

Convolutional Forward Models for X-Ray Computed Tomography*

Kai Zhang[†] and Alireza Entezari[†]

Abstract. This paper presents a framework for efficient and accurate computation of X-ray optics, a key ingredient in optimization-based computed tomography (CT) reconstruction algorithms. Based on an algebraic framework for directional convolution in image space and detector space, we construct forward models for X-ray imaging whose computational cost can be optimized for each specific CT geometry. While the framework allows for modeling various sources of blur in the X-ray imaging process for any CT geometry, we demonstrate and characterize its effectiveness in fan-beam and cone-beam geometries with flat detectors. The experiments show improvements in computational efficiency as well as accuracy, in optics calculations and reconstruction error, of the proposed projector compared to the state-of-the-art methods used in forward- and back-projection algorithms.

Key words. X-ray computed tomography, forward model, model-based iterative reconstruction

MSC codes. 68Q25, 68U10

DOI. 10.1137/21M1464191

1. Introduction. A wide range of applications in science and engineering use principles of computed tomography (CT) and tomographic reconstruction algorithms to form images from projection data [43]. By acquiring X-ray projections from a large, uniformly distributed, set of viewing angles, the filtered back-projection (FBP) algorithm provides a classical choice for image reconstruction. Remarkable technology developments, mainly in hardware, have made CT the modality of choice in a growing list of applications not only in diagnostic but also in image-guided radiation therapy and surgery [27, 29]. Besides the widespread applications in medicine, imaging techniques in biology, seismology, and astronomy use CT reconstruction algorithms.

Imaging from a low-dose X-ray source and/or limited projection views has been the holy grail of CT research for the past several decades. Techniques from applied mathematics, signal processing, and, more recently, machine learning have contributed to formation of modern iterative CT reconstruction algorithms. Model-based iterative reconstruction (MBIR) methods have shown great promise for enabling low-dose imaging by incorporating models of data acquisition and image priors into the CT reconstruction process [4, 17, 54]. Despite promising dose-reduction results from these iterative image reconstruction algorithms [39, 42], the FBP algorithm and its variants, which require high-dose X-ray data, are commonly used in commercial scanners [23, 45, 57]. One of the main impediments in translating MBIR algorithms

*Received by the editors September 26, 2022; accepted for publication (in revised form) April 14, 2023; published electronically October 12, 2023.

<https://doi.org/10.1137/21M1464191>

Funding: The work of the authors was partially supported by National Science Foundation grant CCF-2210866.

[†]Department of Computer and Information Science and Engineering, CISE, University of Florida, Gainesville, FL 32611 USA (kazhang@ufl.edu, entezari@ufl.edu).

to practice has been their computational cost (i.e., long reconstruction time) compared to analytical methods [25, 38, 46].

The success of deep learning in computer vision and image analysis problems has motivated numerous attempts for leveraging neural networks for image reconstruction [32, 54, 58, 60, 61, 66]. These methods use deep learning to improve the image quality by postprocessing of FBP reconstruction from low-dose data or preprocessing projection data in the sinogram domain (and combinations of post- and preprocessing). End-to-end approaches for leveraging deep learning (e.g., [66]) are impractical and difficult to train [35] since they are computationally expensive. Plug-n-play approaches [28, 61] use deep learning for data-driven image priors that are used in MBIR reconstruction. Despite the variety of approaches for reconstruction in CT, MBIR algorithms are considered as the gold standard [53] as robustness to noise, reliability, and image quality issues are well understood [39, 42].

Model-based iterative reconstruction. MBIR formulates the image reconstruction as an optimization problem:

$$(1.1) \quad \arg \min_{\mathbf{c} \in \mathbb{R}^N} \|\mathbf{Ac} - \mathbf{b}\|_Q^2 + \lambda \mathcal{R}(\mathbf{c}).$$

The system matrix $\mathbf{A}$ is a forward model of the continuous-domain X-ray optics (see section 2) on a discretized N -pixel/voxel attenuation coefficient vector $\mathbf{c}$, measured by detector cells, on various viewing angles, leading to the observations $\mathbf{b}$. Here $\mathcal{R}$ is a regularization term incorporating image priors (e.g., neural-network-based, total-variation, sparsity in a transform domain, or Markov random fields) into reconstruction. The parameter λ controls the trade-off between regularization and fidelity of the solution. The fidelity term above is usually a quadratic form to model the noise characteristics of the measurement system with $\|\mathbf{r}\|_Q^2 = \mathbf{r}^T \mathbf{Q} \mathbf{r}$, where the operator $\mathbf{Q}$ models the (co)variance in the noisy observations, $\mathbf{b}$, whose noise is often modeled by a (compound) Poisson distribution [33, 56]. The reconstruction process can also be formulated in a Bayesian framework, often with a penalized likelihood objective, where the solution provides a maximum a posteriori estimate given the observations [16, 20, 47, 52].

For practical imaging resolutions, the size of the system matrix $\mathbf{A}$ is prohibitively large to be stored in main memory and reused during iterations. In a common imaging setup [18, 49], the two-dimensional (2-D) fan-beam forward model requires about 1.6 GB per viewing angle for a moderate imaging resolution of $N = 512 \times 512$. This per-viewing-angle cost increases to 528 TB for a 3-D cone-beam forward model at the resolution of $N = 512 \times 512 \times 512$. Despite utilizing sparse representation, these memory requirements are already impractical for GPU or main memory even in few-view imaging [45] and larger resolutions of up to 4096 with thousands of views are of practical importance [48, 49]. For that reason, in practical applications iterative reconstruction requires the computation of the forward model $\mathbf{A}$ in each iteration for carrying out *forward-projection* $\mathbf{c} \mapsto \mathbf{Ac}$ and *back-projection* $\mathbf{r} \mapsto \mathbf{A}^T \mathbf{r}$. Conceptually, this on-the-fly computation is similar to the application of the fast Fourier transform (FFT) algorithm for computing the discrete Fourier transform (DFT) that avoids storage of the DFT matrix. However, unlike the Fourier transform that lends itself to a divide-and-conquer approach leveraged in the FFT algorithm, the X-ray transform has no known fast discrete transform—necessitating approximations [2, 5, 6, 12]. Earlier research focused on leveraging hardware acceleration (e.g., GPU, multicore) for forward- and back-projection,

which provided a set of useful acceleration tools [3, 41, 46, 55, 59]. However, the physical barriers in increasing clock frequency and the memory latency bottleneck in GPU/parallel computing impose limits in the power of hardware acceleration in addressing computational challenges in clinical-scale CT problems [36].

Forward models and image discretization. The simulation of the acquisition process mathematically models the relationship between the image (i.e., attenuation coefficient map) defined in the continuous domain and the (discrete) detector measurements. This process introduces (at least) two transformations to the image, namely projection of the image domain signal to the sinogram domain followed by the detector's sampling process. The projection step is often modeled by a linear approximation to the Lambert–Beer law, and the detector's sampling process is modeled by the integration of the sinogram domain signal across a detector cell that introduces a blur [62]. While the linearization of the Lambert–Beer law introduces an approximation error that causes various artifacts [11, 24] such as the exponential edge gradient effect [31], it underlies most projection/back-projection methods that are used in practice [38].

Leveraging this mathematical model of acquisition in forward- and back-projection steps of an optimization process requires discretization of the image by its expansion in a set of (basis) functions defined in the continuous domain. This expansion in a basis set is then used to build the forward model, $\mathbf{A}$, that relates the (discrete) coefficients of the image in the expansion, $\mathbf{c}$ (e.g., pixel values), to the (discrete) detector measurements, $\mathbf{b}$, in the sinogram domain $\mathbf{Ac} = \mathbf{b}$. Common choices, in CT, for basis functions are the pixel- and the voxel-basis as well as spherical extension of Kaiser–Bessel functions (aka blobs) that have been proposed for smoother image representations [34]. While the choice of pixel-/voxel-basis provides a partition of unity that guarantees first-order approximation to the underlying continuous-domain image, Kaiser–Bessel functions do not form a partition of unity and require filtering operations [44] to minimize discretization errors. A different class of basis functions, called box splines, have been proposed in [13] for image discretization in the context of CT reconstruction. This class includes the pixel-basis (in two dimensions) and voxel-basis (in three dimensions) as first-order box splines as well as tensor-product B-splines as special cases. From a signal processing perspective, an advantage of representing the image in a basis set is that continuous-domain linear transformations introduced in the imaging pipeline such as projection or blurring effects (e.g., detector blur or focal-spot or motion) can be modeled by corresponding linear transformations applied to the basis functions.

Contributions. In this paper, we demonstrate a continuous-domain framework for directional convolution that allows for incorporation of many linear transformations into the forward model with pixel- and voxel-basis functions (as well as general box splines) in a computationally efficient manner. While previous work has shown efficient X-ray projection with box splines that were limited to parallel geometries [13], the significance of the present framework lies in its flexibility to model divergent ray geometry, making it practical for common scanner geometries. Moreover, the algebraic framework of directional convolutions allows for modeling detector blur into the forward model in a computationally efficient manner. While directional convolutions allow for computationally efficient forward models for any CT geometry, we demonstrate their advantages for fan-beam and cone-beam geometries with flat and equispaced detectors.

Existing techniques for efficient computation of X-ray optics in forward- and back-projection model detector blur by postprocessing the sinogram domain signal. Earlier pixel-driven and ray-driven techniques used interpolation techniques to estimate detector measurements (forward-projection) or pixel corrections (back-projection) from signals sampled at detector or pixel centers, respectively [7, 30, 37, 51]. The distance-driven method uses the image/sinogram signals sampled at pixel/detector boundaries that are then projected to a common plane according to the CT geometry. Then the extent of the overlap between pixels and detectors in that plane is used for the efficient computation the forward model [10]. The drawback in the distance-driven strategy is that it approximates the projection of a pixel/voxel (aka footprint) as a constant function over its support. To improve that approximation, the separable footprints (SF) method [38] provides a separable approximation to footprints and has been demonstrated to be the most accurate among existing methods. A separable approximation to footprints of high order B-splines (as opposed to pixel/voxel) has also been considered in [40]. The quality of this approximation requires higher order B-splines (e.g., cubic) as the support of the tensor product concentrates over a somewhat more isotropic set. This limits the applicability to CT imaging since high order B-splines cannot accurately represent biomedical images that typically contain homogeneous regions with sharp boundaries. Their approximation power can only be realized (for accurate discretization) when the underlying function is smooth (i.e., C^2 for cubic B-splines). More recently a lookup-table-based integration (LTRI) approach was proposed [25] that precomputes the projection of pixels to detectors on a predetermined set of directions and provides speedups at the cost of further errors in approximating the optics.

Our experiments show the computational efficiency of the resulting projectors in comparison to state-of-the-art methods, namely SF and LTRI, and demonstrate that the computational efficiency is achieved without compromising the accuracy in modeling X-ray optics in divergent geometries with detector blur. The experiments also show improvements in the accuracy of forward-projection that also translate to observable improvements in reconstructed images.¹ We remark that our framework provides for exact computation of footprint (of pixel/voxel or higher order box splines) in parallel and divergent geometries as well as exact computation of blur in parallel geometries.

2. X-ray transform. As discussed before, forward models in iterative reconstruction techniques provide a mathematical model for simulating the acquisition process on an image by relating it to what is being measured in the projection domain. Forward models are designed for each particular CT geometry and their computation requires specification of X-ray source and detector geometries on a discretized image.

2.1. Geometry specification. We introduce symbols to describe fan-beam ($d = 2$) and cone-beam ($d = 3$) geometries with a common algebraic notation. To help with this notation, we include a glossary of symbols in Appendix A. Let $\mathbf{u} \in \mathbb{R}^d$ denote a viewing direction that is a unit vector from rotation center to source. We denote θ as the projection angle of source from

¹A preliminary version of this paper in fan-beam geometry was presented at the ISBI conference [65].

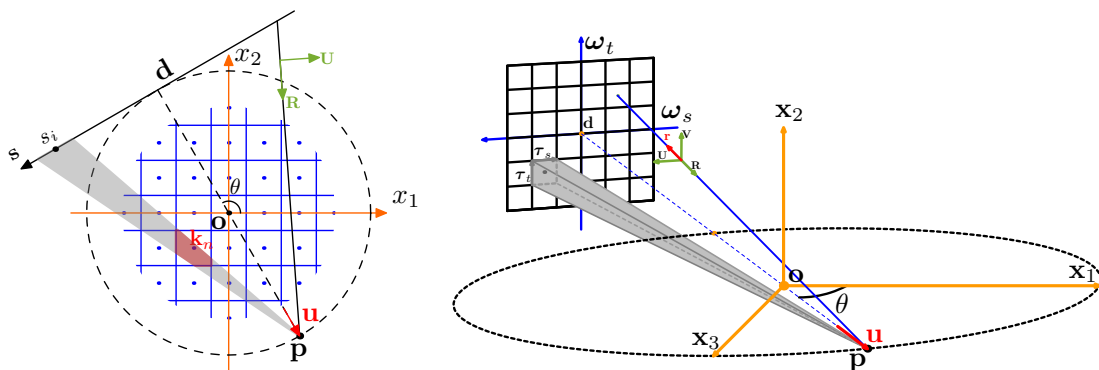


Figure 1. Fan- and cone-beam geometries with flat detectors.

the $\mathbf{x}$ -axis (positive counterclockwise), and the vector $\mathbf{u}$ is represented as $\mathbf{u} = [\cos \theta \quad \sin \theta]^T$ in two dimensions and $\mathbf{u} = [\cos \theta \quad 0 \quad \sin \theta]^T$ in three dimensions with T denoting the transpose. The point $\mathbf{p} = D_{so}\mathbf{u}$ rotating around a circle trajectory of radius D_{so} is the source.

The 1-D detector plane is represented by the local coordinate s in fan-beam geometry, as in Figure 1, left, and (s, t) are the local coordinates of the 2-D detector plane in cone-beam geometry as in Figure 1, right. We use the symbol $\boldsymbol{\omega}$ to denote the detector-space coordinates: $\boldsymbol{\omega} = s$ in fan-beam and $\boldsymbol{\omega} = [t \quad s]^T$ in cone-beam geometries. Together with $\mathbf{u}$, this coordinate system can be described by a view-transformation matrix in fan-beam and cone-beam geometries:

$$(2.1) \quad \mathcal{B}_p = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}, \quad \mathcal{B}_p = \begin{bmatrix} \cos \theta & 0 & -\sin \theta \\ 0 & 1 & 0 \\ \sin \theta & 0 & \cos \theta \end{bmatrix},$$

where in cone-beam geometry, the detector plane and the t -axis are parallel to the $\mathbf{x}_2$ -axis. Let $\mathbf{x} \in \mathbb{R}^d$ be the image domain coordinates, so it can be parameterized as $\mathbf{x} = \mathbf{p} + \lambda \mathbf{r}(\boldsymbol{\omega})$. The vector $\mathbf{r}(\boldsymbol{\omega})$ is a unit direction of each ray calculated by $\mathbf{r}(\boldsymbol{\omega}) = \text{unit}(\mathcal{B}_p[D_{sd} \quad \boldsymbol{\omega}]^T - \mathbf{p})$, where D_{sd} is the distance from source to detector plane, and $\text{unit}()$ returns a unit-length vector by normalizing its input.

2.2. Analytical model. When considering an ideal point source (i.e., no focal-spot effect), the d -dimensional fan-beam/cone-beam X-ray transform $\mathcal{P}$ maps $f(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^d$, into the set of its *line integrals* to form the projection:

$$(2.2) \quad \mathcal{P}_{\mathbf{u}}\{f\}(\boldsymbol{\omega}) = \int_0^\infty f(\mathbf{p} + \lambda \mathbf{r}(\boldsymbol{\omega})) d\lambda.$$

We refer to the above map as $\mathcal{P}_{\mathbf{u}}\{f\}$ for short.

In a simple model of forward-projection, one can do point-sampling on the projected function $\mathcal{P}_{\mathbf{u}}\{f\}$, whereas in more realistic modeling of the transform model the projections are integrated across a detector cell with a finite width. To represent the cell width in each dimension, we use a vector $\boldsymbol{\tau}$, in a local coordinate system in the detector plane. In fan-beam

geometry, $\boldsymbol{\tau} = [\tau_s]$ is a vector in $\mathbb{R}^1$, whereas in cone-beam, $\boldsymbol{\tau} = [\tau_s, \tau_t]$ is a join of two 2-D vectors describing the detector cell orientation and size. To model g_θ —the signal observed by the detector in the sinogram domain (i.e., detector space) at viewing angle θ with viewing direction $\mathbf{u}$ —(2.2) can be extended to model detector blur [62] as

$$(2.3) \quad \begin{aligned} g_\theta(\boldsymbol{\omega}) &= \mathcal{P}_{\mathbf{u}, \boldsymbol{\tau}}\{f\}(\boldsymbol{\omega}) = \int h_\tau(\boldsymbol{\omega} - \boldsymbol{\omega}') \mathcal{P}_{\mathbf{u}}\{f\}(\boldsymbol{\omega}') d\boldsymbol{\omega}' \\ &= (\mathcal{P}_{\mathbf{u}}\{f\} * h_\tau)(\boldsymbol{\omega}). \end{aligned}$$

For a detector located at $\boldsymbol{\omega}_{\mathbf{I}}$, where detector index $\mathbf{I} = i$ in fan-beam and tuple $\mathbf{I} = (i, j)$ in cone-beam, respectively, the response is modeled by sampling: $g_\theta(\boldsymbol{\omega}_{\mathbf{I}})$. The detector blur function, $h_\tau(\boldsymbol{\omega})$, is supported over its cell of width $\boldsymbol{\tau}$. Convolution in (2.3) shows the response of the $\mathbf{I}^{\text{th}}$ detector cell as a 1-D convolution in fan-beam and 2-D convolution in cone-beam geometries.

2.3. Discretized model. Discretization or a finite-dimensional approximation of a continuous-domain signal (image) f utilizes expansion in a basis set:

$$f_N(\mathbf{x}) = \sum_{n=1}^N c_n \varphi(\mathbf{x} - \mathbf{k}_n).$$

This expansion allows for derivation of a *discretized forward model*. Here φ is a basis function, and c_n is the expansion coefficient corresponding to the n^{th} basis function $\varphi_n := \varphi(\cdot - \mathbf{k}_n)$, which is the translation of φ to the grid point $\mathbf{k}_n$. The combination of integral transform (which is linear) and this expansion provides the discretized forward model:

$$(2.4) \quad \mathcal{P}_{\mathbf{u}, \boldsymbol{\tau}}\{f_N\}(\boldsymbol{\omega}_{\mathbf{I}}) = \sum_{n=1}^N c_n \mathcal{P}_{\mathbf{u}, \boldsymbol{\tau}}\{\varphi_n\}(\boldsymbol{\omega}_{\mathbf{I}}),$$

where

$$(2.5) \quad \begin{aligned} \mathcal{P}_{\mathbf{u}, \boldsymbol{\tau}}\{\varphi_n\}(\boldsymbol{\omega}_{\mathbf{I}}) &= \int h_\tau(\boldsymbol{\omega}_{\mathbf{I}} - \boldsymbol{\omega}) \mathcal{P}_{\mathbf{u}}\{\varphi_n\}(\boldsymbol{\omega}) d\boldsymbol{\omega} \\ &= (\mathcal{P}_{\mathbf{u}}\{\varphi_n\} * h_\tau)(\boldsymbol{\omega}) \end{aligned}$$

is the *detector blur* of a basis function and $\mathcal{P}_{\mathbf{u}}\{\varphi_n\}(\boldsymbol{\omega})$ is called the footprint of the n^{th} basis function φ_n . Therefore, the linear model (2.4) indicates that the forward-projection of the discretized model is the linear combination of the detector blur of basis functions. For a discretized model with N basis functions and a set of M detector measurements ($M = \text{number of viewing angles} \times \text{number of detectors}$), the forward model is an $M \times N$ matrix $\mathbf{A}$ which is calculated using (2.5) during forward- and back-projection. Specifically the contribution of the n^{th} pixel ($1 \leq n \leq N$) to the m^{th} measurement ($1 \leq m \leq M$) coming from the detector cell at $\mathbf{I}$ on the viewing direction $\mathbf{u}$ (with m being a linearized indexing of the detector and viewing direction indices), we have $[\mathbf{A}]_{m,n} = \mathcal{P}_{\mathbf{u}, \boldsymbol{\tau}}\{\varphi_n\}(\boldsymbol{\omega}_{\mathbf{I}})$. Efficient calculation of this forward model eliminates the need for precomputation and storage of $\mathbf{A}$ by allowing for on-the-fly computation during the forward- and back-projection (i.e., matched pair) steps for solving (1.1).

The most common choice of basis function φ is the pixel-basis in two dimensions and the voxel-basis in three dimensions that are simply 1 inside the pixel/voxel and 0 outside. As described in (2.5), the footprint and detector blur transform the basis function φ defined in image space to a function defined in the detector space $\mathcal{P}_{\mathbf{u},\boldsymbol{\tau}}\{\varphi\}$. The significance of this function is that it determines the contribution of a pixel/voxel to a particular detector. Efficient computation of $\mathcal{P}_{\mathbf{u},\boldsymbol{\tau}}\{\varphi\}$ is key to fast computation of projectors in forward- and back-projection. Since the ray direction $\mathbf{u}$ in line integrals and convolution directions $\boldsymbol{\tau}$ in detector blur do not align with pixel/voxel orientations, $\mathcal{P}_{\mathbf{u},\boldsymbol{\tau}}\{\varphi\}$ is a *nonseparable function*. Approximating this function by separable functions has been considered for efficiency purposes [38]. We demonstrate that this nonseparable function $\mathcal{P}_{\mathbf{u},\boldsymbol{\tau}}\{\varphi\}$ can be computed algebraically by *directional convolution* in the continuous domain (see (4.2) in section 4). The directional convolution of footprint and blurring effects (e.g., detector blur) with pixel and voxel basis result in special piecewise polynomial functions whose structure we use for efficient computation of detector blur and footprint transformations. Other sources of blur such as focal-spot and motion blurs can be integrated into the forward model by space-varying directional convolutions transforming the image domain basis into piecewise polynomials with higher degrees; however, in this paper we limit the discussion to the detector blur for simplicity of presentation.

3. Directional convolution and parallel projection. Box splines are piecewise polynomial functions that are defined based on a number of directions. The pixel- or voxel-basis itself can be viewed as simple (piecewise constant) box splines with unit orthogonal directions in two and three dimensions (see Figures 3 and 4). The benefit of viewing pixel- or voxel-basis as box splines becomes apparent once we establish that their footprint and detector blur integrals in $\mathcal{P}_{\mathbf{u},\boldsymbol{\tau}}\{\varphi\}$ result in a box spline whose directions are formed by the union of detector blur directions $\boldsymbol{\tau}$ with the unit orthogonal directions in pixel-/voxel-basis projected along the ray direction $\mathbf{u}$.

3.1. Algebra of directional convolutions. Box splines generalize B-splines to the multivariate setting where they include tensor-product B-splines as a special case, but are generally nonseparable functions. An *elementary* box spline, $M_{\boldsymbol{\xi}}$, associated with a vector $\boldsymbol{\xi} \in \mathbb{R}^d$ can be thought of as the indicator function of the set $\{t\boldsymbol{\xi} | 0 \leq t < 1\}$ and is formally defined as a Dirac-line distribution (generalized function) by its *directional convolution* with a (test) function $f(\mathbf{x})$ in $\mathbb{R}^d$: $(M_{\boldsymbol{\xi}} * f)(\mathbf{x}) = \int_0^1 f(\mathbf{x} - t\boldsymbol{\xi}) dt$. Given a set of $N \geq d$ directions, arranged in columns, as $\boldsymbol{\Xi} := [\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \dots, \boldsymbol{\xi}_N]$, the associated piecewise polynomial function is defined by

$$(3.1) \quad M_{\boldsymbol{\Xi}}(\mathbf{x}) = (M_{\boldsymbol{\xi}_1} * \dots * M_{\boldsymbol{\xi}_N})(\mathbf{x}),$$

and this is illustrated, for $d = 2$, in Figure 2. When the directions are orthogonal, $M_{\boldsymbol{\Xi}}$ is a tensor-product B-spline with the repeats of a direction elevating the degree of the B-spline. A box spline $M_{\boldsymbol{\Xi}}$ can be evaluated using de Boor–Höllig recurrence relations [9] that are in general computationally expensive. Efficient evaluation for a given set of directions $\boldsymbol{\Xi}$ is possible by leveraging their piecewise polynomial forms [15, 21].

3.2. Projection in parallel geometry. Previous work [13] has demonstrated that in *parallel* geometry the projection of a box spline $M_{\boldsymbol{\Xi}}$ defined in image space (e.g., pixel/voxel) results in another box spline, $M_{\mathbf{Z}}$ in the detector space. The directions $\mathbf{Z}$ of the latter in detector

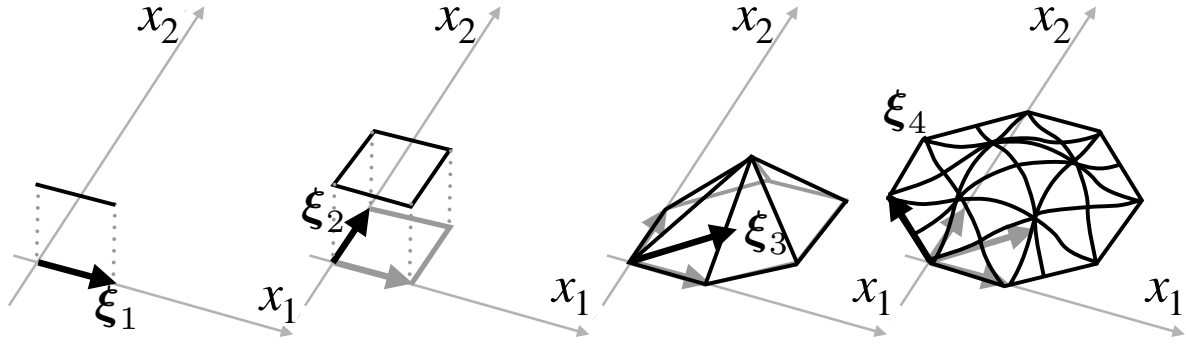


Figure 2. Directional convolution along a set of vectors with pixel basis viewed as a box spline corresponding to two unit orthogonal vectors.

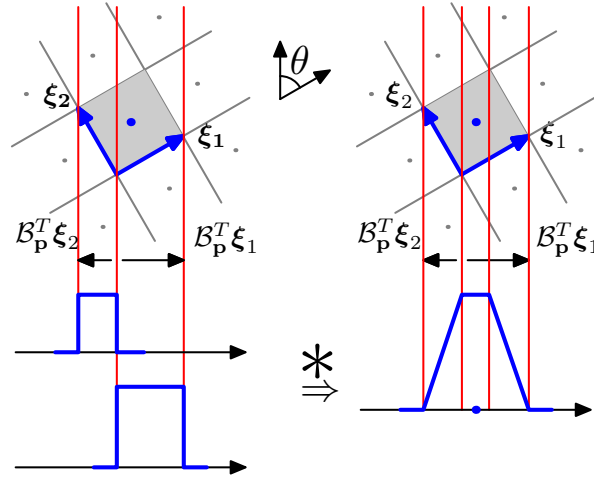


Figure 3. Parallel projection of pixel basis, $\varphi = M_{\Xi}$ viewed as a directional convolution.

space are the *geometric* projection of the directions Ξ in image space. Let $\mathcal{R}_{\mathbf{u}}$ denote the parallel projection for a viewing direction specified by vector $\mathbf{u}$, then we have

$$(3.2) \quad \mathcal{R}_{\mathbf{u}}\{M_{\Xi}\} = M_{\mathbf{Z}} = M_{\zeta_1} * \cdots * M_{\zeta_N},$$

where $\zeta_n := \mathcal{B}_p^T \xi_n(\tilde{\mathbf{1}})$ (for $1 \leq n \leq N$) is the geometric projection of the vector ξ_n by dropping the first coordinate of $\mathcal{B}_p^T \xi_n$ after view transformation by $\mathcal{B}_p^T$. By this notation, we have $\mathbf{Z} := \mathcal{B}_p^T \Xi(\tilde{\mathbf{1}})$. Figure 3 shows the pixel-basis viewed as a box spline, $\varphi = M_{\Xi}$, defined in $\mathbb{R}^2$ specified by two directions $\Xi = [\xi_1, \xi_2]$ with $\xi_1 = [1 \ 0]^T$, and $\xi_2 = [0 \ 1]^T$. When projected to the detector space, the projection is a box spline, $M_{\mathbf{Z}}$, defined in $\mathbb{R}^1$ with two directions $\mathbf{Z} = [\zeta_1, \zeta_2] = [\cos \theta, -\sin \theta] = \mathcal{B}_p^T \Xi(\tilde{\mathbf{1}})$, which is the 1-D convolution of two elementary box splines. Similarly in three dimensions, as shown in Figure 4, the voxel-basis can be viewed as a box spline, $\varphi = M_{\Xi}$, defined in $\mathbb{R}^3$ specified by three directions $\Xi = [\xi_1, \xi_2, \xi_3]$ where $\xi_1 = [1 \ 0 \ 0]^T$, $\xi_2 = [0 \ 1 \ 0]^T$, and $\xi_3 = [0 \ 0 \ 1]^T$. Its projection in the detector space is described again by a box spline $M_{\mathbf{Z}}$ where $\mathbf{Z} = [\zeta_1, \zeta_2, \zeta_3] = \mathcal{B}_p^T \Xi(\tilde{\mathbf{1}})$ are the geometric projection of directions in Ξ by view-transformation $\mathcal{B}_p$.

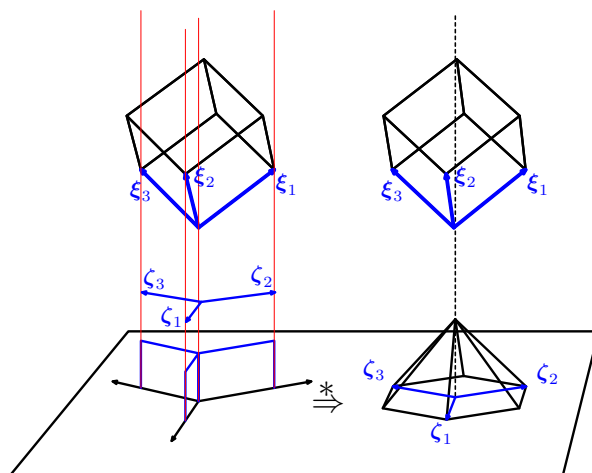


Figure 4. Parallel projection of voxel basis as convolution.

3.3. Detector blur. Having the footprint represented as a directional convolution in the detector space allows us to model detector blur with additional directional convolutions aligned with detector cells τ . The detector sensitivity is often modeled as a constant function over the detector cell or with a drop-off at the cell boundary [22], and the blur function h_τ can be modeled as an elementary box spline: $h_\tau = M_\tau$. Sensitivity functions with a faster decay on cell boundary can also be modeled with higher order box splines. The parallel beam projection of pixel-/voxel-basis with detector blur is a directional convolution in the detector space:

$$(3.3) \quad \mathcal{R}_{\mathbf{u}, \tau}\{M_{\Xi}\} = \mathcal{R}_{\mathbf{u}}\{M_{\Xi}\} * M_\tau = M_{\mathbf{Z}} * M_\tau = M_{[\mathbf{Z}, \tau]}.$$

This allows for adding detector blur (and similarly focal-spot blurs) to previous results in parallel beam geometry [13]; moreover, we can leverage the geometric relationship between *directions of blur* in image space and detector space to extend this result to divergent geometries. To that end let ν be a vector in image domain parallel to the detector plane whose geometric projection coincides with detector blur directions: $\tau = \mathcal{B}_p^T \nu(\tilde{1})$. Using the Fourier slice-projection theorem we can show that projection commutes with convolution. This allows us to show that the detector blur τ can also be modeled by additional directional blurs ν , applied to basis φ , in the image domain:

$$(3.4) \quad \mathcal{R}_{\mathbf{u}, \tau}\{M_{\Xi}\} = M_{[\mathbf{Z}, \tau]} = \mathcal{R}_{\mathbf{u}}\{M_{[\Xi, \nu]}\}.$$

An illustration is presented in Figure 5, where $\nu = [\nu_s, \nu_t]$.

4. Directional convolution in divergent ray geometries. In fan-/cone-beam geometries, the divergent nature of rays prevents a direct application of directional convolution equivalence in image and detector spaces since rays are not perpendicular to the detector plane. To circumvent this problem, we formulate the projection in a coordinate system designed for each ray.

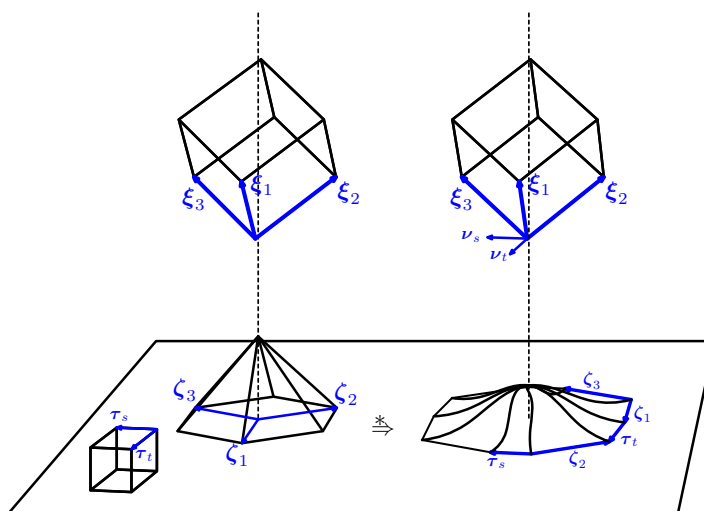


Figure 5. Parallel projection of voxel-basis with detector blur.

4.1. Virtual detector plane and ray coordinate system. The coordinate systems for particular rays are shown as the $\mathbf{RU}$ -axis in green in Figure 1, left, and the $\mathbf{RUV}$ -axis in Figure 1, right. Given a point in detector space with coordinate ω , the $\mathbf{R}$ -axis in this coordinate system has the ray's orientation, in the reverse direction: $\mathbf{R}(\omega) = -\mathbf{r}(\omega)$. We refer to the plane orthogonal to the ray as the *virtual detector plane* denoted by $\mathbf{R}(\omega)^\perp$. In the fan-beam case, a unit vector $\mathbf{U}(\omega)$ in the virtual detector plane defines the ray-dependent view-transformation matrix $\mathcal{B}_\omega = [\mathbf{R}(\omega) \ \mathbf{U}(\omega)]$. In the cone-beam case, we pick a pair of orthogonal vectors in the virtual detector space that form the ray-dependent view-transformation matrix: $\mathcal{B}_\omega = [\mathbf{R}(\omega) \ \mathbf{U}(\omega) \ \mathbf{V}(\omega)]$. As illustrated in Figure 6 the intersection of virtual detector and real detector planes includes ω .

4.2. Projection in divergent ray geometry. In parallel geometry the relationship between directions in image space (e.g., voxel directions) and directions in detector space (e.g., blur directions) was fixed for all rays in a particular viewing angle. This relationship was characterized by a single $\mathcal{B}_p$ which implied footprint and detector blurs are precisely box splines as shown in (3.2) and (3.4). The coordinate system defined for each ray provides a similar transformation, $\mathcal{B}_\omega$, that we leverage for relating the directional convolutions in image and detector spaces. However, since the ray coordinate system changes for each point ω in the detector space, this relationship is no longer constant for a viewing angle. Nonetheless, we show that the resulting footprint function can be evaluated by a transformation of directional convolution.²

Theorem 4.1. *In a divergent ray geometry, the footprint of image space directional convolution is a space-varying directional convolution:*

$$(4.1) \quad \mathcal{P}_{\mathbf{u}}\{M_{\Xi}\}(\omega) = M_{\mathbf{Z}_\omega}(\mathbf{p}_\omega).$$

²A preliminary approach experimented with fan-beam geometry was presented in a conference talk [64].

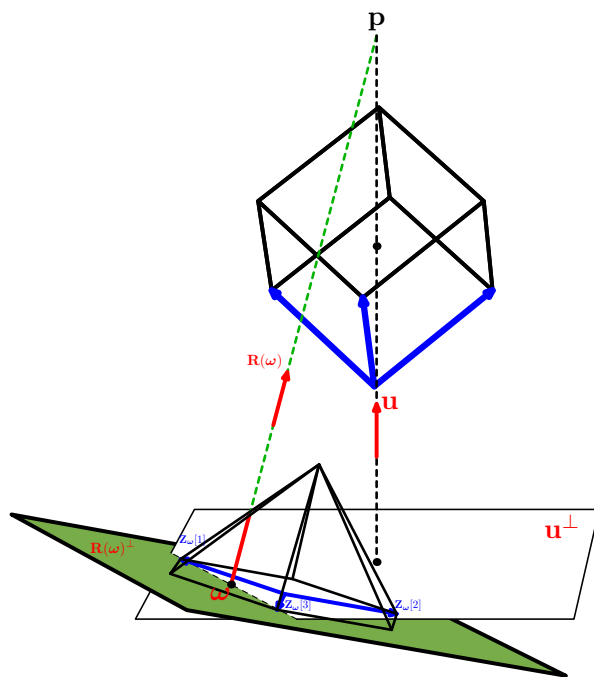
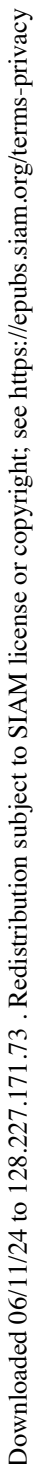


Figure 6. Cone-beam projection of voxel-basis in the real detector plane (white) $\mathbf{u}^\perp$ at a point ω computed via the projection of source $\mathbf{p}$ on the virtual detector plane (green) $\mathbf{R}(\omega)^\perp$. The intersection of the two planes contains the detector point ω .

Here the image space directions Ξ are projected on the virtual detector plane $\mathbf{R}(\omega)^\perp$ as $\mathbf{Z}_\omega := \mathcal{B}_\omega^T \Xi(\tilde{\mathbf{i}})$, and the resulting directional convolution is evaluated at the projection of the source, $\mathbf{p}$, on the virtual detector plane $\mathbf{p}_\omega := \mathcal{B}_\omega^T \mathbf{p}(\tilde{\mathbf{i}})$.

The arguments for establishing this result are detailed in Appendix B. A direct implication is that the footprints of a pixel-basis in the fan-beam and a voxel-basis in the cone-beam are space-varying directional convolutions with two and three directions, respectively. The cone-beam projection of the voxel-basis is illustrated in Figure 6.

4.3. Detector blur via image space effective blur. Since the set of directions $\mathbf{Z}_\omega$ in the footprint function vary with the detector space coordinates ω , a further convolution with detector function h_τ (with fixed directions τ) cannot be formed directly by joining the directions as in (3.3). Since (4.1) provides exact projection not only for the voxel-basis but also for any set of directions, Ξ , we can back-project the detector cell directions, τ , to the voxel and augment the voxel directions with the back-projected detector cell directions before applying the exact footprint computation (4.1). The divergent nature of rays leads to a perspective scaling of the detector blur directions τ during back-projection. Considering parallel planes perpendicular to the ray slicing a single voxel, the perspective scaling varies for each slice depending on its depth. We introduce the notion of effective blur that applies perspective scaling according to the center slice (passing through the center of voxel) for back-projection of detector blur. Figure 7 shows the back-projection of detector cell τ to a plane parallel to the virtual detector plane running through the center of the pixel/voxel resulting in the



Downloaded 06/11/24 to 128.227.171.73 . Redistribution subject to SIAM license or copyright; see <https://epubs.siam.org/terms-privacy>

Downloaded 06/11/24 to 128.227.171.73 . Redistribution subject to SIAM license or copyright; see <https://epubs.siam.org/terms-privacy>

Downloaded 06/11/24 to 128.227.171.73 . Redistribution subject to SIAM license or copyright; see <https://epubs.siam.org/terms-privacy>

Downloaded 06/11/24 to 128.227.171.73 . Redistribution subject to SIAM license or copyright; see <https://epubs.siam.org/terms-privacy>

Downloaded 06/11/24 to 128.227.171.73 . Redistribution subject to SIAM license or copyright; see <https://epubs.siam.org/terms-privacy>

Downloaded 06/11/24 to 128.227.171.73 . Redistribution subject to SIAM license or copyright; see <https://epubs.siam.org/terms-privacy>

Downloaded 06/11/24 to 128.227.171.73 . Redistribution subject to SIAM license or copyright; see <https://epubs.siam.org/terms-privacy>

footprint (CNSF).³ As it has been demonstrated that LD (LTRI) and TR (SF) are the most time-efficient methods in LTRI and SF, and LL (LTRI) and TT (SF) are the most accurate methods in LTRI and SF, respectively, we choose the LD method in the computational efficiency experiments and LL and TT in all accuracy experiments.

All fast projectors designed for practical forward- and back-projection strike a balance between computational cost and accuracy of projection in (2.3). Our experiments, below, demonstrate CNSF provides significant improvements in computational efficiency that are achieved without compromising the accuracy. Indeed, we observe slight improvements in accuracy of forward-projection that impact reconstruction accuracy.

5.1. Forward-projection: A single voxel study. To assess the accuracy of forward-projection, we first examine a “microscopic” view of a single voxel when projected by different methods. We simulate a cone-beam system with a flat detector with a single voxel centered at radiation isocenter with size 1 mm^3 . The detector bin width $\tau = (\tau_s, \tau_t) = (0.5 \text{ mm}, 0.5 \text{ mm})$, the source to rotation center distance $D_{so} = 5 \text{ mm}$, and the source-detector distance $D_{sd} = 11 \text{ mm}$. In practice the detector bin width τ is identical to the sampling distance Δ in the detector space (i.e., sinogram domain). In this study, however, in order to illustrate the differences in a close-up view of projection of a single voxel, we oversample the continuous sinogram domain by choosing a very small $\Delta = (0.01 \text{ mm}, 0.01 \text{ mm})$.

We used the MIRT toolbox [19] and LTRI [26] systems to generate the comparisons and we implemented the CNSF projector in CUDA by organizing the voxels intersecting a ray in a block to utilize the warp shuffle feature. Figure 8 shows the profile from the projection of voxel-basis for the four projectors with detector blur. The first column shows the ground truth projection from different viewing angles and the three columns show the absolute errors made by each fast projector. This study shows that the proposed method consistently provides the most accurate projector in all of the viewing angles.

5.2. Forward-projection: A field-of-view study. The previous study focused on showing the accuracy of the projection of a single voxel, in a restricted geometric setting designed to enlarge errors in projectors. To assess the accuracy in a more realistic geometric setting, we focus on the performance of projectors across different angles and spatial distribution of errors made in various locations in the field of view. In this experiment we study forward-projection of a cubic region of size 32 mm^3 located at the radiation isocenter. To analyze spatial distribution of errors, we discretized the cube at the resolution of (32^3) such that the size of each voxel is 1 mm^3 , and the attenuation coefficient of each voxel was $(1/\text{mm})$. The source-rotation center distance $D_{so} = 200 \text{ mm}$, source-detector distance $D_{sd} = 261 \text{ mm}$, and detector bin width τ was $(1 \text{ mm} \times 1 \text{ mm})$. To cover the field of view, we used an array of detectors of size $(N_s \times N_t) = (123 \times 109)$ and placed each of them at the distance of 1 mm in each of the s and t directions. In order to analyze performance of projectors in different spatial positions (voxels), we picked the density in the cubic region to be constant across all voxels.

To investigate the global accuracy performance of all methods, we categorized the experiments into two groups that show different aspects of projectors. In the first group,

³We focus on cone-beam geometry, as experiments in fan-beam geometry appeared in a conference proceedings [65].

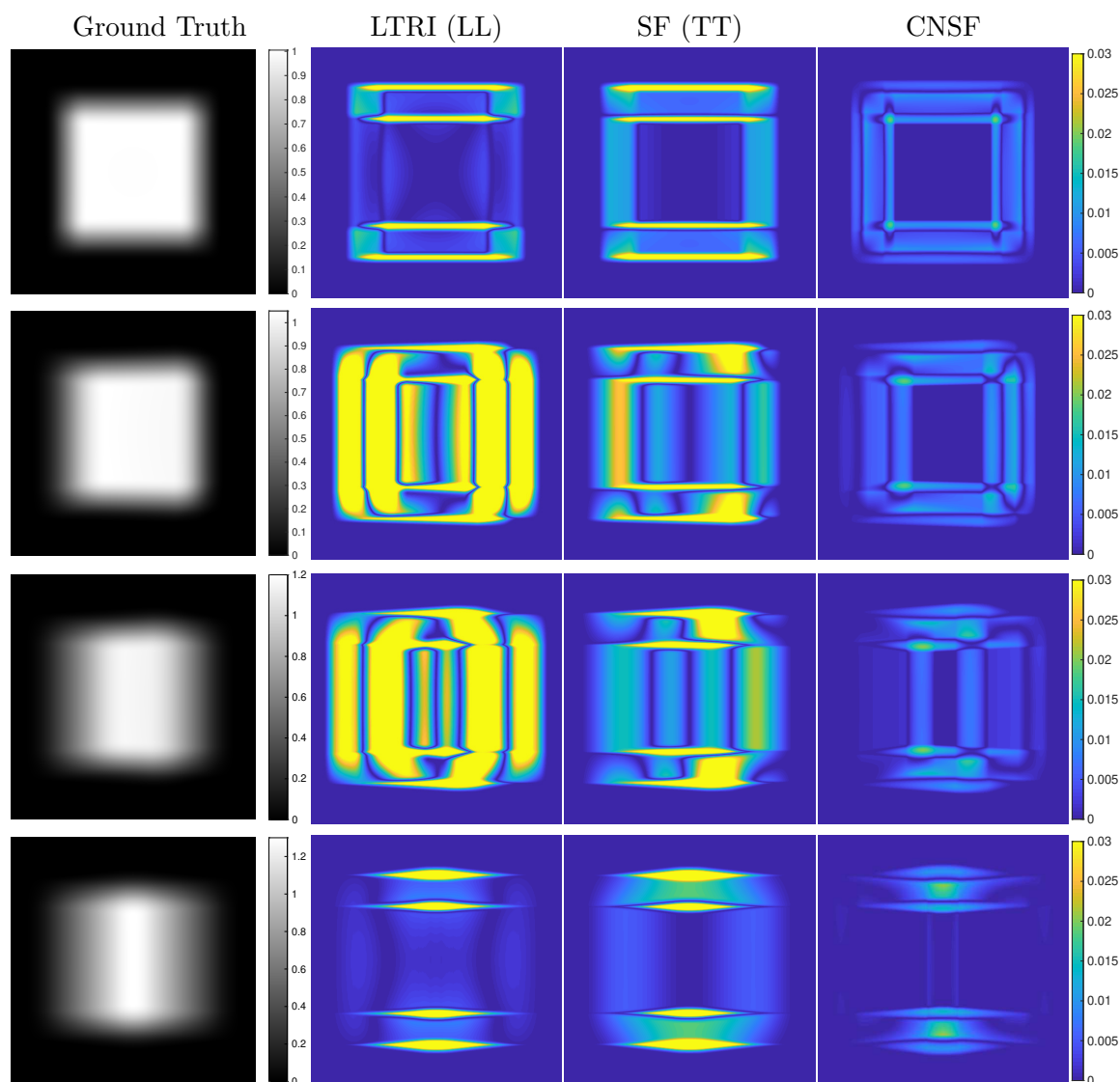


Figure 8. Accuracy of forward-projection from the viewing angles: 0° , 15° , 35° , and 45° (rows 1 through 4). Columns 2 through 4 are the absolute difference images between different approximated projectors and reference projector. The unitless color bar values are accumulated values by voxels along each ray.

we show the root mean square error (RMSE) in the detector plane for each viewing angle θ : $\text{RMSE}(\theta) = (\frac{1}{N} \sum_{i=1}^N (g_\theta(\omega_i) - \tilde{g}_\theta(\omega_i))^2)^{1/2}$, where g is the response of the reference projector according to (2.3) and $\tilde{g}$ is the response provided by the fast projectors in our study. In the second group, to illustrate spatial distribution of errors, we characterize the error in each detector bin along different viewing angles. This is measured by $\text{RMSE}(\omega) = (\frac{1}{N} \sum_{i=1}^N (g_{\theta_i}(\omega) - \tilde{g}_{\theta_i}(\omega))^2)^{1/2}$. The results of these measurements are shown in Figure 9. This study shows that the error of each projector fluctuates depending on

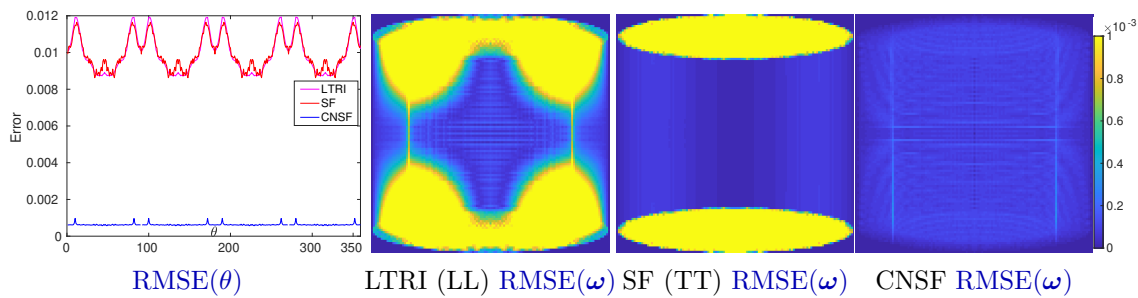


Figure 9. Average projection error (unitless) over detector space for each viewing angle (leftmost) and over viewing angles for each detector space point (the rest).

Table 1

Average fan-beam projection time in milliseconds.

Resolution	256 ²	512 ²	1024 ²	2048 ²	4096 ²
System Size	102.2M	815M	6.3G	50.7G	355.4G
LTRI (LD)	0.136	0.310	1.000	3.950	14.080
CNSF	0.002	0.011	0.038	0.124	0.446

projection angles and spatial locations, with our proposed method showing consistently smaller distribution of errors in both the angular and spatial dimensions.

5.3. Computational efficiency. The computational efficiency of our approach stems from efficient computation of the right-hand side of (4.2) that relies on evaluation of a box spline in (4.1), which, in turn, translates to evaluating a specific polynomial determined from the directions of convolution. The polynomial computation eliminates the need for accessing memory hierarchy (e.g., a lookup table in LTRI) that limits the GPU performance. In this experiment, we constructed a fan-beam system with flat detectors for pixels of size 1 mm² with $N_s = 409, 815, 1627, 3250, 6499$ detectors with $\tau = 1$ mm corresponding to image resolutions of 256², 512², 1024², 2048², and 4096². For the 256² resolution we had $D_{so} = 200$ mm, $D_{sd} = 400$ mm, which were doubled for each of the higher resolutions. We measured average projection time for 360 views, which are displayed in Table 1.

As the GPU implementation of the SF method is protected by a patent, it is not possible to perform GPU comparisons. However, its CPU implementation is available as part of the MIRT package [19]. We implemented our CNSF method with the Intel Threading Building Blocks library to compare the run time with SF with CPU execution in a fan-beam geometry setup. In the CPU experiment, we simulated a fan-beam system with flat detectors for pixel sizes of 1mm² with $N_s = 125, 251, 409, 815, 1627, 3250$ detectors with $\tau = 1$ mm corresponding to image resolutions of 64², 128², 256², 512², 1024², and 2048². For the 64² resolution we have $D_{so} = 50$ mm, $D_{sd} = 100$ mm, which were doubled for each of the higher resolutions. Figure 10 shows the speedup of average projection of 360 views over SF. Since the computational resources on CPU are limited, the CPU speedups are less significant than those accomplished by GPU computation of CNSF. Nevertheless, this experiment shows the projector reduces the computational burden of fan-beam projectors by a factor of 2 compared to the SF method.

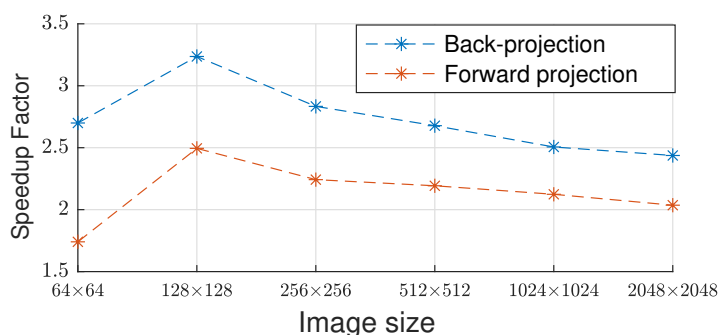


Figure 10. CPU run time comparisons with SF.

Table 2

Average cone-beam projection time (milliseconds) in GPU execution.

Resolution	16 ³	32 ³	64 ³	128 ³	256 ³
System Size	16.52M	528.64M	16.52G	528.64G	16.52T
LTRI (LD)	1.18	1.54	2.34	6.88	27.26
CNSF	0.04	0.07	0.21	0.30	17.89

We also set up a cone-beam flat detector X-ray CT system for voxels of size 1 mm^3 with $(N_s \times N_t) = (31 \times 32), (63 \times 65), (126 \times 130), (252 \times 261), (505 \times 523)$ detectors corresponding to image resolutions of 16^3 , 32^3 , 64^3 , 128^3 , and 256^3 . For all these experiments, $D_{sd} = 1300 \text{ mm}$, $D_{so} = 1000 \text{ mm}$, τ was $(1 \text{ mm} \times 1 \text{ mm})$, and we observed the average projection time using 360 views.

The computational demands for cone-beam geometry makes it impractical to compare the fast projectors in CPU computations. Table 2 documents the average time in forward-projection for various resolutions in GPU execution. As discussed before, the GPU implementation of the SF method is unavailable for comparison in this experiment. However, it has been argued previously [25] that both LD (LRTI) and TR (SF) are the most efficient projectors, and that they perform similarly in GPU computations. Therefore, we only conducted comparisons with LTRI (LD) as reported in this table. Yet for a specific comparison with SF, based on a recent GPU improvement that reduces memory access [8], we note that at 256^3 resolution the average projection time for SF is reported to be 129msec (corresponding to the last column in the table). All experiments were performed on an NVIDIA-TitanX GPU with CUDA 11.2, Intel i7 6-core CPU. As a reference, the *System size* row in the table documents the memory size necessary to store the forward model, a limitation that fast projectors avoid by enabling on-the-fly computation of the forward model.

Eliminating the need to access a lookup table explains the observed performance improvements in the GPU implementation of our approach compared to LTRI. As the polynomial computation (4.1) is carried out using the registers available to the CUDA kernel, with increase in resolution, the number of threads in a block increases the demand for registers, which in turn hampers the performance as slower memory is used more frequently. As the speedup attained by our method shown in Table 2 at 64^3 or 128^3 compared to 256^3 shows, availability

of more registers/kernels (e.g., GPUs more recent than ours) is necessary to maintain an order of magnitude speedup over LTRI. Without having access to more resourceful GPUs, one can also minimize the number of registers necessary for computing the polynomials in (4.1) to boost parallel computations. This may be done using symbolic manipulation as well as compiler optimization tools.

5.4. Reconstruction. We have observed improvements in accuracy of the forward model provided by the directional convolution framework over other projectors. To assess whether these improvements can impact the image reconstruction, in the presence of regularization or the nonlinearity in the Lambert–Beer law, the following study compares the reconstruction accuracy of the fast projectors. In terms of accuracy, the reference reconstructions, provided by the computationally expensive reference projector, provide upper bounds on the practical projectors SF (TT), LTRI (LL), and CNSF in this study.

5.4.1. 2-D reconstruction. As mentioned, our framework in 2-D fan-beam geometry was presented in a conference paper [65]. For assessing the impact of accurate forward-projection on reconstruction with regularization, we present a similar study using the Forbield head phantom [63], designed for assessment of image reconstruction algorithms, and refer to [65] for more studies. In this experiment, we obtained the sinogram data by evaluating the line integrals (2.2) analytically followed by numerical integration across each detector cell. The solver we used for iterative reconstruction is ASD-POCS [50] and we optimized the hyperparameters (i.e., total-variation regularization parameter and number of iterations) separately for the lowest value of objective function possible in each case.

The size of the phantom we used is 25 mm $\times$ 25 mm, and we reconstructed it by using an image at the resolution of 512 $\times$ 512 resulting in 0.0488 mm $\times$ 0.0488 mm pixels discretizing the object. We simulated a flat detector fan-beam CT system with $D_{sd} = 78.125$ mm and $D_{so} = 39.0625$ mm, and a flat detector with 1627 bins (i.e., N_s) and 0.0488 mm bin size (i.e., τ). The reconstruction result is shown in Figure 11. The insets show the close-up views of ROI in Figure 11. The result achieved by the reference projection is almost perfect with some small artifacts due to the limited number of views available. The images reconstructed by all the other methods contain heavier artifacts around the right black dots; as evident in this experiment CNSF provides an image that is closest to the reference reconstruction, compared to other projectors.

We also conducted an experiment to examine the impact of improved accuracy of projectors in presence of the nonlinearity brought about by the Lambert–Beer law that is always present in practice. In this experiment, we obtained the sinogram data by evaluating the line integrals (2.2) analytically whose exponentials were integrated numerically across each detector cell. We removed regularization in this experiment to better study the difference between projectors (CNSF, SF, and LTRI) when nonlinearity in data is present. Using 360 uniformly spaced projections, we employed a commonly used iterative solver (i.e., the simultaneous algebraic reconstruction technique [1]) for each projector. While the presence of artifacts coming from the nonlinearity makes the visual differences between reconstructions more subtle, the improvements brought about by CNSF are observable in error images, as shown in Figure 12. While in our experiments we observed CNSF needed fewer iterations, compared to LTRI and SF, to reach a desired error level (due to differences in spectra of forward models), we ran all

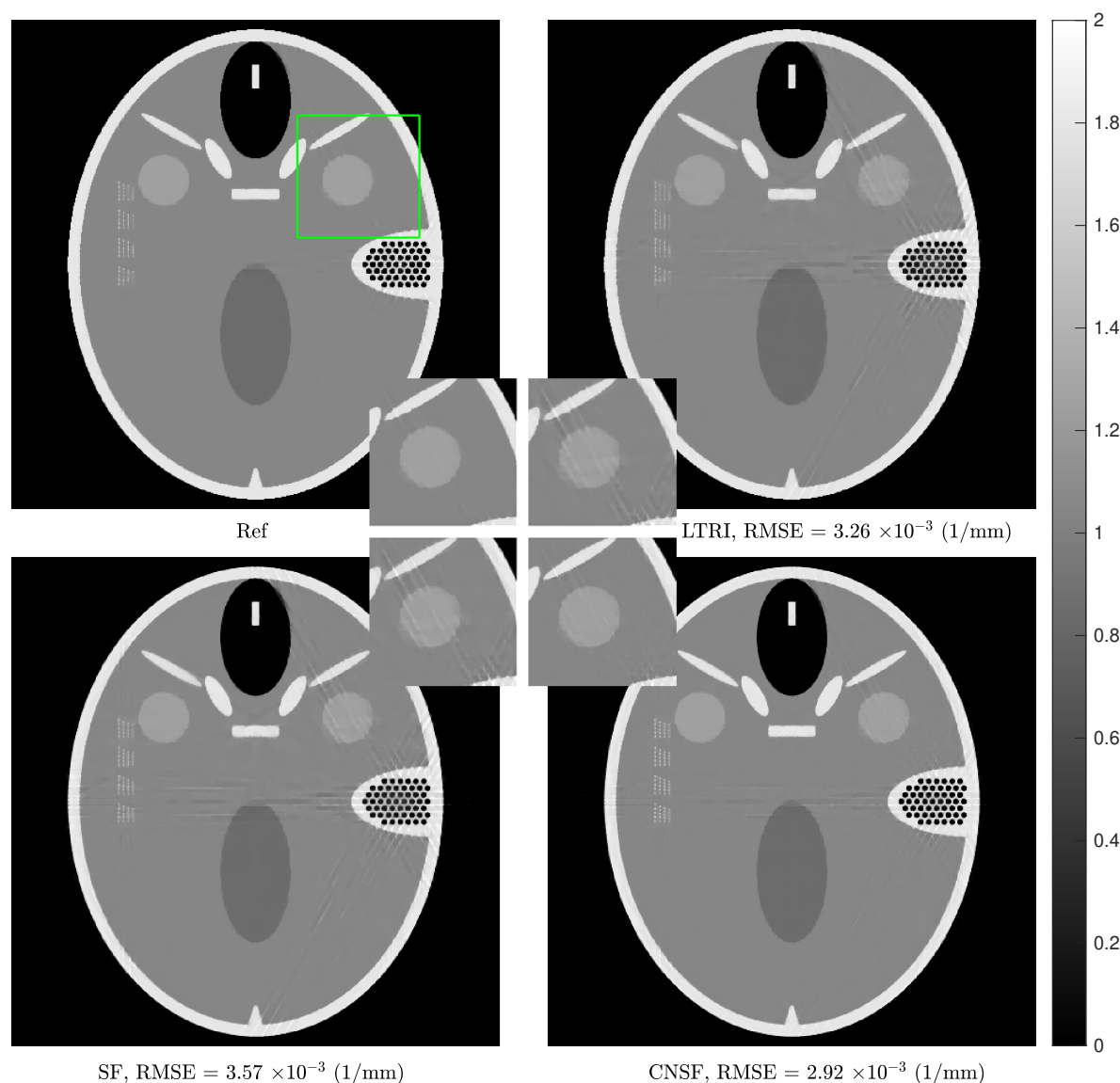


Figure 11. Reconstruction of FORBILD head phantom from 36 uniformly spaced projections using ASD-POCS. The RMSE measures deviation from reconstruction provided by the reference projector (top left). The insets in the center are the zoomed-in view of the green box area.

methods to a large number of iterations (i.e., 2000) so the reconstructions reflect the highest accuracy possible for each projector. We also observed that, in this experiment, applying the regularization makes the errors even smaller but again CNSF achieves the lowest error among the fast projectors.

5.4.2. 3-D reconstruction. In this experiment, we used Shepp–Logan with resolution of 512^3 and 1mm^3 voxel size as the benchmark dataset. The simulated flat detector X-ray system was configured with $N_s = 600$ and $N_t = 700$, $D_{so} = 970$ mm, $D_{sd} = 1337$ mm. The

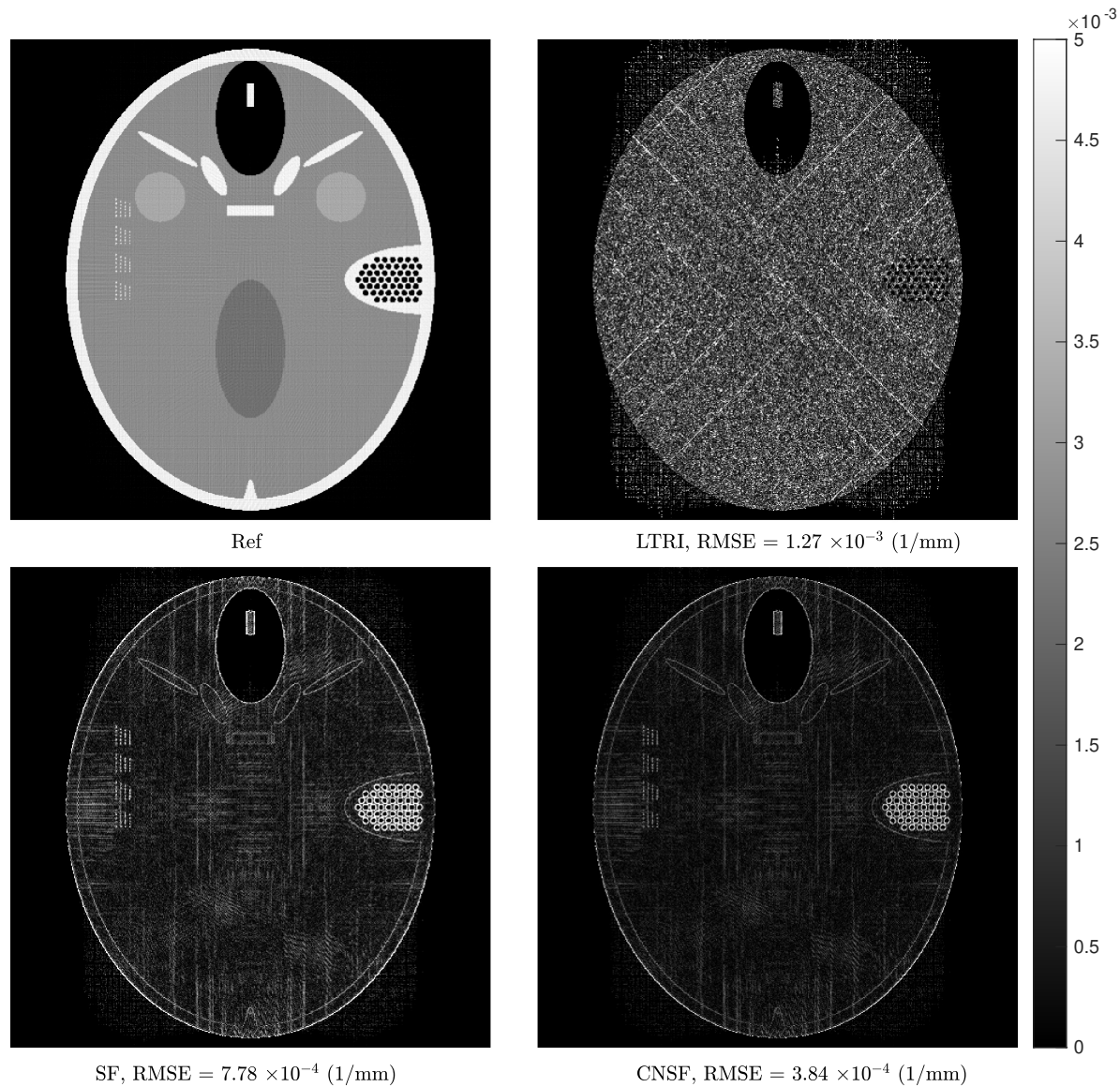


Figure 12. Reconstruction of FORBILD head phantom from 360 uniformly spaced projections, in the presence of the nonlinearity in the Lambert–Beer law, using SART with 2000 iterations. The reconstruction by the reference projector is shown in the top left image and its differences with reconstructions from fast projectors are shown in the other three images.

detectors were spaced by $\Delta = [1 \ 1]^T$ mm with bin width $\tau = [1 \ 1]^T$ mm and 360 views were uniformly spaced over 360° . To avoid providing data from the discretized solution to the inverse problem, the projection data was generated by the reference projector acting on an image with a higher resolution that mismatches the reconstruction resolution. Moreover, to assess the impact of inaccuracies in fast projectors on the final reconstructed images, we avoided the use of (explicit) regularization in this experiment and used the SART solver.

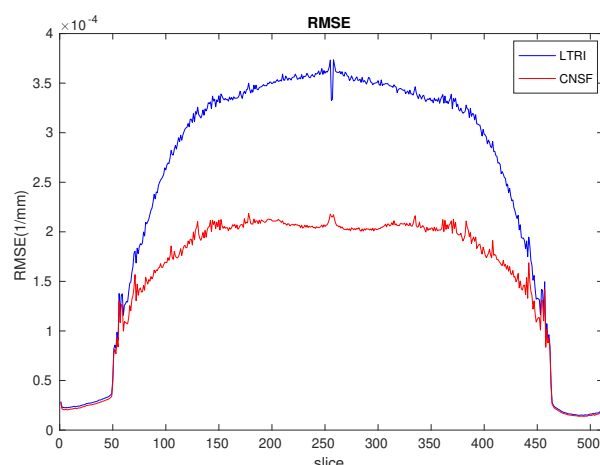


Figure 13. RMSE of 3-D reconstructions by LTRI and CNSF projectors.

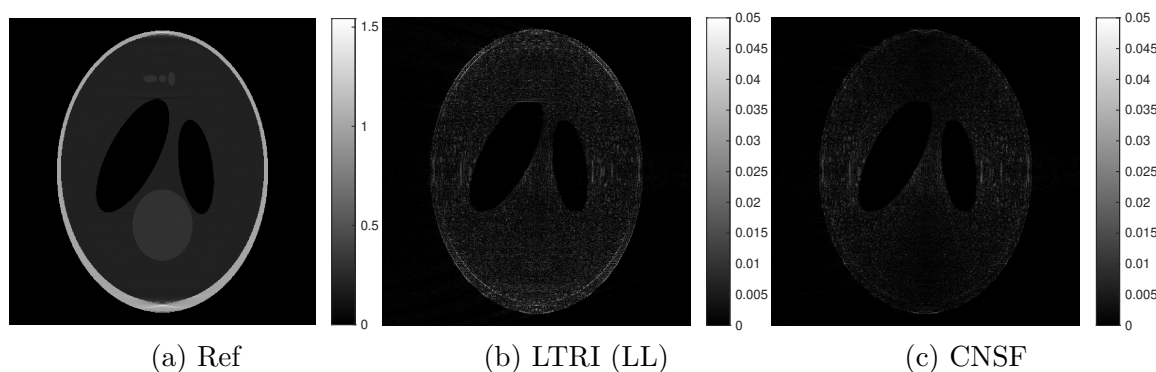


Figure 14. A slice from the 3-D cone-beam reconstruction, by SART, at resolution $512 \times 512 \times 512$. Reconstruction by the reference projector (a) and differences between reconstructions by LTRI (b) and our CNSF (c) from the reference reconstruction.

Figure 13 shows an assessment of accuracy in reconstruction by these projectors from 360 views, where the reference projector reconstruction shows the highest accuracy achievable from the given data. The result shows that, numerically, the quality of the projector does contribute to the accuracy of the reconstruction, and CNSF's improvements in the forward model are observable in the final reconstruction. A slice of this 3-D reconstructed image is shown in Figure 14 along with the absolute differences from reference reconstruction which are scaled by 512 for visualization purposes.

The imperfect reconstruction achieved by the reference projector in Figure 14(a) illustrates the best reconstruction possible from the given projection views, if we are not concerned about computational cost. The approximations introduced by fast projectors further reduce the reconstruction accuracy. This experiment shows that the CNSF method provides quantifiable improvements over LTRI and results in a reconstruction that achieves near identical quality as the one provided by the reference projector in this experiment.

6. Conclusion and future work. This paper introduces a directional convolution framework for modeling X-ray optics for iterative methods in CT reconstruction. The framework allows for modeling projection as well as detector blur in a computationally efficient manner into the forward model commonly used in MBIR. While the framework has the flexibility to model any CT geometry and any integral transform (e.g., blur effects) in the continuous domain, we demonstrated its effectiveness in fan-beam and cone-beam geometries with detector blur. From the computationally efficiency viewpoint, we demonstrate the method provides a significant speedup compared to state-of-the-art methods in fan-beam and cone-beam CT. Our experiments show that the observed speedup is accomplished without compromising the accuracy in modeling X-ray optics, and also the accuracy improvements were observed both in forward-projection and in the final reconstructed images.

Our current CUDA implementation can be optimized in terms of register allocation and other strategies exploiting the access pattern of ray-volume intersection that increase the throughput for boosting performance. Moreover, other sources of blur such as focal-spot and motion blur can be integrated into the directional convolution framework.

Appendix A. Glossary of symbols.

Glossary.

- M_{Ξ} Box spline associated with direction set Ξ .
- $\mathcal{P}_{\mathbf{u},\tau}$ Fan-/cone-beam X-ray transform from direction $\mathbf{u}$ with detector blur of cell width τ .
- $\mathcal{P}_{\mathbf{u}}$ Fan-/cone-beam X-ray transform from direction $\mathbf{u}$.
- $\mathcal{R}_{\mathbf{u}}$ Parallel beam X-ray projection from direction $\mathbf{u}$.
- $\mathcal{R}_{\mathbf{u},\tau}$ Parallel beam X-ray projection from direction $\mathbf{u}$ with detector of cell with τ .
- Ξ A set of vectors in image space arranged in columns of matrix Ξ for directional convolution.
- $\mathbf{Z}$ A set consists of geometric projection of directions in Ξ onto detector space.
- $\xi \in \Xi$, A d -dimensional unit direction vector in image space.
- $\zeta \in \mathbf{z}$, A $d - 1$ -dimensional unit direction vector in sinogram domain.
- ω Coordinates in detector space.
- $\mathbf{p}$ Location of source.
- $\mathbf{r}$ Unit direction of each X-ray.
- τ Vector(s) representing detector cell boundary (detector bin).
- $\mathbf{u}$ Unit viewing directions.
- $\mathbf{x}$ Coordinate in image space.
- D_{sd} Distance from source to the detector plane.
- D_{so} Distance from source to the rotation center.
- $\mathcal{B}_p$ The transformation matrix mapping coordinates from image space to detector space.
- θ Viewing angle.
- φ Image space basis function.
- h_{τ} Detector blur function over support of τ .

Appendix B. Proof of Theorem 4.1.

Proof. As in (2.2), the line integral of a box spline is

$$(B.1) \quad \mathcal{P}_{\mathbf{u}}\{M_{\Xi}\}(\omega) = \int_0^\infty M_{\Xi}(\mathbf{p} + \lambda \mathbf{r}(\omega)) d\lambda.$$

The line equation $l_{\omega}(\lambda) = \mathbf{p} + \lambda \mathbf{r}(\omega)$ can also be reparameterized by changing the starting point and the direction to $l_{\omega}(\lambda) = \mathcal{B}_{\omega}[\frac{0}{\nu}] - \lambda \mathbf{r}(\omega)$, where $\nu \in \mathbb{R}^{d-1}$ is the local coordinate on a virtual plane that can be computed by the orthogonal projection of $\mathbf{p}$ to the virtual plane: $\nu = \mathcal{B}_{\omega} \mathbf{p}(\tilde{\mathbf{I}})$. For simplicity, as the virtual plane is orthogonal to the $\mathbf{R}$ -axis, we use the notation $\mathbf{R}^{\perp} \in \mathbb{R}^{d \times (d-1)}$ to represent the basis of the virtual plane. According to the definition of $\mathbf{R}$ -axis whose direction is the opposite of the ray direction, we have $\mathbf{R} = -\mathbf{r}(\omega)$. Therefore, the line equation becomes

$$(B.2) \quad \mathcal{P}_{\mathbf{u}}\{M_{\Xi}\}(\omega) = \int_0^\infty M_{\Xi}(\mathbf{R}^{\perp} \nu + \lambda \mathbf{R}) d\lambda,$$

which is the exact form of the X-ray projection of box spline in parallel geometry in [14] and [13]. Therefore the right-hand side of (B.2) can be derived as

$$(B.3) \quad \begin{aligned} \int_0^\infty M_{\Xi}(\mathbf{R}^{\perp} \nu + \lambda \mathbf{R}) d\lambda &= \mathcal{R}_{\mathbf{R}}\{M_{\Xi}\}(\nu) \\ &= M_{\mathcal{B}_{\omega}^T \Xi(\tilde{\mathbf{I}})}(\nu) \\ &= M_{\mathbf{Z}(\omega)}(\nu). \end{aligned}$$

Thus we can conclude that $\mathcal{P}_{\mathbf{u}}\{M_{\Xi}\}(\omega) = M_{\mathbf{Z}(\omega)}(\nu)$ by combining (B.2) and (B.3) (see Figure 6). ■

Acknowledgment. We thank the anonymous reviewers, whose comments and suggestions improved various aspects of this paper.

REFERENCES

- [1] A. H. ANDERSEN AND A. C. KAK, *Simultaneous algebraic reconstruction technique (SART): A superior implementation of the ART algorithm*, Ultrasonic Imaging, 6 (1984), pp. 81–94.
- [2] S. BASU AND Y. BRESLER, *$O(N^3 \log N)$ backprojection algorithm for the 3-D Radon transform*, IEEE Trans. Med. Imaging, 21 (2002), pp. 76–88.
- [3] A. BIGURI, M. DOSANJH, S. HANCOCK, AND M. SOLEIMANI, *TIGRE: A MATLAB-GPU toolbox for CBCT image reconstruction*, Biomed. Phys. Eng. Express, 2 (2016), 055010.
- [4] C. BOUMAN AND K. SAUER, *A generalized Gaussian image model for edge-preserving MAP estimation*, IEEE Trans. Image Process., 2 (1993), pp. 296–310.
- [5] M. L. BRADY, *A fast discrete approximation algorithm for the Radon transform*, SIAM J. Comput., 27 (1998), pp. 107–119.
- [6] A. BRANDT, J. MANN, M. BRODSKI, AND M. GALUN, *A fast and accurate multilevel inversion of the Radon transform*, SIAM J. Appl. Math., 60 (2000), pp. 437–462.
- [7] M. H. BYONOCORE, W. R. BRODY, AND A. MACOVSKI, *A natural pixel decomposition for two-dimensional image reconstruction*, IEEE Trans. Biomed. Eng., 28 (1981), pp. 69–78.
- [8] C. CHAPDELAIN, N. GAC, A.-M. DJAFARI, AND E. PARRA-DENIS, *New GPU implementation of separable footprint (SF) projector and backprojector: First results*, in Proceedings of the Fifth International Conference on Image Formation in X-Ray Computed Tomography, 2018, pp. 314–317.

- [9] C. DE BOOR, K. HÖLLIG, AND S. RIEMENSCHNEIDER, *Box Splines*, Appl. Math. Sci. 98, Springer-Verlag, Berlin, 1993.
- [10] B. DE MAN AND S. BASU, *Distance-driven projection and backprojection in three dimensions*, Phys. Med. Biol., 49 (2004), 2463.
- [11] B. DE MAN, J. NUYTS, P. DUPONT, G. MARCHAL, AND P. SUETENS, *Metal streak artifacts in X-ray computed tomography: A simulation study*, in Proceedings of the Nuclear Science Symposium and Medical Imaging Conference, Vol. 3, IEEE, 1998, pp. 1860–1865.
- [12] D. L. DONOHO AND O. LEVI, *Fast X-ray and beamlet transforms for three-dimensional data*, Modern Signal Process., 46 (2002), pp. 79–116.
- [13] A. ENTEZARI, M. NILCHIAN, AND M. UNSER, *A box spline calculus for the discretization of computed tomography reconstruction problems*, IEEE Trans. Med. Imaging, 31 (2012), pp. 1532–1541.
- [14] A. ENTEZARI AND M. UNSER, *A box spline calculus for computed tomography*, in Proceedings of the International Symposium on Biomedical Imaging: From Nano to Macro, IEEE, 2010, pp. 600–603.
- [15] A. ENTEZARI, D. VAN DE VILLE, AND T. MÖLLER, *Practical box splines for reconstruction on the body centered cubic lattice*, IEEE Trans. Vis. Comput. Graphics, 14 (2008), pp. 313–328.
- [16] H. ERDOGAN AND J. A. FESSLER, *Ordered subsets algorithms for transmission tomography*, Phys. Med. Biol., 44 (1999), pp. 2835–2851.
- [17] J. A. FESSLER, *Statistical image reconstruction methods for transmission tomography*, Handb. Med. Imaging, 2 (2000), pp. 1–70.
- [18] J. A. FESSLER, *Fundamentals of CT reconstruction in 2D and 3D*, in Comprehensive Biomedical Physics, Vol. 2, Elsevier, Amsterdam, 2014, pp. 263–295.
- [19] J. A. FESSLER, *Michigan Image Reconstruction Toolbox (MIRT)*, <http://web.eecs.umich.edu/~fessler/irt/irt>, 2014.
- [20] J. A. FESSLER, *Medical Image Reconstruction: A Brief Overview of Past Milestones and Future Directions*, <https://arxiv.org/abs/1707.05927>, 2017
- [21] B. FINKBEINER, A. ENTEZARI, D. VAN DE VILLE, AND T. MÖLLER, *Efficient volume rendering on the body centered cubic lattice using box splines*, Comput. Graph., 34 (2010), pp. 409–423.
- [22] R. GARCÍA-MOLLÁ, R. LINARES, AND R. AYALA, *A study of DQE dependence with beam quality, GE Senographe essential detector for mammography*, in World Congress on Medical Physics and Biomedical Engineering, Springer, Munich, 2009, pp. 886–889.
- [23] L. L. GEYER, U. J. SCHOEPF, F. G. MEINEL, J. W. NANCE, JR., G. BASTARRIKA, J. A. LEIPSIC, N. S. PAUL, M. RENGO, A. LAGHI, AND C. N. DE CECCO, *State of the art: Iterative CT reconstruction techniques*, Radiology, 276 (2015), pp. 339–357.
- [24] G. GLOVER AND N. PELC, *Nonlinear partial volume artifacts in x-ray computed tomography*, Med. Phys., 7 (1980), pp. 238–248.
- [25] S. HA AND K. MUELLER, *A look-up table-based Ray integration framework for 2-D/3-D Forward and back projection in X-Ray CT*, IEEE Trans. Med. Imaging, 37 (2018), pp. 361–371.
- [26] S. HA AND K. MUELLER, *Look-Up Table-Based Ray Integration Framework (LTTRI)*, <https://github.com/bsmind/LTRI>, 2018
- [27] G. T. HERMAN, *Fundamentals of Computerized Tomography: Image Reconstruction from Projections*, Springer, Berlin, 2009.
- [28] T. HIGAKI, Y. NAKAMURA, J. ZHOU, Z. YU, T. NEMOTO, F. TATSUGAMI, AND K. AWAI, *Deep learning reconstruction at CT: Phantom study of the image characteristics*, Acad. Radiol., 27 (2020), pp. 82–87.
- [29] J. HSIEH, *Computed Tomography: Principles, Design, Artifacts, and Recent Advances*, Wiley, New York, 2003.
- [30] P. M. JOSEPH, *An improved algorithm for reprojecting rays through pixel images*, IEEE Trans. Med. Imaging, 1 (1982), pp. 192–196.
- [31] P. M. JOSEPH AND R. SPITAL, *The exponential edge-gradient effect in x-ray computed tomography*, Phys. Med. Biol., 26 (1981), pp. 473–487.
- [32] E. KANG, J. MIN, AND J. C. YE, *A deep convolutional neural network using directional wavelets for low-dose X-ray CT reconstruction*, Med. Phys., 44 (2017), pp. e360–e375.

- [33] G. M. LASIO, B. R. WHITING, AND J. F. WILLIAMSON, *Statistical reconstruction for x-ray computed tomography using energy-integrating detectors*, Phys. Med. Biol., 52 (2007), pp. 2247–2266.
- [34] R. M. LEWITT, *Multidimensional digital image representations using generalized Kaiser-Bessel window functions*, J. Optical Soc. Amer., 7 (1990), pp. 1834–1846.
- [35] W. LI, G. T. BUZZARD, AND C. A. BOUMAN, *Direct sparse-view CT reconstruction using LSTM processing of stacked back projections*, in Proceedings of the International Conference on Computational Photography, 2020.
- [36] X. LI, Y. LIANG, W. ZHANG, T. LIU, H. LI, G. LUO, AND M. JIANG, *cuMBIR: An efficient framework for low-dose x-ray CT image reconstruction on GPUs*, in Proceedings of the International Conference on Supercomputing, 2018, pp. 184–194.
- [37] S.-C. LO, *Strip and line path integrals with a square pixel matrix: A unified theory for computational CT projections*, IEEE Trans. Med. Imaging, 7 (1988), pp. 355–363.
- [38] Y. LONG, J. A. FESSLER, AND J. M. BALTER, *3D forward and back-projection for X-ray CT using separable footprints*, IEEE Trans. Med. Imaging, 29 (2010), pp. 1839–1850.
- [39] C. H. MCCOLLOUGH, G. H. CHEN, W. KALENDER, S. LENG, E. SAMEI, K. TAGUCHI, G. WANG, L. YU, AND R. I. PETTIGREW, *Achieving routine submillisievert CT scanning: Report from the summit on management of radiation dose in CT*, Radiology, 264 (2012), pp. 567–580.
- [40] F. MOMEY, L. DENIS, C. BURNIER, E. THIÉBAUT, J.-M. BECKER, AND L. DESBAT, *Spline driven: High accuracy projectors for tomographic reconstruction from few projections*, IEEE Trans. Image Process., 24 (2015), pp. 4715–4725.
- [41] K. MUELLER AND F. XU, *Practical considerations for GPU-accelerated CT*, in Proceedings of the 3rd International Symposium on Biomedical Imaging: Nano to Macro, IEEE, 2006, pp. 1184–1187.
- [42] NATIONAL LUNG SCREENING TRIAL RESEARCH TEAM, *Reduced lung-cancer mortality with low-dose computed tomographic screening*, New Engl. J. Med., 365 (2011), pp. 395–409.
- [43] F. NATTERER, *The Mathematics of Computerized Tomography*, SIAM, Philadelphia, 2001.
- [44] M. NILCHIAN, J. P. WARD, C. VONESCH, AND M. UNSER, *Optimized Kaiser-Bessel window functions for computed tomography*, IEEE Trans. Image Process., 24 (2015), pp. 3826–3833.
- [45] X. PAN, E. Y. SIDKY, AND M. VANNIER, *Why do commercial CT scanners still employ traditional, filtered back-projection for image reconstruction?*, Inverse Problems, 25 (2009), 123009.
- [46] A. SABNE, X. WANG, S. J. KISNER, C. A. BOUMAN, A. RAGHUNATHAN, AND S. P. MIDKIFF, *Model-based iterative CT image reconstruction on GPUs*, ACM SIGPLAN Notices, 52 (2017), pp. 207–220.
- [47] K. SAUER AND C. BOUMAN, *Bayesian estimation of transmission tomograms using segmentation based optimization*, IEEE Trans. Nuclear Sci., 39 (1992), pp. 1144–1152.
- [48] E. Y. SIDKY, R. CHARTRAND, Y. DUCHIN, C. ULLBERG, AND X. PAN, *High resolution image reconstruction with constrained, total-variation minimization*, in Proceedings of the Nuclear Science Symposium & Medical Imaging Conference, IEEE, 2010, pp. 2617–2620.
- [49] E. Y. SIDKY, Y. DUCHIN, X. PAN, AND C. ULLBERG, *A constrained, total-variation minimization algorithm for low-intensity x-ray CT*, Med. Phys., 38 (2011), pp. S117–S125.
- [50] E. Y. SIDKY AND X. PAN, *Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization*, Phys. Med. Biol., 53 (2008), pp. 4777–4807.
- [51] M. TABEI AND M. UEDA, *Backprojection by upsampled Fourier series expansion and interpolated FFT*, IEEE Trans. Image Process., 1 (1992), pp. 77–87.
- [52] S. TILLEY, M. JACOBSON, Q. CAO, M. BREHLER, A. SISNIEGA, W. ZBIJEWSKI, AND J. W. STAYMAN, *Penalized-likelihood reconstruction with high-fidelity measurement models for high-resolution cone-beam imaging*, IEEE Trans. Med. Imaging, 37 (2017), pp. 988–999.
- [53] G. WANG, *A perspective on deep imaging*, IEEE Access, 4 (2016), pp. 8914–8924.
- [54] G. WANG, J. C. YE, K. MUELLER, AND J. A. FESSLER, *Image reconstruction is a new frontier of machine learning*, IEEE Trans. Med. Imaging, 37 (2018), pp. 1289–1296.
- [55] X. WANG, A. SABNE, P. SAKDHNAGOO, S. J. KISNER, C. A. BOUMAN, AND S. P. MIDKIFF, *Massively parallel 3D image reconstruction*, in Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, ACM, 2017, 3.

- [56] B. R. WHITING, P. MASSOUMZADEH, O. A. EARL, J. A. O'SULLIVAN, D. L. SNYDER, AND J. F. WILLIAMSON, *Properties of preprocessed sinogram data in x-ray computed tomography*, Med. Phys., 33 (2006), pp. 3290–3303.
- [57] M. J. WILLEMINK AND P. B. NOËL, *The evolution of image reconstruction for CT from filtered back projection to artificial intelligence*, Eur. Radiol., 29 (2019), pp. 2185–2195.
- [58] T. WÜRFL, M. HOFFMANN, V. CHRISTLEIN, K. BREININGER, Y. HUANG, M. UNBERATH, AND A. K. MAIER, *Deep learning computed tomography: Learning projection-domain weights from image domain in limited angle problems*, IEEE Trans. Med. Imaging, 37 (2018), pp. 1454–1463.
- [59] G. YAN, J. TIAN, S. ZHU, Y. DAI, AND C. QIN, *Fast cone-beam CT image reconstruction using GPU hardware*, J. X-Ray Sci. Technol., 16 (2008), pp. 225–234.
- [60] D. H. YE, G. T. BUZZARD, M. RUBY, AND C. A. BOUMAN, *Deep back projection for sparse-view CT reconstruction*, in Proceedings of the Global Conference on Signal and Information Processing, IEEE, 2018, pp. 1–5.
- [61] D. H. YE, S. SRIVASTAVA, J.-B. THIBAUT, K. SAUER, AND C. BOUMAN, *Deep residual learning for model-based iterative CT reconstruction using plug-and-play framework*, in Proceedings of the International Conference on Acoustics, Speech and Signal Processing, IEEE, 2018, pp. 6668–6672.
- [62] H. YU AND G. WANG, *Finite detector based projection model for high spatial resolution*, J. X-Ray Sci. Technol., 20 (2012), pp. 229–238.
- [63] Z. YU, F. NOO, F. DENNERLEIN, A. WUNDERLICH, G. LAURITSCH, AND J. HORNEGGER, *Simulation tools for two-dimensional experiments in x-ray computed tomography using the FORBILD head phantom*, Phys. Med. Biol., 57 (2012), pp. N237–52, <https://doi.org/10.1088/0031-9155/57/13/N237>.
- [64] K. ZHANG AND A. ENTEZARI, *Box spline projection in non-parallel geometry*, in Proceedings of the 16th International Symposium on Biomedical Imaging, IEEE, 2019, pp. 1844–1847.
- [65] K. ZHANG AND A. ENTEZARI, *A convolutional framework for forward and back-projection in fan-beam geometry*, in Proceedings of the 16th International Symposium on Biomedical Imaging, IEEE, 2019, pp. 1455–1458.
- [66] B. ZHU, J. Z. LIU, S. F. CAULEY, B. R. ROSEN, AND M. S. ROSEN, *Image reconstruction by domain-transform manifold learning*, Nature, 555 (2018), pp. 487–492.