar{\mathbf{g}}$ via

$$\mathbf{g} = \operatorname{Poisson}(\bar{\mathbf{g}}), \tag{16}$$

where $\operatorname{Poisson}(\alpha)$ denotes a Poisson distributed random vector with mean α, cf., [13, 31]. Note that the tracer distribution function f may be constructed via the coefficient vector $\mathbf{f}$ which may be estimated from data $\bar{\mathbf{g}}$ via system (15) using the maximum likelihood estimation. Specifically, the regularized maximum likelihood estimate $\mathbf{f}_*$ is obtained by maximizing the conditional *a posteriori* probability

$\mathrm{p}(\mathbf{f}|\mathbf{g})$, the probability that $\mathbf{f}$ occurs when $\mathbf{g}$ is observed. This probability may be computed using the Bayes law:

$$\mathrm{p}(\mathbf{f}|\mathbf{g}) \propto \mathrm{p}(\mathbf{g}|\mathbf{f})\mathrm{p}(\mathbf{f}), \tag{17}$$

where $\alpha \propto \beta$ means that the scalar α is proportional to the scalar β. By taking the logarithm of both sides of equation (17), the estimate can be calculated by solving the optimization problem

$$\mathbf{f}_* = \arg\max \left\{ \ln \mathrm{p}(\mathbf{g}|\mathbf{f}) + \ln \mathrm{p}(\mathbf{f}) \colon \mathbf{f} \in \mathbb{R}_+^{4p^2} \right\}. \tag{18}$$

In the right-hand side of equation (18), the first term serves as a fidelity term, while the second term serves as a regularization term. Recalling the Poissonian nature of the detected data, we can calculate the likelihood function $\mathrm{p}(\mathbf{g}|\mathbf{f})$ as

$$\mathrm{p}(\mathbf{g}|\mathbf{f}) = \prod_{k=1}^{N_d} \frac{(g_k)^{\mathbf{g}_k} \exp(-g_k)}{\mathbf{g}_k!}.$$

Substituting (15) and (16) into the right-hand side of the equation above yields

$$\mathrm{p}(\mathbf{g}|\mathbf{f}) = \prod_{k=1}^{N_d} \frac{((\mathbf{A}\mathbf{f})_k + \gamma_k)^{\mathbf{g}_k} \exp(-(\mathbf{A}\mathbf{f})_k - \gamma_k)}{\mathbf{g}_k!}. \tag{19}$$

For the regularization term, applying the Gibbs priors [8, 15], we obtain that

$$\mathrm{p}(\mathbf{f}) \propto \exp(-\lambda \mathcal{R}(\mathbf{f})), \tag{20}$$

where λ is a positive regularization parameter and $\mathcal{R}(\mathbf{f})$ is a real-valued convex function defined on $\mathbb{R}_+^{4p^2}$ to be specified later. Substituting (19) and (20) into model (18), and converting the maximum problem to its minimum form, we then obtain the following minimization model:

$$\mathbf{f}_* = \arg\min \left\{ \langle \mathbf{A}\mathbf{f}, \mathbf{1} \rangle - \langle \ln(\mathbf{A}\mathbf{f} + \gamma), \mathbf{g} \rangle + \lambda \mathcal{R}(\mathbf{f}) \colon \mathbf{f} \in \mathbb{R}_+^{4p^2} \right\}. \tag{21}$$

We note that here '$\mathbf{1}$' denotes the vector with all components equal to 1. A more detailed derivation of the above model can be found in the references [12].

The regularization term is constructed by using the hypothesis that the tracer distribution function f is piecewise smooth. Unlike the discrete models where the traditional total variation (TV) regularization was used [12], for the Lagrange interpolatory basis described in the last section for the approximation of f, we need to design a regularizer that penalizes the discontinuity of the basis functions. Following the idea of [11], we propose a regularizer based on a multiscale representation [5, 6, 21, 22] of the tracer distribution function to penalize its discontinuity. For the $p \times p$ square mesh, and $i \in \{1, 2, \ldots, p^2\}$, let $S_i \coloneqq \{i, i+1, i+p, i+p+1\}$ denote the ith index set and let

$$\mathbb{X}_i \coloneqq \operatorname{span}\{\varphi_{l,j} \colon l \in S_i,\ j = 1, 2, 3, 4\}$$

denote the ith overlapping subspace of the trial function space $\mathbb{X}$. Clearly, the subspaces $\mathbb{X}_i$ satisfy $\mathbb{X} = \oplus_{i=1}^{p^2} \mathbb{X}_i$. We note that the periodic boundary condition is used for the definition of index set S_i when the ith grid is a boundary grid. We next construct a coarse scale subspace $\mathbb{Y}_i$ of the space $\mathbb{X}_i$. First, we construct the coarse grid $\chi_i \coloneqq \bigcup_{l \in S_i} \omega_l$ by taking the union of four neighboring fine grids. We

then scale the four basis functions $\varphi_{i,j}$, $j = 1, 2, 3, 4$, to form four basis functions $\varphi_{i,j}(2^{-1}x, 2^{-1}y)$, with support χ_i, for $j = 1, 2, 3, 4$. Thus, the coarse scale space $\mathbb{Y}_i$ of $\mathbb{X}_i$ can be constructed by

$$\mathbb{Y}_i \coloneqq \operatorname{span}\{\varphi_{i,j}(2^{-1}x, 2^{-1}y) \colon j = 1, 2, 3, 4\}.$$

Let $\mathbb{W}_i$ denote the orthogonal complement of $\mathbb{Y}_i$ in $\mathbb{X}_i$, that is, $\mathbb{X}_i = \mathbb{Y}_i \oplus^{\perp} \mathbb{W}_i$. We readily have that the dimension of $\mathbb{W}_i$ is $\dim(\mathbb{W}_i) = \dim(\mathbb{X}_i) - \dim(\mathbb{Y}_i) = 12$. We construct an orthonormal basis $\{\psi_{i,n} \colon n = 1, 2, \ldots, 12\}$ for $\mathbb{W}_i$ by using the basis functions of $\mathbb{X}_i$. For $n = 1, 2, \ldots, 12$, let

$$\psi_{i,n} \coloneqq \sum_{l \in S_i} \sum_{j=1}^{4} s_{l,j} \varphi_{l,j},$$

where $s_{l,j}$ are coefficients determined by the equations

$$\left\langle \psi_{i,n}(x, y), \varphi_{i,j}(2^{-1}x, 2^{-1}y) \right\rangle = 0, \quad j = 1, 2, 3, 4. \tag{22}$$

The above system is underdetermined. By choosing a group solution $\{s_{l,j}\}$ for 12 basis functions of $\mathbb{W}_i$, we are able to obtain specific $\psi_{i,n}$. Furthermore, applying the Gram-Schmidt process to orthonormalize the resulting basis functions, we obtain an orthonormal wavelet basis $\{\psi_{i,n} \colon n = 1, 2, \ldots, 12\}$ for $\mathbb{W}_i$.

Now, we return to the construction of the regularization term. Based on the orthogonality between the functions $\psi_{i,n}$ and $\varphi_{i,j}$, we use the inner products of the wavelet basis function and the trial function as the "derivative" of the tracer distribution function. Then we use the absolute sum of the inner products as a regularization to promote the sparsity of the derivative. This gives rise to the regularization term

$$\mathcal{R}(\mathbf{f}) \coloneqq \sum_{n=1}^{12} \sum_{i=1}^{p^2} \left| \langle \psi_{i,n}, \tilde{f} \rangle \right|. \tag{23}$$

This regularization term enables us to minimize the discontinuity and staircase artifacts of the reconstructed image, which is consistent with the piecewise smoothness nature of the tracer distribution function.

We next identify the regularization term as a composition of a convex function with a matrix. Such a form is convenient for us to apply the existing fixed-point proximity algorithm to solve the resulting nonsmooth minimization problem. For $n = 1, 2, \ldots, 12$, let Φ_n denote the l^1-norm on $\mathbb{R}^{p^2}$, that is,

$$\Phi_n(\mathbf{x}_n) \coloneqq \sum_{i=1}^{p^2} |x_{n,i}|, \ \mathbf{x}_n \in \mathbb{R}^{p^2}.$$

Then for $\mathbf{x} \coloneqq [\mathbf{x}_1^{\top}, \mathbf{x}_2^{\top}, \ldots, \mathbf{x}_{12}^{\top}]^{\top} \in \mathbb{R}^{12p^2}$, we define the function $\mathbf{\Phi}$ by

$$\mathbf{\Phi}(\mathbf{x}) \coloneqq \sum_{n=1}^{12} \Phi_n(\mathbf{x}_n). \tag{24}$$

The function $\mathbf{\Phi}$ defined by (24) is actually the l^1-norm of the vector in $\mathbb{R}^{12p^2}$, and thus, it is a convex function.

We next introduce a matrix $\mathbf{B}$ with 12 sub-matrices forming by the inner products of the 12 wavelet basis functions $\psi_{i,n}$ and the fine scale Lagrange interpolatory basis functions. Specifically, for each $n = 1, 2, \ldots, 12$, we define the sub-matrix $\mathbf{B}_n \in \mathbb{R}^{p^2 \times 4p^2}$ as

$$\mathbf{B}_n \coloneqq \begin{bmatrix} \mathbf{B}_{n,1} & \mathbf{B}_{n,2} & \mathbf{B}_{n,3} & \mathbf{B}_{n,4} \end{bmatrix}. \tag{25}$$

For $j = 1, 2, 3, 4$, the entries of $\mathbf{B}_{n,j}$ have the forms

$$(\mathbf{B}_{n,j})_{i,l} \coloneqq \langle \psi_{i,n}, \varphi_{l,j} \rangle, \quad i = 1, 2, \ldots, p^2, \ l = 1, 2, \ldots, p^2, \tag{26}$$

and let

$$\mathbf{B} \coloneqq \begin{bmatrix} \mathbf{B}_1^\top & \mathbf{B}_2^\top & \ldots & \mathbf{B}_{12}^\top \end{bmatrix}^\top.$$

Based on the above convex function $\mathbf{\Phi}$ and matrix $\mathbf{B}$, we have the following proposition for the regularization.

Proposition 5.1. *If the regularization $\mathcal{R}$ is defined by* (23)*, then for each $\mathbf{f} \in \mathbb{R}^{4p^2}$, $\mathcal{R}(\mathbf{f}) = \mathbf{\Phi}(\mathbf{B}\mathbf{f})$.*

Proof. We re-express the regularier (23) in a form of matrix-vector multiplication. Substituting the trail function $\tilde{f}$ in (12) into the right-hand side of (23), we obtain that

$$\mathcal{R}(\mathbf{f}) = \sum_{n=1}^{12} \sum_{i=1}^{p^2} \left| \sum_{l=1}^{p^2} \sum_{j=1}^{4} \langle \psi_{i,n}, \varphi_{l,j} \rangle f_{l,j} \right|. \tag{27}$$

According to the definition of the entries for matrix $\mathbf{B}_n$ in (26), equation (27) can be written as

$$\mathcal{R}(\mathbf{f}) = \sum_{n=1}^{12} \sum_{i=1}^{p^2} \left| \sum_{j=1}^{4} \sum_{l=1}^{p^2} (\mathbf{B}_{n,j})_{i,l} f_{i,j} \right|.$$

In the above equation, the sum $\sum_{j=1}^{4} \sum_{l=1}^{p^2} (\mathbf{B}_{n,j})_{i,l} f_{i,j}$ is the product of the ith row of $\mathbf{B}_n$ and the vector $\mathbf{f}$. Moreover, the sum $\sum_{i=1}^{p^2} |\sum_{j=1}^{4} \sum_{l=1}^{p^2} (\mathbf{B}_{n,j})_{i,l} f_{i,j}|$ can be written as the absolute sum of the vector $\mathbf{B}_n\mathbf{f}$, that is, $\Phi_n(\mathbf{B}_n\mathbf{f})$. Finally, the regularier $\mathcal{R}$ can be written as

$$\mathcal{R}(\mathbf{f}) = \sum_{n=1}^{12} \Phi_n(\mathbf{B}_n\mathbf{f}) = \mathbf{\Phi}(\mathbf{B}\mathbf{f}), \quad \text{for} \ \ \mathbf{f} \in \mathbb{R}^{4p^2},$$

proving the desired expression. □

We next describe the algorithm for solving the minimization problem (21). Proposition 5.1 allows us to use an existing fixed-point proximity algorithm [11, 12, 16, 20, 34], that is, the preconditioned proximity gradient algorithm (PPGA) to efficiently solve the minimization problem. To this end, we denote by $\mathbf{I}_p$ the $p \times p$ identity matrix, and for $\tau_n > 0, n = 1, 2, \ldots, 12$, define a parameter matrix $\mathbf{\Gamma}$ as

$$\mathbf{\Gamma} \coloneqq \operatorname{diag}(\tau_1 \mathbf{I}_{p^2}, \tau_2 \mathbf{I}_{p^2}, \ldots, \tau_{12} \mathbf{I}_{p^2}).$$

For $\mathbf{x} \in \mathbb{R}^{12p^2}$, we have that

$$\mathrm{prox}_{\mathbf{\Gamma}\mathbf{\Phi}^*}(\mathbf{\Gamma x}) \coloneqq \left[\mathrm{prox}_{\tau_1\Phi_1^*}(\tau_1\mathbf{x}_1), \mathrm{prox}_{\tau_2\Phi_2^*}(\tau_2\mathbf{x}_2), \ldots, \mathrm{prox}_{\tau_{12}\Phi_{12}^*}(\tau_{12}\mathbf{x}_{12})\right]^\top,$$

where $\mathbf{\Phi}^*$ is the conjugate function of $\mathbf{\Phi}$, and

$$\mathrm{prox}_{\tau_n\Phi_n^*}(\tau_n\mathbf{x}_n) \coloneqq \left[\mathrm{prox}_{\tau_n\Phi_n^*}(\tau_n x_{n,1}), \mathrm{prox}_{\tau_n\Phi_n^*}(\tau_n x_{n,2}), \ldots, \mathrm{prox}_{\tau_n\Phi_n^*}(\tau_n x_{n,p^2})\right]^\top,$$
$$\mathrm{prox}_{\tau_n\Phi_n^*}(\tau_n x_{n,i}) \coloneqq \min\{|\tau_n x_{n,i}|, 1\}\mathrm{sign}(\tau_n x_{n,i}).$$

The PPGA is shown as follows.

Algorithm 1 PPGA for PET image reconstruction

1: Given parameters: $\lambda > 0, \beta > 0$, $\tau_n > 0$ for $n = 1, 2, \ldots, 12$
2: Initialization: $\mathbf{f}^0 = \mathbf{f}, \mathbf{b}^0 = \mathbf{0}$
3: **for** $r = 0, 1, 2, \ldots$ **do**
4: $\quad \mathbf{b}^{r+1} = \mathrm{prox}_{\mathbf{\Gamma}\mathbf{\Phi}^*}(\mathbf{b}^r + \mathbf{\Gamma B f}^r)$
5: $\quad \Delta\mathbf{b}^{r+1} = 2\mathbf{b}^{r+1} - \mathbf{b}^r$
6: $\quad \mathbf{f}^{r+1} = \mathbf{f}^r - \dfrac{\beta}{\lambda}\mathbf{S}\mathbf{A}^\top\left(\mathbf{1} - \dfrac{\mathbf{g}}{\mathbf{A}\mathbf{f}^r + \gamma}\right) - \beta\mathbf{S}\mathbf{B}^\top\Delta\mathbf{b}^{r+1}$
7: **end for**

It is worthwhile pointed out that the PPGA can be accelerated via incremental update methods such as relaxed ordered subsets [30] and momentum techniques [17].

6. Convergence Analysis for PPGA

In this section, we establish convergence analysis for the PPGA algorithm. Specifically, we provide specific ranges of parameters appearing in the PPGA algorithm which ensure its convergence.

The following convergence result for Algorithm 1 was established in [12].

Theorem 6.1. *Suppose that* $\mathbf{S}$ *is a positive diagonal matrix and* $\mathbf{\Phi}$ *is a positive homogeneous convex function on* $\mathbb{R}^{12p^2}$*. If positive parameters* λ, β *and* τ_n*,* $n = 1, 2, \ldots, 12$*, are chosen to satisfy the conditions*

$$0 < \frac{\beta}{\lambda} \le \frac{(1-\varepsilon)\gamma^2}{2\|\mathbf{g}\|_\infty\|\mathbf{A}\|_2^2\|\mathbf{S}\|_2}, \quad 0 < \beta\tau_n \le \frac{1-\varepsilon}{2\|\mathbf{B}_n\|_2^2\|\mathbf{S}\|_2}, \tag{28}$$

for some $\varepsilon \in (0, 1)$*, then for any initial pair* $(\mathbf{f}^0, \mathbf{b}^0) \in \mathbb{R}_+^{4p^2} \times \mathbb{R}^{12p^2}$*, the sequence* $\{(\mathbf{f}^r, \mathbf{b}^r)\}$ *generated by Algorithm 1 converges to a solution of the minimization problem* (21)*.*

Note that the ranges (28) of the choices of parameters guaranteed convergence of Algorithm 1 involve $\|\mathbf{A}\|_2$ and $\|\mathbf{B}_n\|_2$. We next estimate these quantities to provide specific parameter range conditions for convergence of the PPGA algorithm.

It is nontrivial to compute $\|\mathbf{A}\|_2^2$. We turn to estimating its upper bound by estimating the l^1-norm and l^∞-norm of matrix $\mathbf{A}$. Let $\mathbf{c} \in \mathbb{R}_+^{4p^2}$, $\tilde{\mathbf{c}} \in \mathbb{R}_+^{N_d}$ denote

the positive vectors, whose components are given respectively by

$$c_q \coloneqq \max_{\boldsymbol{x}\in\omega_i}\left\{\frac{9}{4\pi}\int_0^{\pi}\mathcal{K}_{\mathrm{a}}(\boldsymbol{x},\rho)\,\mathrm{d}\rho\right\}, \qquad \text{for } \ q=1,2,\ldots,4p^2,$$

$$\tilde{c}_k \coloneqq \frac{9}{\pi}\int_{\Omega_k}\int_{\rho_{\min,k}}^{\rho_{\max,k}}\mathcal{K}_{\mathrm{a}}(\boldsymbol{x},\phi_k-\rho)\,\mathrm{d}\rho\,\mathrm{d}\boldsymbol{x}, \ \text{ for } \ k=1,2,\ldots,N_d.$$

We define the positive constant

$$C_a \coloneqq \max(\mathbf{c})\max(\tilde{\mathbf{c}}). \tag{29}$$

Lemma 6.2. *If the matrix* $\mathbf{A}$ *is defined by* (14)*, then* $\|\mathbf{A}\|_2^2 \le C_a h^2$.

Proof. We give an upper bound of $\|\mathbf{A}\|_2^2$ by estimating the upper bound of its l^1-norm and l^∞-norm. Because $0 \le |\varphi_{i,j}| \le 9/4$, together with (14), we obtain that

$$\sum_{k=1}^{N_d}|\tilde{a}_{k,q}| \le \frac{1}{\pi}\int_{\omega_i}\int_0^{\pi}\mathcal{K}_{\mathrm{a}}(\boldsymbol{x},\phi)|\varphi_{i,j}(\boldsymbol{x})|\,\mathrm{d}\rho\,\mathrm{d}\boldsymbol{x} \le \frac{9}{4\pi}\int_{\omega_i}\int_0^{\pi}\mathcal{K}_{\mathrm{a}}(\boldsymbol{x},\phi)\,\mathrm{d}\rho\,\mathrm{d}\boldsymbol{x} \le c_q h^2,$$

which yields that

$$\|\mathbf{A}\|_1 = \max_{1\le q\le 4p^2}\sum_{k=1}^{N_d}|\tilde{a}_{k,q}| \le h^2\max(\mathbf{c}). \tag{30}$$

Moreover, $\sum_{i=1}^{p^2}\sum_{j=1}^{4}|\varphi_{i,j}(\boldsymbol{x})| \le 9$, we then obtain that

$$\sum_{q=1}^{4p^2}|a_{k,q}| \le \frac{1}{\pi}\int_{\Omega_k}\int_{\rho_{\min,k}}^{\rho_{\max,k}}\mathcal{K}_{\mathrm{a}}(\boldsymbol{x},\phi_k-\rho)\,\mathrm{d}\rho\left(\sum_{i=1}^{p^2}\sum_{j=1}^{4}|\varphi_{i,j}(\boldsymbol{x})|\right)\mathrm{d}\boldsymbol{x} \le \tilde{c}_k,$$

which leads to

$$\|\mathbf{A}\|_\infty = \max_{1\le k\le N_d}\sum_{q=1}^{4p^2}|a_{k,q}| \le \max(\tilde{\mathbf{c}}). \tag{31}$$

Finally, according to estimates (31) and (30), we conclude that

$$\|\mathbf{A}\|_2^2 \le \|\mathbf{A}\|_1\|\mathbf{A}\|_\infty \le \max(\mathbf{c})\max(\tilde{\mathbf{c}})h^2 = C_a h^2,$$

proving the desired result. □

We now calculate the l^2-norm of matrices $\mathbf{B}_n$, $n=1,2,\ldots,12$. Due to the large size of these matrices, it may be a challenge to calculate their norm directly. Since $\|\mathbf{B}_n\|_2^2$ is equal to the largest eigenvalue of $\mathbf{B}_n^\top\mathbf{B}_n$, and the eigenvalues of $\mathbf{B}_n^\top\mathbf{B}_n$ are not easy to obtain, we shall construct a similar matrix of $\mathbf{B}_n^\top\mathbf{B}_n$, whose eigenvalues are easy to calculate. Thus, by calculating the largest eigenvalue of the similar matrix, we can obtain $\|\mathbf{B}_n\|_2^2$.

Before the construction of the similar matrix for $\mathbf{B}_n^\top\mathbf{B}_n$, we analyze the structure of $\mathbf{B}_n$, and then construct its concise form by using the Kronecker product of some simple matrices. According to the definition (26) of the entries for $\mathbf{B}_{n,j}$, the matrices $\mathbf{B}_{n,j}$ have the same structural characteristics for all $n=1,2,\ldots,12$ and $j=1,2,3,4$. Without loss of generality, we only give the construction procedure for one of $\mathbf{B}_{n,j}$ and others can be constructed in the same way. Considering the wavelet functions $\psi_{i,n}$ supported in χ_i, and the basis functions $\varphi_{l,j}$ supported in

ω_l, we have that for $i = 1, 2, \ldots, p^2$, the inner product $\langle \psi_{i,n}, \varphi_{l,j} \rangle = 0$, if $l \notin S_i$. Thus, we only need to consider the non-zero entries of $\mathbf{B}_{n,j}$, which are the entries with $l = i, i+1, i+p, i+p+1$. Moreover, since the shift invariance of the wavelet function $\psi_{i,n}$ and the basis function $\varphi_{l,j}$ on different grid, we can readily verify that the value of the inner product $\langle \psi_{i,n}, \varphi_{l,j} \rangle$ is independent of the index i, but depends on the relation between the indices i and l. Thereby, let $\otimes$ denotes the Kronecker product, the matrix $\mathbf{B}_{n,j}$ can be written as a sparse block matrix with two kinds of non-zero block matrices $\mathbf{Z}_{n,j,1}$ and $\mathbf{Z}_{n,j,2}$, that is,

$$\mathbf{B}_{n,j} = \mathbf{I}_p \otimes \mathbf{Z}_{n,j,1} + \mathbf{E}_p \otimes \mathbf{Z}_{n,j,2}, \tag{32}$$

where $\mathbf{E}_p$ denotes a $p \times p$ circulant matrix and its specific form is given by

$$\mathbf{E}_p \coloneqq \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ 1 & & & 0 \end{bmatrix},$$

and the two non-zero block matrices are $\mathbf{Z}_{n,j,1} \coloneqq \langle \psi_{i,n}, \varphi_{i,j} \rangle \mathbf{I}_p + \langle \psi_{i,n}, \varphi_{i+1,j} \rangle \mathbf{E}_p$ and $\mathbf{Z}_{n,j,2} \coloneqq \langle \psi_{i,n}, \varphi_{i+p,j} \rangle \mathbf{I}_p + \langle \psi_{i,n}, \varphi_{i+p+1,j} \rangle \mathbf{E}_p$. We now give a Kronecker product decomposition for $\mathbf{B}_{n,j}$. Let the indices k_1, k_2 defined as

(33)
$$k_1 \coloneqq \begin{cases} \mathrm{mod}(n+1, 2) + 1, & 1 \le n \le 4, \\ \mathrm{mod}(n+1, 2) + 3, & 5 \le n \le 12, \end{cases} \quad k_2 \coloneqq \begin{cases} \lfloor (n+1)/2 \rfloor + 2, & 1 \le n \le 4, \\ \lfloor (n+1)/2 \rfloor - 2, & 5 \le n \le 12, \end{cases}$$

where $\mathrm{mod}(\alpha, \beta)$ denotes the modulo operation that returns the remainder of α divided by β, $\lfloor \alpha \rfloor$ denotes the floor function that give the greatest integer less than or equal to α. Based on the specific wavelet functions, we construct the following eight circulant matrices $\mathbf{H}_{k,r}$ by

$$\mathbf{H}_{k,r} \coloneqq a_{k,r} \mathbf{I}_p + b_{k,r} \mathbf{E}_p, \quad \text{for} \quad k = 1, 2, 3, 4, \ r = 1, 2, \tag{34}$$

to satisfy the conditions

$$\begin{aligned} a_{k_1,1} a_{k_2,1} &= \langle \psi_{i,n}, \varphi_{i,j} \rangle, & a_{k_1,2} b_{k_2,1} &= \langle \psi_{i,n}, \varphi_{i+1,j} \rangle, \\ b_{k_1,1} a_{k_2,2} &= \langle \psi_{i,n}, \varphi_{i+p,j} \rangle, & b_{k_1,2} b_{k_2,2} &= \langle \psi_{i,n}, \varphi_{i+p+1,j} \rangle. \end{aligned}$$

Specifically, these non-zero entries $a_{k,r}$ and $b_{k,r}$ are given by

$$\begin{array}{llllllll} a_{1,1} = & \frac{2\sqrt{h}}{\sqrt{21}}, & a_{2,1} = -\frac{\sqrt{h}}{8\sqrt{7}}, & a_{3,1} = & \frac{5\sqrt{h}}{4\sqrt{39}}, & a_{4,1} = & \frac{7\sqrt{h}}{8\sqrt{13}}, \\ b_{1,1} = & \frac{\sqrt{h}}{2\sqrt{21}}, & b_{2,1} = & \frac{5\sqrt{h}}{8\sqrt{7}}, & b_{3,1} = & \frac{7\sqrt{h}}{4\sqrt{39}}, & b_{4,1} = -\frac{11\sqrt{h}}{8\sqrt{13}}, \\ a_{1,2} = & \frac{\sqrt{h}}{\sqrt{21}}, & a_{2,2} = & \frac{3\sqrt{h}}{8\sqrt{7}}, & a_{3,2} = -\frac{11\sqrt{h}}{4\sqrt{39}}, & a_{4,2} = -\frac{5\sqrt{h}}{8\sqrt{13}}, \\ b_{1,2} = -\frac{\sqrt{h}}{2\sqrt{21}}, & b_{2,2} = & \frac{9\sqrt{h}}{8\sqrt{7}}, & b_{3,2} = -\frac{\sqrt{h}}{4\sqrt{39}}, & b_{4,2} = & \frac{9\sqrt{h}}{8\sqrt{13}}, \end{array} \tag{35}$$

where h is the grid size defined in section 4. With the above eight matrices $\mathbf{H}_{k,r}$ in (34) and the relation between the index n and indices k_1, k_2 in (33), the matrices $\mathbf{B}_{n,j}$ in (32) can be expressed in the form of the Kronecker product as follows, for $n = 1, 2, \ldots, 12$ and $j = 1, 2, 3, 4$,

$$\mathbf{B}_{n,j} = \mathbf{H}_{k_1,r_1} \otimes \mathbf{H}_{k_2,r_2}, \ r_1 \coloneqq \mathrm{mod}(j-1, 2) + 1, r_2 \coloneqq \lfloor (j-1)/2 \rfloor + 1. \tag{36}$$

Finally, according to (25) and (36), the matrices $\mathbf{B}_n$ can be written as

$$\mathbf{B}_n = \begin{bmatrix} \mathbf{H}_{k_1,1} \otimes \mathbf{H}_{k_2,1} & \mathbf{H}_{k_1,2} \otimes \mathbf{H}_{k_2,1} & \mathbf{H}_{k_1,1} \otimes \mathbf{H}_{k_2,2} & \mathbf{H}_{k_1,2} \otimes \mathbf{H}_{k_2,2} \end{bmatrix}. \tag{37}$$

With the above concise form for matrices $\mathbf{B}_n$ for $n = 1, 2, \ldots, 12$, we construct a similar matrix for $\mathbf{B}_n^\top \mathbf{B}_n$. To facilitate the construction, we define the matrices $\mathbf{V}_k$ by

$$\mathbf{V}_k \coloneqq \begin{bmatrix} \mathbf{H}_{k,1} & \mathbf{H}_{k,2} \end{bmatrix}, \quad \text{for} \quad k = 1, 2, 3, 4. \tag{38}$$

We then have the following lemma for $\mathbf{B}_n^\top \mathbf{B}_n$ and its similar matrix.

Lemma 6.3. *If matrix $\mathbf{B}_n$ can be written as the form in* (37)*, then the matrix $\mathbf{B}_n^\top \mathbf{B}_n$ is similar to $\mathbf{V}_{k_2}^\top \mathbf{V}_{k_2} \otimes \mathbf{V}_{k_1}^\top \mathbf{V}_{k_1}$, for all $n = 1, 2, \ldots, 12$ and its corresponding index pair (k_1, k_2) defined by* (33)*.*

Proof. For any $n = 1, 2, \ldots, 12$, and the index pair (k_1, k_2), based on the permutation equivalence of the Kronecker product [27], there exists the permutation matrices $\mathbf{P} \in \mathbb{R}^{p^2 \times p^2}$ and $\mathbf{Q} \in \mathbb{R}^{2p^2 \times 2p^2}$ such that

$$\mathbf{V}_{k_1} \otimes \mathbf{H}_{k_2,1} = \mathbf{P}^\top \left(\mathbf{H}_{k_2,1} \otimes \mathbf{V}_{k_1} \right) \mathbf{Q}.$$

Since the matrix $\mathbf{H}_{k_2,2}$ has the same structure with $\mathbf{H}_{k_2,1}$, the permutation matrices $\mathbf{P}$ and $\mathbf{Q}$ for $\mathbf{V}_{k_1} \otimes \mathbf{H}_{k_2,2}$ are the same as these of $\mathbf{V}_{k_1} \otimes \mathbf{H}_{k_2,1}$. Thus, we have

$$\mathbf{V}_{k_1} \otimes \mathbf{H}_{k_2,2} = \mathbf{P}^\top \left(\mathbf{H}_{k_2,2} \otimes \mathbf{V}_{k_1} \right) \mathbf{Q}.$$

Let $\widetilde{\mathbf{Q}} \coloneqq \operatorname{diag}(\mathbf{Q}, \mathbf{Q}) \in \mathbb{R}^{4p^2 \times 4p^2}$. According to (37), and together with the above two equations, we have

$$\begin{aligned} \mathbf{B}_n &= \begin{bmatrix} \mathbf{V}_{k_1} \otimes \mathbf{H}_{k_2,1} & \mathbf{V}_{k_1} \otimes \mathbf{H}_{k_2,2} \end{bmatrix} \\ &= \mathbf{P}^\top \begin{bmatrix} \mathbf{H}_{k_2,1} \otimes \mathbf{V}_{k_1} & \mathbf{H}_{k_2,2} \otimes \mathbf{V}_{k_1} \end{bmatrix} \begin{bmatrix} \mathbf{Q} & \\ & \mathbf{Q} \end{bmatrix} \\ &= \mathbf{P}^\top \left(\mathbf{V}_{k_2} \otimes \mathbf{V}_{k_1} \right) \widetilde{\mathbf{Q}}. \end{aligned}$$

The permutation matrix $\mathbf{P}$ satisfy $\mathbf{P}\mathbf{P}^\top = \mathbf{I}_{p^2}$. Further, we obtain that

$$\begin{aligned} \mathbf{B}_n^\top \mathbf{B}_n &= \widetilde{\mathbf{Q}}^\top \left(\mathbf{V}_{k_2}^\top \otimes \mathbf{V}_{k_1}^\top \right) \left(\mathbf{V}_{k_2} \otimes \mathbf{V}_{k_1} \right) \widetilde{\mathbf{Q}} \\ &= \widetilde{\mathbf{Q}}^\top \left(\mathbf{V}_{k_2}^\top \mathbf{V}_{k_2} \otimes \mathbf{V}_{k_1}^\top \mathbf{V}_{k_1} \right) \widetilde{\mathbf{Q}}. \end{aligned}$$

It is easy to verify that $\widetilde{\mathbf{Q}}$ is a permutation matrix. Thus, we have that matrix $\mathbf{B}_n^\top \mathbf{B}_n$ is similar to $\mathbf{V}_{k_2}^\top \mathbf{V}_{k_2} \otimes \mathbf{V}_{k_1}^\top \mathbf{V}_{k_1}$. □

According to the above Lemma 6.3, $\mathbf{B}_n^\top \mathbf{B}_n$ is similar to $\mathbf{V}_{k_2}^\top \mathbf{V}_{k_2} \otimes \mathbf{V}_{k_1}^\top \mathbf{V}_{k_1}$, thus, these two matrices have the same eigenvalues. We shall calculate $\|\mathbf{B}_n\|_2^2$ by calculating the largest eigenvalue of $\mathbf{V}_{k_2}^\top \mathbf{V}_{k_2} \otimes \mathbf{V}_{k_1}^\top \mathbf{V}_{k_1}$. Based on the spectrum property of the Kronecker product [14], the eigenvalues of $\mathbf{V}_{k_2}^\top \mathbf{V}_{k_2} \otimes \mathbf{V}_{k_1}^\top \mathbf{V}_{k_1}$ are the product of the eigenvalues of $\mathbf{V}_{k_2}^\top \mathbf{V}_{k_2}$ and $\mathbf{V}_{k_1}^\top \mathbf{V}_{k_1}$. Therefore, the calculation of $\|\mathbf{B}_n\|_2^2$ turns into calculating the eigenvalues of matrices $\mathbf{V}_k^\top \mathbf{V}_k$. We give the following lemma regarding the eigenvalues of $\mathbf{V}_k^\top \mathbf{V}_k$ for $k = 1, 2, 3, 4$.

Lemma 6.4. *If the matrix* $\mathbf{V}_k$ *is defined by* (38) *for* $k = 1, 2, 3, 4$*, then the non-zero eigenvalues of* $\mathbf{V}_k^\top \mathbf{V}_k$ *are given by for* $l = 1, 2, \ldots, p$*,*

$$\delta_{k,l} \coloneqq (a_{k,1}^2 + b_{k,1}^2 + a_{k,2}^2 + b_{k,2}^2) + 2(a_{k,1}b_{k,1} + a_{k,2}b_{k,2}) \cos \frac{2l\pi}{p}. \tag{39}$$

Proof. Since $\mathbf{V}_k^\top \mathbf{V}_k$ has the same non-zero eigenvalues as $\mathbf{V}_k \mathbf{V}_k^\top$, it suffices to calculate eigenvalues of $\mathbf{V}_k \mathbf{V}_k^\top$. By the definition (38) for $\mathbf{V}_k$, and the definition (34) for $\mathbf{H}_{k,r}$, we have that

$$\begin{aligned} \mathbf{V}_k \mathbf{V}_k^\top &= \mathbf{H}_{k,1}\mathbf{H}_{k,1}^\top + \mathbf{H}_{k,2}\mathbf{H}_{k,2}^\top \\ &= (a_{k,1}^2 + b_{k,1}^2 + a_{k,2}^2 + b_{k,2}^2)\mathbf{I}_p + (a_{k,1}b_{k,1} + a_{k,2}b_{k,2})(\mathbf{E}_p + \mathbf{E}_p^\top). \end{aligned} \tag{40}$$

For eigenvalues of the above matrix, we have a somewhat more general result. If a matrix can be expressed as $a\mathbf{I}_p + b(\mathbf{E}_p + \mathbf{E}_p^\top)$, for $a, b \in \mathbb{R}$, then its eigenvalues are given by

$$\delta_l \coloneqq a + 2b \cos \frac{2l\pi}{p}, \quad \text{for} \quad l = 1, 2, \ldots, p.$$

In fact, by selecting the eigenvector $\mathbf{z} \in \mathbb{R}^p$ with components $z_k \coloneqq \sin(k - 1/2)\theta$, and θ satisfy $\cos p\theta = 1$, we can directly verify that

$$(a\mathbf{I}_p + b(\mathbf{E}_p + \mathbf{E}_p^\top))\mathbf{z} = (a + 2b\cos\theta)\mathbf{z}.$$

For $l = 1, 2, \ldots, p$, we choose $\theta_l = 2l\pi/p$ to satisfy $\cos p\theta = 1$. This gives the desired general result. The result of this lemma follows directly from (40) and the above general result. □

We next present the l^2-norm of $\mathbf{B}_n$. For the sake of brevity, we define

$$\begin{aligned} \xi_{k,1} &\coloneqq (a_{k,1} - b_{k,1})^2 + (a_{k,2} - b_{k,2})^2, \\ \xi_{k,2} &\coloneqq (a_{k,1} + b_{k,1})^2 + (a_{k,2} + b_{k,2})^2. \end{aligned} \tag{41}$$

Lemma 6.5. *For* $n = 1, 2, \ldots, 12$*, if the matrix* $\mathbf{B}_n$ *is defined by* (37)*, then*

$$\|\mathbf{B}_n\|_2^2 = \max\{\xi_{k_1,1}, \xi_{k_1,2}\} \max\{\xi_{k_2,1}, \xi_{k_2,2}\}.$$

Proof. The matrices $\mathbf{B}_n^\top \mathbf{B}_n$ and $\mathbf{V}_{k_2}^\top \mathbf{V}_{k_2} \otimes \mathbf{V}_{k_1}^\top \mathbf{V}_{k_1}$ have the same eigenvalues. Based on the spectrum property of the Kronecker product [14], the eigenvalues of $\mathbf{V}_{k_2}^\top \mathbf{V}_{k_2} \otimes \mathbf{V}_{k_1}^\top \mathbf{V}_{k_1}$ are the product of the eigenvalues of $\mathbf{V}_{k_2}^\top \mathbf{V}_{k_2}$ and $\mathbf{V}_{k_1}^\top \mathbf{V}_{k_1}$. According to Lemma 6.4, the non-zero eigenvalues of $\mathbf{V}_k^\top \mathbf{V}_k$ are $\delta_{k,l}$ in (39), for $l = 1, 2, \ldots, p$. Hence, the non-zero eigenvalues of $\mathbf{V}_{k_2}^\top \mathbf{V}_{k_2} \otimes \mathbf{V}_{k_1}^\top \mathbf{V}_{k_1}$ are given by

$$\Lambda_{l_1,l_2} \coloneqq \delta_{k_1,l_1}\delta_{k_2,l_2}, \quad \text{for} \quad l_1 = 1, 2, \ldots, p,\ l_2 = 1, 2, \ldots, p.$$

Thus, the non-zero eigenvalues of $\mathbf{B}_n^\top \mathbf{B}_n$ are Λ_{l_1,l_2}.

We now calculate the largest eigenvalue of $\mathbf{B}_n^\top \mathbf{B}_n$, which is $\|\mathbf{B}_n\|_2^2$. In the definition (39) of $\delta_{k,l}$, $a_{k,1}, b_{k,1}, a_{k,2}, b_{k,2}$ are the constants that defined by (35), and $\cos(2l\pi/p) \in [-1, 1]$. For all $l = 1, 2, \ldots, p$, it is easy to verify that $\delta_{k,l}$ chooses its extreme value $\xi_{k,1}$ when $l = p/2$, and chooses $\xi_{k,2}$ when $l = p$. We remarked here that the size p of mesh is an even. Thus, the largest eigenvalue of $\mathbf{V}_k^\top \mathbf{V}_k$ is

$$\max_{1 \le l \le p} \delta_{k,l} = \max\{\xi_{k,1}, \xi_{k,2}\}.$$

Since the eigenvalues of matrices $\mathbf{V}_{k_1}^\top \mathbf{V}_{k_1}$ and $\mathbf{V}_{k_2}^\top \mathbf{V}_{k_2}$ are independent, the largest eigenvalue of matrix $\mathbf{B}_n^\top \mathbf{B}_n$ is given by

$$\max_{1\le l_1,l_2\le p} \Lambda_{l_1,l_2} = \max_{1\le l_1\le p} \delta_{k_1,l_1} \max_{1\le l_2\le p} \delta_{k_2,l_2} = \max\{\xi_{k_1,1},\xi_{k_1,2}\}\max\{\xi_{k_2,1},\xi_{k_2,2}\}.$$

We thus prove the desired result of this lemma. □

We finally consider the specific l^2-norm of $\mathbf{B}_n$. By substituting the specific nonzero values (35) of matrix $\mathbf{H}_{k,r}$ into Lemma 6.5, we then obtain the l^2-norm of 12 matrices as

$$\begin{array}{llll} \|\mathbf{B}_1\|_2^2 = h^2/7, & \|\mathbf{B}_2\|_2^2 = 15h^2/91, & \|\mathbf{B}_3\|_2^2 = 65h^2/336, \\ \|\mathbf{B}_4\|_2^2 = 25h^2/112, & \|\mathbf{B}_5\|_2^2 = h^2/7, & \|\mathbf{B}_6\|_2^2 = 65h^2/336, \\ \|\mathbf{B}_7\|_2^2 = 15h^2/91, & \|\mathbf{B}_8\|_2^2 = 25h^2/112, & \|\mathbf{B}_9\|_2^2 = 36h^2/169, \\ \|\mathbf{B}_{10}\|_2^2 = 15h^2/52, & \|\mathbf{B}_{11}\|_2^2 = 15h^2/52, & \|\mathbf{B}_{12}\|_2^2 = 25h^2/64. \end{array} \tag{42}$$

For $n = 1, 2, \ldots, 12$, we define constants $D_n := \|\mathbf{B}_n\|_2^2/h^2$. By using the formulas in (42) we observe that

$$\begin{array}{llll} D_1 = 1/7, & D_2 = 15/91, & D_3 = 65/336, & D_4 = 25/112, \\ D_5 = 1/7, & D_6 = 65/336, & D_7 = 15/91, & D_8 = 25/112, \\ D_9 = 36/169, & D_{10} = 15/52, & D_{11} = 15/52, & D_{12} = 25/64. \end{array} \tag{43}$$

We finally provide the choice of parameters that guarantees convergence of Algorithm 1.

Theorem 6.6. *Suppose that* $\mathbf{S}$ *is a positive diagonal matrix,* $\mathbf{\Phi}$ *is a positive homogeneous convex function on* $\mathbb{R}^{12p^2}$*,* C_a *is a constant defined by* (29) *and* D_n *are the constants defined by* (43)*. If positive parameters* λ, β *and* τ_n*,* $n = 1, 2, \ldots, 12$*, are chosen to satisfy the conditions*

$$0 < \frac{\beta}{\lambda} \le \frac{(1-\varepsilon)\gamma^2}{2C_a h^2 \|\mathbf{g}\|_\infty \|\mathbf{S}\|_2}, \quad 0 < \beta\tau_n \le \frac{1-\varepsilon}{2D_n h^2 \|\mathbf{S}\|_2}, \tag{44}$$

for some $\varepsilon \in (0,1)$*, then for any initial pair* $(\mathbf{f}^0, \mathbf{b}^0) \in \mathbb{R}_+^{4p^2} \times \mathbb{R}^{12p^2}$*, the sequence* $\{(\mathbf{f}^r, \mathbf{b}^r)\}$ *generated by Algorithm 1 converges to a solution of the minimization problem* (21)*.*

Proof. It suffices to prove that the conditions for the parameter λ, β and τ_n in (44) imply those in (28). Using the norm estimation in Lemma 6.2 for system matrix $\mathbf{A}$, we obtain that

$$\|\mathbf{g}\|_\infty \|\mathbf{A}\|_2^2 \|\mathbf{S}\|_2 \le C_a h^2 \|\mathbf{g}\|_\infty \|\mathbf{S}\|_2.$$

Thus, we naturally obtain the inequality

$$\frac{(1-\varepsilon)\gamma^2}{2C_a h^2 \|\mathbf{g}\|_\infty \|\mathbf{S}\|_2} \le \frac{(1-\varepsilon)\gamma^2}{2\|\mathbf{g}\|_\infty \|\mathbf{A}\|_2^2 \|\mathbf{S}\|_2}. \tag{45}$$

For the second inequality in (44), using the norms (42) of the matrices $\mathbf{B}_n$, we have that

$$\|\mathbf{B}_n\|_2^2 \|\mathbf{S}\|_2 = D_n h^2 \|\mathbf{S}\|_2, \quad n = 1, 2, \ldots, 12. \tag{46}$$

The desired result of this theorem is thus obtained from Theorem 6.1 with (45) and (46). □

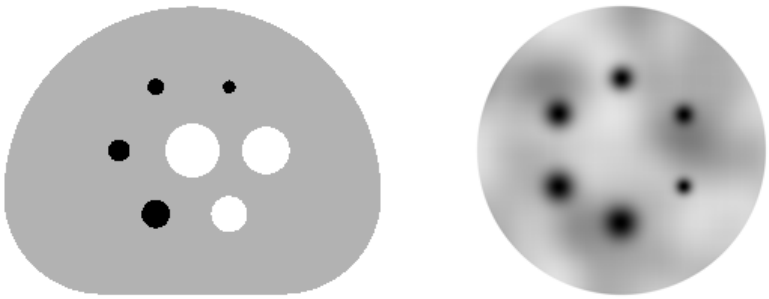
7. Numerical Experiments

We conduct in this section numerical experiments to evaluate the performance of the proposed integral equation model for PET image reconstruction. Comparisons are presented for both the projection and image domains for the proposed model with a conventional discrete model. In addition, we compare the reconstruction performance of the proposed model discretized by two different bases with their corresponding regularization to verify the advantages of piecewise linear discretization.

We first describe the simulation setup and experiment data. The simulation model used in this work is based on a 366.7 mm radius cylindrical detector ring with 576 detectors with widths of 4 mm. The FOV diameter is set to 300 mm, with 576 projection angles and 83 parallel projection bins, and the mesh of the reconstructed images is 256×256. Two phantoms are used in the experiments. The NEMA-IEC body phantom (Figure 2A) is mainly used for the projection comparisons due to its well-defined structure and obvious edges of its sinogram. Whereas, the smooth Lumpy phantom (Figure 2B) is used for the reconstruction experiments, which represents a more variable tracer distribution as might be seen in patient studies. The choice of the preconditioner $\mathbf{S}$ appearing in Algorithm 1 is given by

$$\mathbf{S} \coloneqq \operatorname{diag}\left(\mathbf{f}^{\bar{r}}/\mathbf{t}\right)$$

where $\mathbf{t} \coloneqq \max\{\mathbf{A}^\top \mathbf{1}, C_s h^2\}$, with C_s being an empirically selected positive constant, and $\mathbf{f}^{\bar{r}}$ is $\bar{r}$th iteration solution of $\mathbf{f}$. The division of two vectors $\mathbf{f}^{\bar{r}}$ and $\mathbf{t}$ in the above equation is the componentwise division of the vectors. When the iteration step $r+1$ is less than $\bar{r}$, $\mathbf{f}$ in $\mathbf{S}$ is chosen as $\mathbf{f}^r$; while $\mathbf{S}$ is fixed when the iteration step exceeds $\bar{r}$.



(A) (B)

FIGURE 2. (A) Piecewise constant NEMA-IEC phantom; (B) smooth Lumpy phantom.

The reference projection data for the true coincidences, without the positron range and the detector effects, were generated for the experiment; its implementation is based on a modified version of Radon transform [23]. An ultra-high resolution phantom (8192×8192) and a highly dense detector ring with 4608 detectors

of 0.5 mm width were used to obtain high accurate projection data. The resulting sinograms are much larger (4608×664) than those typically used. To obtain conventionally sized sinograms, each group of 8 adjacent detectors are combined to form 576 groups. The 64 individual projection angles between the detectors within each two groups are combined as well. Summing up all the counts on the projection angles between each two detector groups, we obtain a 576×83 reference sinogram.

To evaluate the performance of the proposed model, we use the following image quality metrics.

(1) Peak signal-to-noise-ratio (PSNR): PSNR of images is used to measure the image quality between two images. The higher the value of PSNR, the better the image quality. PSNR is defined as a ratio of the maximum possible power of a signal to the power of corrupting noise that affects representation fidelity

$$\mathrm{PSNR} := 10 * \lg \frac{N \max(\mathbf{f}_*)^2}{\|\mathbf{f} - \mathbf{f}_*\|_2^2},$$

where N is the pixel number of images.

(2) Structural similarity index measure (SSIM [36]): SSIM is a method for predicting the perceived quality of images. It gives a comparison of local patterns of pixel intensities that have been normalized for luminance and contrast. SSIM is used for measuring similarity between two images. The closer the value is to 1, the higher similarity between two images there is. Conversely, the closer the value is to 0, the lower similarity there is. The specific form of SSIM is given by [36]

$$\mathrm{SSIM} := \left(\frac{2\bar{x}\bar{y} + c_1}{\bar{x}^2 + \bar{y}^2 + c_1}\right)^{\alpha} \left(\frac{2\sigma_x\sigma_y + c_2}{\sigma_x^2 + \sigma_y^2 + c_2}\right)^{\beta} \left(\frac{\sigma_{xy} + c_3}{\sigma_x\sigma_y + c_3}\right)^{\gamma}.$$

where $\bar{x}$ and σ_x are the mean value and the variance of x, respectively, σ_{xy} is the covariance of x and y. In addition, c_1, c_2 and c_3 are three variables to stabilize the division with weak denominator, and α, β and γ are the weights for three fractions.

(3) Normalized mean square error (NMSE): NMSE is defined by

$$\mathrm{NMSE} := \|\mathbf{f} - \mathbf{f}_*\|_2^2 / \|\mathbf{f}_*\|_2^2,$$

where $\mathbf{f}$ and $\mathbf{f}_*$ denote the reconstructed image and original image, respectively.

(4) Normalized standard deviation (NSD): NSD used here is to measure the amount of variation for the residual sinogram and it is normalized by the mean value of the reference sinogram. The lower the value of NSD is, the lower the error of the residual sinogram has.

(5) Normalized mean (NMean): NMean is a measurement for the residual sinogram. It provides the amount of the residual and is normalized by the mean of the reference sinogram.

To examine the effect of Poisson noise on reconstruction results, the noisy projection data are used for the reconstruction. These noisy data are obtained by adding Poisson noise to the reference sinogram. We define the measurement of the noise

level by

$$\frac{\|(\mathbf{s}-\mathbf{s}_*)-\text{mean}(\mathbf{s}-\mathbf{s}_*)\|_2}{\sqrt{N}\text{mean}(\mathbf{s}_*)}\%,$$

where $\mathbf{s}$ and $\mathbf{s}_*$ denote the noisy sinogram and the noiseless sinogram, respectively, and N denotes the dimension of the sinograms.

We next turn to discussing the competing model for the proposed model. The competing model here is a conventional discretization of the CtoD ART model. The CtoD ART model is of infinite dimension. It requires a numerical method to solve it for a practical use. For each detector pair, there are an infinite number of line integrals on the corresponding LOR. The discretization of these line integrals is computationally expensive. Also, because this model is dependent on the line integral, the intersections of LOR and the basis functions may be difficult to represent. Instead, some conventional discrete methods mentioned in [10] were used. For example, previously uniformly spaced SORs were interpolated to redistribute the counts on a rectangular sampling grid [7, 9], or as an alternative example, a small number of approximate sampling parallel projections can be used [7, 18, 25, 28].

We use the conventional parallel projection discretization of the CtoD ART model as a competing model in this section. This method first simplifies the integral range $\mathcal{S}_k$, and then assumes that the attenuation for each LOR between detector pairs is the same. The resulting system is then discretized. We denote by $\widetilde{\mathcal{S}}_k \subset \mathcal{S}_k$ a finite subset with very few elements for the set $\mathcal{S}_k$, and by

$$\mathcal{L}_k \coloneqq \{\tilde{l}(u,\phi)\colon (u,\phi)\in\widetilde{\mathcal{S}}_k\}$$

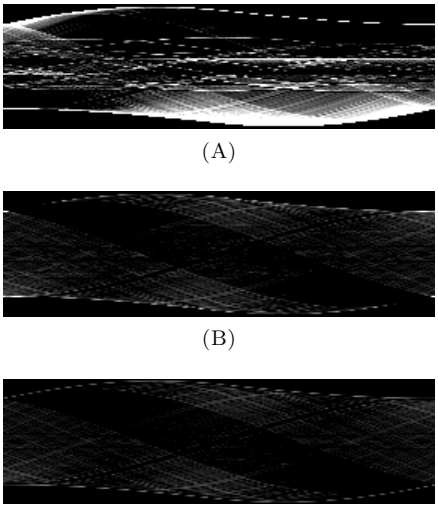
a LOR set for kth detector pair. Each $\tilde{l}(u,\phi)$ in $\mathcal{L}_k$ is parallel to the v_k-axis for kth SOR. We discretize the model (9) by using the collocation principle with the piecewise constant basis. With the mesh mentioned in the above discretization of the proposed model, for $i = 1, 2, \ldots, p^2$, we define f_i as the coefficient of the piecewise constant basis for the ith grid. Then a basic LOR intersection based DM derived from the parallel-beam sampling model is given by,

$$g_k = \widetilde{\mathcal{K}}_{\text{a}}(\bar{u}_k,\phi_k)\sum_{l_q\in\mathcal{L}_k}\sum_{i}^{p^2}\Delta l_{q,i}f_i,\ \text{ for }\ k=1,2,\ldots,N_d,$$

where $\Delta l_{q,i}$ is the intersection length of l_q and ith grid. The contribution weights of the image grid to a detector pair here are based on attenuation and intersection length of the LORs and grids. Further, for high accuracy and less computation, the intersection area of the SOR and grid is used as a part of the contribution weight instead of the intersection length. The improved SOR intersection-based DM [18, 25, 28] is now given by

$$g_k = \widetilde{\mathcal{K}}_{\text{a}}(\bar{u}_k,\phi_k)\sum_{i}^{p^2}\Delta s_{k,i}f_i,\ \text{ for } k=1,2,\ldots,N_d, \tag{47}$$

where $\Delta s_{k,i}$ is the intersection area of the kth SOR and the ith grid. Note that the LOR intersection-based DM is an approximation of the SOR intersection-based DM using the composite rectangle integration rule. For a description of these two



(A)

(B)

(C)

FIGURE 3. The residual sinograms of the three models: (A) Conventional discrete model; (B) Integral equation model with the piecewise constant basis; (C) Integral equation model with the piecewise linear basis for NEMA-IEC.

models, refer to the Figure 2 of [35]. The linear system for SOR intersection based model can be obtained in a manner similar to that of equations (13) to (15).

We now compare the proposed model with the competing discrete model in both the projection and image domains. The two proposed methods are system models discretized using either the piecewise constant basis (IE-PC), or the piecewise linear basis (IE-PL), leading to the discrete systems (13). The competing method is the conventional DM (C-DM) [18, 25, 28] for the CtoD ART model, resulting in the discrete system (47).

We first compare the performance of two proposed models with the conventional discrete model in the projection domain. We present the residuals between the sinogram of three models and the reference sinogram in Figure 3. These residuals show the agreement between the projection data of three models and the reference projection data. The smaller the residuals, the higher the agreement, illustrating the accuracy of the models. From the figure, we see that the two proposed models performed significantly better than the conventional DM. The proposed model with the piecewise linear discretization performed the best overall, showing the closest agreement with the reference, especially near edges. The major source of the improvements appear to be improved data fidelity at the edges and in textured regions. Further, we give quantitative comparison between the projections in Table 1. Compared to the conventional DM, NSD and NMean of the proposed model with the piecewise constant discretization are reduced by factors of 3.5 and 2.9, respectively. In addition, PSNR of the proposed method is 10.8 dB higher than that of the conventional DM. Discretization using the LIB further improves the

results with respect to the piecewise constant discretization. Compared with the conventional DM, NSD and NMean of the proposed model with the piecewise linear discretization are reduced by factors of 4.7 and 3.5, respectively, and PSNR is improved by 13.5 dB. In summary, the projection data generated by the proposed integral equation model are more consistent with those by the finely sampled Radon transform than those by the conventional DM.

TABLE 1. Agreement of the sinograms of three methods (C-DM, IE-PC and IE-PL) with the reference.

Method	PSNR	NSD	NMean
C-DM	43.65	1.07×10^{-2}	5.24×10^{-3}
IE-PC	54.45	3.09×10^{-3}	1.81×10^{-3}
IE-PL	**57.12**	$\mathbf{2.27 \times 10^{-3}}$	$\mathbf{1.51 \times 10^{-3}}$

Now, we compare quality of the reconstructed images generated by the proposed method and the conventional method for reconstructing the smooth Lumpy phantom. We have conducted experiments from projection data with five noise levels. In these experiments, we used the TV regularizer for both the conventional discrete model and the proposed model discretized with the piecewise constant basis, while used the wavelet-based regularizer for the proposed model with the piecewise linear discretization.

TABLE 2. Agreement of the reconstructed images (noise free) with the reference images.

Method	PSNR	SSIM	NMSE
C-DM	33.74	0.9580	4.09×10^{-3}
IE-PC	40.60	0.9902	8.44×10^{-4}
IE-PL	**53.79**	**0.9991**	$\mathbf{4.00 \times 10^{-5}}$

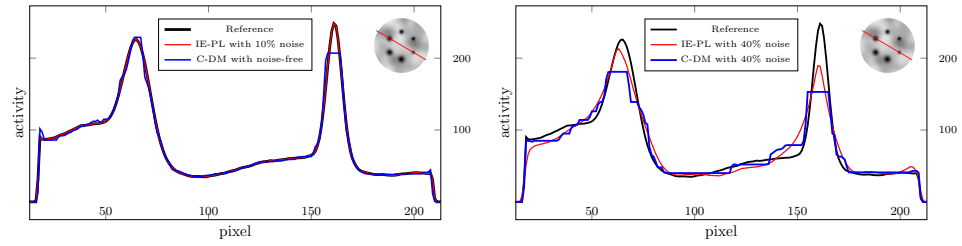
Table 2 shows PSNR, SSIM and NMSE of the images reconstructed by the three competing methods for the noise-free projection data. From the table we see that PSNR of the images reconstructed by the proposed model with the piecewise constant and linear discretization is respectively 7 dB and 20 dB higher than that by the conventional DM. These results reveal the substantial improvements of image quality of the integral equation method over the conventional DM.

We next compare the reconstruction results of the three methods for noisy data with noise levels 10%, 20%, 30%, and 40%, which correspond to 5274k, 1235k, 536k, and 320k photon counts, respectively. The numerical results are summarized in Table 3. We observe that the proposed integral equation method has remarkable advantages over the conventional discrete method in all cases. In particular, PSNR produced by the proposed method with the piecewise linear discretization improved by 5 dB and 1.5 dB for data with the 10% and 40% noise level, respectively, compared to that by the conventional discrete method. Furthermore, for data with 30% and 40% noise levels, the proposed method achieves the same reconstruction quality as the conventional DM for data with 10% and 20% noise levels, respectively.

TABLE 3. Agreement of the reconstructed images with the reference images in the presence of noise.

Noise Level	Method	PSNR	SSIM	NMSE
10% (5274k)	C-DM	30.35	0.8814	8.93×10^{-3}
	IE-PC	31.39	0.8892	7.04×10^{-3}
	IE-PL	**36.31**	**0.9750**	$\mathbf{2.27 \times 10^{-3}}$
20% (1235k)	C-DM	28.39	0.8476	1.40×10^{-2}
	IE-PC	28.97	0.8672	1.26×10^{-2}
	IE-PL	**31.91**	**0.9210**	$\mathbf{5.70 \times 10^{-3}}$
30% (536k)	C-DM	27.50	0.8329	1.72×10^{-2}
	IE-PC	27.61	0.8347	1.68×10^{-2}
	IE-PL	**29.73**	**0.9224**	$\mathbf{1.03 \times 10^{-2}}$
40% (320k)	C-DM	26.69	0.8085	2.08×10^{-2}
	IE-PC	26.69	0.8114	2.07×10^{-2}
	IE-PL	**28.14**	**0.8951**	$\mathbf{1.48 \times 10^{-2}}$

When comparing SSIM of images reconstructed by the three competing methods, the results of the proposed method for data with the 40% noise level are even better than those of the discrete method for data with 10% noise level. Such results show that even with higher noise in the projection data, the proposed method can still produce better reconstructed images than the conventional discrete method, which may help physicians and radiologists better identify tumors. This is very encouraging for the proposed method to be used for low-dose PET imaging due to the fact that SSIM is based on the human visual system.



(A) (B)

FIGURE 4. The radial line profile passing through the two hot spheres for the Lumpy phantom.

Finally, we show in Figure 4 line profiles of the images reconstructed by the proposed method and the conventional discrete method. Figure 4A shows that the reconstructed image of the proposed method for projection data with 10% Poisson noise is better than that of the conventional discrete method for the noise-free projection data, in terms of both the smoothness and the contrast of the hot spheres. We then show line profiles of the images reconstructed by the two methods using 40% noise projection data in Figure 4B. As shown in this figure, the line profile of

the discrete method presents an overall stepped shape, and poorer contrast recovery of the two hot spheres. Conversely, the overall smoothness of the reconstructed image and the contrast of the hot spheres by the proposed method are much better than those by the conventional discrete method. The figure shows that the image reconstructed by the proposed method can produce high consistency with the reference image. Whereas for the discrete method, non-negligible deviations appear in the contrast of the hot sphere and the boundary regions of the reconstructed image. We concluded that the proposed method can suppress noise, and may have the potential to improve the image quality in the low-dose PET image reconstruction.

8. Conclusions

In this work, we proposed an integral equation model for PET imaging. The proposed model allows the accurate modeling of the physical effects that would otherwise impose a fundamental bottleneck on the approximation accuracy of a PET system. Furthermore, the proposed model provides a general approach for accurately modeling the imaging process of PET that allows one to flexibly add additional physical effects into the model. We have shown that the proposed integral equation model is mathematically equivalent to the continuous-to-discrete ART model, an existing idealized model in terms of line integrals. Although the idealized model is accurate, it is not well suited for numerical approximation. The proposed integral equation model with the wavelet-based regularization produces encouraging numerical results, which demonstrate it significantly outperforms the conventional discrete model in all areas including approximation errors, noise suppression and reconstructed image quality.

Acknowledgments

X. Tang was supported in part by the Special Project on High-performance Computing under the National Key R&D Program (No. 2016YFB0200602), by the 2018NFGD under Grant 20186700031680001, by the Imaging and Radiation Sciences sub-award of the MSK Cancer Center Support Grant/Core Grant (P30 CA008748). C. Ross Schmidtlein was supported in part by the Imaging and Radiation Sciences sub-award of the MSK Cancer Center Support Grant/Core Grant (P30 CA008748). S. Li was supported in part by the Natural Science Foundation of China under grant 11601537, and 11571384, by the Science and Technology Program of Guangzhou under Grant 201804020053. Y. Xu was supported in part by US National Science Foundation under grant DMS-1912958, by Natural Science Foundation of China under grants 11771464.

References

[1] D. L. Bailey and J. Humm, Nuclear Medicine Physics: a Handbook for Teachers and Students, International Atomic Energy Agency, Vienna, Austria, 2014.

[2] A. Beer, Bestimmung der absorption des rothen lichts in farbigen flussigkeiten, Annalen der Physik, 162 (1852), pp. 78–88.

[3] B. Bendriem and D. W. Townsend, The Theory and Practice of 3D PET, vol. 32, Kluwer Academic Publishers, Dutch, 1998.

[4] M. Cecchetti, S. Moehrs, N. Belcari, and A. Del Guerra, Accurate and efficient modeling of the detector response in small animal multi-head PET systems, Physics in Medicine & Biology, 58 (2013), pp. 6713.

[5] Z. Chen, C. Micchelli, and Y. Xu, A construction of interpolating wavelets on invariant sets, Mathematics of Computation, 68 (1999), pp. 1569–1587.
[6] Z. Chen, C. A. Micchelli, and Y. Xu, Multiscale Methods for Fredholm Integral Equations, vol. 28, Cambridge University Press, Cambridge, UK, 2015.
[7] P. E. V. Dale L. Bailey, David W. Townsend and M. N. Maisey, Positron Emission Tomography: Basic Science, Springer, London, UK, 2003.
[8] S. Geman and D. Geman, Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images, IEEE Transactions on Pattern Analysis and Machine Intelligence, 6 (1984), pp. 721–741.
[9] Z. Hu, W. Wang, E. E. Gualtieri, Y. L. Hsieh, J. S. Karp, S. Matej, M. J. Parma, C. H. Tung, E. S. Walsh, M. Werner, et al., An LOR-based fully-3D PET image reconstruction using a blob-basis function, in IEEE Nuclear Science Symposium. Conference Record, vol. 6, Honolulu, USA, 2007, IEEE, pp. 4415–4418.
[10] A. Iriarte, R. Marabini, S. Matej, C. O. S. Sorzano, and R. M. Lewitt, System models for PET statistical iterative reconstruction: A review, Computerized Medical Imaging and Graphics, 48 (2016), pp. 30–48.
[11] Y. Jiang, S. Li, and Y. Xu, A higher-order polynomial method for SPECT reconstruction, IEEE Transactions on Medical Imaging, 38 (2019), pp. 1271–1283.
[12] A. Krol, S. Li, L. Shen, and Y. Xu, Preconditioned alternating projection algorithms for maximum a posteriori ECT reconstruction, Inverse Problems, 28 (2012), pp. 115005.
[13] K. Lange, R. Carson, et al., EM reconstruction algorithms for emission and transmission tomography, Journal of Computer Assisted Tomography, 8 (1984), pp. 306–16.
[14] A. Laub, Matrix Analysis for Scientists and Engineers, vol. 91, Siam, Philadelphia, USA, 2005.
[15] R. M. Leahy and J. Qi, Statistical approaches in quantitative positron emission tomography, Statistics and Computing, 10 (2000), pp. 147–165.
[16] S. Li, J. Zhang, A. Krol, C. R. Schmidtlein, L. Vogelsang, L. Shen, E. Lipson, D. Feiglin, and Y. Xu, Effective noise-suppressed and artifact-reduced reconstruction of SPECT data using a preconditioned alternating projection algorithm, Medical Physics, 42 (2015), pp. 4872–4887.
[17] Y. Lin, C. R. Schmidtlein, Q. Li, S. Li, and Y. Xu, A Krasnoselskii-Mann algorithm with an improved EM preconditioner for PET image reconstruction, IEEE Transactions on Medical Imaging, 38 (2019), pp. 2114–2126.
[18] A. Lougovski, F. Hofheinz, J. Maus, G. Schramm, E. Will, and J. Van den Hoff, A volume of intersection approach for on-the-fly system matrix calculation in 3D PET image reconstruction, Physics in Medicine & Biology, 59 (2014), pp. 561.
[19] Y. Lu, L. Shen, and Y. Xu, Integral equation models for image restoration: high accuracy methods and fast algorithms, Inverse Problems, 26 (2010), pp. 045006.
[20] C. A. Micchelli, L. Shen, and Y. Xu, Proximity algorithms for image models: denoising, Inverse Problems, 27 (2011), pp. 045009.
[21] C. A. Micchelli and Y. Xu, Using the matrix refinement equation for the construction of wavelets on invariant sets, Applied and Computational Harmonic Analysis, 1 (1994), pp. 391–401.
[22] C. A. Micchelli and Y. Xu, Reconstruction and decomposition algorithms for biorthogonal multiwavelets, Multidimensional Systems and Signal Processing, 8 (1997), pp. 31–69.
[23] F. Monard, RTG IPDE Summer School 2011 University of Washington-Seattle. https://sites.math.washington.edu/ipde/summer2011/index.html, June 2011.
[24] E. U. Mumcuoglu, R. M. Leahy, S. R. Cherry, and E. Hoffman, Accurate geometric and physical response modelling for statistical image reconstruction in high resolution PET, in IEEE Nuclear Science Symposium. Conference Record, vol. 3, Anaheim, USA, 1996, IEEE, pp. 1569–1573.
[25] J. M. Ollinger and A. S. Goggin, Maximum likelihood reconstruction in fully 3D PET via the SAGE algorithm, in IEEE Nuclear Science Symposium. Conference Record, vol. 3, Anaheim, USA, 1996, IEEE, pp. 1594–1598.
[26] J. Qi, R. M. Leahy, S. R. Cherry, A. Chatziioannou, and T. H. Farquhar, High-resolution 3D Bayesian image reconstruction using the microPET small-animal scanner, Physics in Medicine & Biology, 43 (1998), pp. 1001.

[27] D. J. Rose, Matrix identities of the fast Fourier transform, Linear algebra and its applications, 29 (1980), pp. 423–443.
[28] J. J. Scheins, F. Boschen, and H. Herzog, Analytical calculation of volumes-of-intersection for iterative, fully 3-D PET reconstruction, IEEE Transactions on Medical Imaging, 25 (2006), pp. 1363–1369.
[29] O. Scherzer, Handbook of Mathematical Methods in Imaging, Springer Science & Business Media, New York, USA, 2010.
[30] C. R. Schmidtlein, Y. Lin, S. Li, A. Krol, B. J. Beattie, J. L. Humm, and Y. Xu, Relaxed ordered subset preconditioned alternating projection algorithm for PET reconstruction with automated penalty weight selection, Medical Physics, 44 (2017), pp. 4083–4097.
[31] L. A. Shepp and Y. Vardi, Maximum likelihood reconstruction for emission tomography, IEEE Transactions on Medical Imaging, 1 (1982), pp. 113–122.
[32] A. Terstegge, S. Weber, H. Herzog, H. W. Muller-Gartner, and H. Halling, High resolution and better quantification by tube of response modelling in 3D PET reconstruction, in IEEE Nuclear Science Symposium. Conference Record, vol. 3, Anaheim, USA, 1996, IEEE, pp. 1603–1607.
[33] P. Toft, The Radon Transform - Theory and Implementation, phdthesis, Technical University of Denmark (DTU), Copenhagen, Denmark, Nov. 1996.
[34] Z. Wu, S. Li, X. Zeng, Y. Xu, and A. Krol, Reducing staircasing artifacts in SPECT reconstruction by an infimal convolution regularization, Journal of Computational Mathematics, 34 (2016), pp. 624–645.
[35] W. Zheng, S. Li, A. Krol, C. R. Schmidtlein, X. Zeng, and Y. Xu, Sparsity promoting regularization for effective noise suppression in SPECT image reconstruction, Inverse Problems, 35 (2019), pp. 115011.
[36] Zhou Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Transactions on Image Processing, 13 (2004), pp. 600–612.

[1] School of Computer Science and Engineering, Guangdong Province Key Lab of Computational Science, Sun Yat-sen University, Guangzhou 510275, China
E-mail: tangxinhuang@live.com

[2] Department of Medical Physics, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA
E-mail: schmidtr@mskcc.org
URL: https://www.mskcc.org/profile/charles-schmidtlein

[3] School of Computer Science and Technology, Guangdong University of Technology, Guangzhou 510006, China
E-mail: reesiloveu@163.com

[4] Department of Mathematics and Statistics, Old Dominion University, Norfolk, VA 23529, USA
E-mail: y1xu@odu.edu