# Fast projection/backprojection and incremental methods applied to synchrotron light tomographic reconstruction

**Camila de Lima and Elias Salomão Helou***

Institute of Mathematical Science and Computation, University of São Paulo, SP, Brazil.
*Correspondence e-mail: elias@icmc.usp.br

Iterative methods for tomographic image reconstruction have the computational cost of each iteration dominated by the computation of the (back)projection operator, which take roughly $O(N^3)$ floating point operations (flops) for $N \times N$ pixels images. Furthermore, classical iterative algorithms may take too many iterations in order to achieve acceptable images, thereby making the use of these techniques unpractical for high-resolution images. Techniques have been developed in the literature in order to reduce the computational cost of the (back)projection operator to $O(N^2 \log N)$ flops. Also, incremental algorithms have been devised that reduce by an order of magnitude the number of iterations required to achieve acceptable images. The present paper introduces an incremental algorithm with a cost of $O(N^2 \log N)$ flops per iteration and applies it to the reconstruction of very large tomographic images obtained from synchrotron light illuminated data.

## 1. Introduction

One of the approaches used in tomographic image reconstruction is to consider that the function to be reconstructed (a tomographic image is modeled as a function from the plane to the set of non-negative real numbers) lies in a finite dimensional space and then to solve the resulting linear system of equations. Iterative techniques may be required because of the very large dimensions of the system, which may have a coefficient matrix reaching sizes of $10^7 \times 10^7$ or more, and because of the unstructured and sparse nature of the coefficients matrix. In general, noise in images obtained by such techniques is smaller when compared against images reconstructed by the FBP algorithm (Bracewell, 1965). On the other hand, image quality can still be affected by low photon counts and other sources of noise in the data, allied to the ill-conditioning of the system matrix (Helou, 2009).

In order to reduce the effects of poor photon counts, more advanced strategies have been developed, such as the statistical model, also known as the maximum-likelihood (ML) (Vardi *et al.*, 1985) model, which advocates the maximization of the likelihood that the data have been generated by the reconstructed image. ML models in tomographic reconstruction have first been introduced for the emission tomography problem and the expectation maximization (EM) algorithm was proposed for its solution (Shepp & Vardi, 1982) with good resulting image quality. Despite the improvements in image quality, application of the EM algorithm could not be made practical because of the large amount of iterations required to obtain reasonable reconstructions.
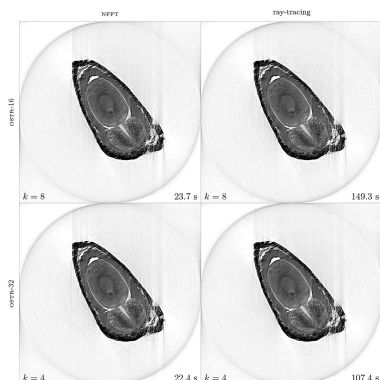
Therefore, Hudson & Larkin (1994) introduced an EM-based technique, the ordered subsets expectation maximization (OSEM) algorithm, which processes subsets of the data at each step within the iteration, updating the image in between the processing of each of the data subsets. OSEM brought a major speed-up to maximum-likelihood techniques and, while not having a fully satisfying convergence theory to support its use, it pioneered the use of incremental techniques in the realm of maximum-likelihod tomographic image reconstruction and spurred a host of new efficient methods for tomographic reconstruction through the ML model. Among these, we can mention: the row-action maximum-likelihood algorithm (RAMLA) (Browne & De Pierro, 1996), which processes one line at a time between updates and has the advantage of being convergent; dynamic RAMLA (DRAMA), proposed by Tanaka & Kudo (2003), whose convergence was demonstrated by Helou & De Pierro (2005); block sequential regularized expectation maximization (BSREM), an extension of RAMLA which uses the maximum *a posteriori* (MAP) regularization technique (De Pierro & Yamagishi, 2001); string-averaging expectation-maximization for maximum-likelihood (SAEM), an extension of EM that uses string-averaging (Helou *et al.*, 2014); and ordered subsets for transmission tomography (OSTR), which applies the idea of data subdivided into ordered subsets to separable paraboloidal surrogates (SPS) (Erdoğan & Fessler, 1999). OSTR, unlike the other methods mentioned in the present paragraph, is meant to be used for transmission tomography imaging.

Non-incremental algorithms such as the EM method require the projection (the forward operator in tomography) and its adjoint to be computed on every iteration, which amounts to $O(N^3)$ floating point operations (flops) for $N \times N$ images under reasonable data sampling using traditional on-the-fly algorithms for the computation, including straightforward ray-tracing techniques to trace the voxels along a certain projection ray (Siddon, 1985; Han & You, 1999) and the slant stack, where each projection can be obtained by summing the columns of a image slanted by an angle (Thorson, 1978; Hawkes, 2006). In order to decrease the per-iteration overhead of these algorithms, several fast techniques have been devised to reduce this figure to $O(N^2 \log N)$ flops, such as the fast slant stack (Averbuch *et al.*, 2008), the hierarchical decomposition algorithm (George & Bresler, 2007), the technique based on Fourier in log-polar grid (Andersson, 2005) and the non-equispaced fast Fourier transform (NFFT) (Fessler & Sutton, 2003; Potts & Steidl, 2000).

To the best of our knowledge, however, there is no work in the literature merging the ideas from the previous two paragraphs. That is, no incremental algorithm with $O(N^2 \log N)$ complexity seems to have been studied so far. The main contribution of the

present paper is to show that this is indeed a practical possibility by applying the concept to tomographic reconstruction of large transmission tomography images from data obtained by synchrotron light illumination.

## 2. Tomographic reconstruction problem

One of the fundamental mathematical concepts in tomography is the Radon transform (RT), formulated by Johann Radon in 1917 (Radon, 1986). The image reconstruction problem in tomography is to recover a function $f : \mathbb{R}^2 \to \mathbb{R}$ from its arc length integrals along straight lines. Thus, $f$ is to be determined from its Radon transform $\mathcal{R}[f]$:

$$\mathcal{R}[f](\theta, t) := \int_{\mathbb{R}} f\left[ t \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} + s \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix} \right] ds. \quad (1)$$

We will also use the alternative notation $\mathcal{R}_\theta[f](t) := \mathcal{R}[f](\theta, t)$. The function $p_\theta = \mathcal{R}_\theta[f]$ will be called the projection with relation to the angle $\theta$. A geometric representation of the RT is given on the left in Fig. 1. In this figure, a Sheep–Logan phantom, which is an image composed of ten ellipses described by Kak & Slaney (1988), is centralized in axes $x$ and $y$. The $t$ axis, whose slope is determined by the angle $\theta$, is also shown. For the point $t = t'$, the perpendicular dashed line represents the integration path, and the graph of $\mathcal{R}_\theta[f](t)$ is plotted. The representation of the RT in the plane $\theta \times t$ is called the sinogram. The sinogram of the Sheep–Logan phantom is presented on the right in Fig. 1.

An alternative to obtain the inverse of the Radon transform $\mathcal{R}_\theta^{-1}[f](t)$ in order to reconstruct the image $f$ is the Fourier slice theorem (FST). This is an important result that relates a projection to the image in the Fourier space. The Fourier transform (FT) definition, some considerations and this theorem are presented in the following subsection.

### 2.1. Fourier slice theorem

The representation of a function in the Fourier or frequency space is given by the Fourier transform. Denoting by $\langle \omega, \mathbf{x} \rangle$ the inner product between the vectors $\omega, \mathbf{x} \in \mathbb{R}^n$, the Fourier
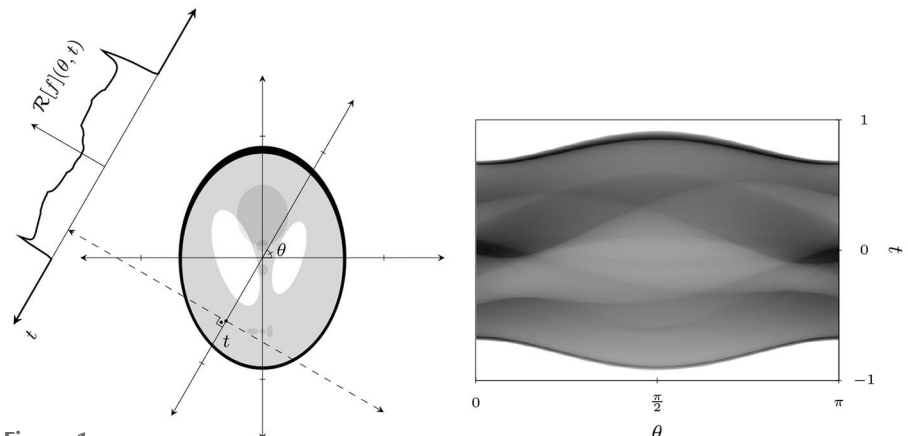


**Figure 1**
Geometric representation of the RT.

transform $\mathcal{F}[f]$ (or $\hat{f}$) of a function $f : \mathbb{R}^n \to \mathbb{C}$ is defined as follows,

$$\mathcal{F}[f](\boldsymbol{\omega}) = \int\limits_{\mathbb{R}^n} f(\mathbf{x}) \exp(-\iota\langle\boldsymbol{\omega}, \mathbf{x}\rangle)\, d\mathbf{x}, \qquad (2)$$

where $\iota = \sqrt{-1}$. In turn, the inverse Fourier transform $\mathcal{F}^{-1}[f]$ of the function $f$ is defined by the expression

$$\mathcal{F}^{-1}[f](\mathbf{x}) = \frac{1}{(2\pi)^n} \int\limits_{\mathbb{R}^n} f(\boldsymbol{\omega}) \exp(\iota\langle\mathbf{x}, \boldsymbol{\omega}\rangle)\, d\boldsymbol{\omega}.$$

If $f, \hat{f} \in L^1_{\mathbb{R}^n}$, then the inverse Fourier transform retrieves the original function such that $\mathcal{F}^{-1}[\hat{f}] = f$.

Fourier analysis is important in image reconstruction from projections due to the Fourier slice theorem, which states that the Fourier transform of a projection with angle $\theta$ is equal to a slice with the same angle of the Fourier transform of the image (Herman, 1980; Natterer, 1986; Kak & Slaney, 1988).

*Theorem 1 (Fourier slice theorem).* Let $f : \mathbb{R}^2 \to \mathbb{C}$ be defined such that $f \in L^1_{\mathbb{R}^2}$ and $\omega \in \mathbb{R}$ follows

$$\hat{p}_\theta(\omega) = \hat{f}(\omega\cos\theta, \omega\sin\theta). \qquad (3)$$

Therefore, determining $\hat{p}_\theta(\omega)$ for any $(\theta, \omega)$ allows $\hat{f}$ to be known at any point. Using the inverse Fourier transform it is then possible to reconstruct the image. Since only a finite number of samples are determined, other samples of the function $f$ in the frequency space can be obtained by interpolating the radial samples. Methods that use this strategy and then reconstruct the function $f$ through the inverse Fourier transform are known as Fourier methods. However, they have inferior accuracy when compared with iterative methods, which are presented in the following sections. This is caused because higher frequencies are insufficiently sampled due to the sampling distribution of the Fourier domain being much denser near the origin, and higher frequencies contain the finer details of the image.

## 2.2. Discrete model

Considering that the function to be recovered lies on a finite dimensional space generated by the base $\{f_1, f_2, \ldots, f_n\}$, the aim becomes to find a vector $x$ such that the function $f = \sum_{i=j}^n x_j f_j$ satisfies the equality (1). Knowing that the dataset provided by the tomographic scanner consists of a finite number of samples of the RT, we discretize the problem as

$$\mathbf{Rx} = \mathbf{b}, \qquad (4)$$

where $\mathbf{x} \in \mathbb{R}^n$ determines the desired image, $\mathbf{b} \in \mathbb{R}^m$ contains the (approximate) RT samples $b_i \approx \mathcal{R}[f](\theta_i, t_i)$, $i \in \{1, 2, \ldots, m\}$, obtained during the tomographic scan, and $\mathbf{R} \in \mathbb{R}^{m \times n}$ is the discretized RT, with coefficients given as

$$r_{ij} = \mathcal{R}[f_j](\theta_i, t_i). \qquad (5)$$

One of the existing strategies for solving this problem is presented in the following section.

## 3. Statistical model for transmission tomography

Following Erdoğan & Fessler (1999), the negative log-likelihood function for independent transmission data is given by

$$-L(\mathbf{x}) = \sum_{i=1}^m h_i\big[(\mathbf{Rx})_i\big], \qquad (6)$$

where

$$h_i(l) = \varphi_i \exp(-l) + d_i - \rho_i \log\big[\varphi_i \exp(-l) + d_i\big], \qquad (7)$$

$d_i$ is the mean number of detected background photons, $\varphi_i$ is the blank scan photon count and $\rho_i$ is the tomographic scan photon count. Thus, the model consists of solving

$$\min_{\mathbf{x} \in \mathbb{R}^n_+} -L(\mathbf{x}). \qquad (8)$$

In the above statement, the function $-L$ to be minimized is also known as the objective function of the minimization problem.

The next subsection presents the method of ordered subsets for transmission tomography aiming at solving the problem proposed above. Before proceeding, note that

$$h_i'(l) = \left[\frac{\rho_i}{\varphi_i \exp(-l) + d_i} - 1\right] \varphi_i \exp(-l).$$

These derivatives will be used to simplify exposition of the algorithm in the following.

## 3.1. Ordered subsets for transmission tomography

The ordered subsets for the transmission tomography method (OSTR), proposed by Erdoğan & Fessler (1999), arises from the application of the ordered subsets principle to the separable paraboloidal surrogates algorithm (SPS) (Erdoğan & Fessler, 1998), which can be used to solve the statistical model for transmission tomography without penalty defined by (8). Again, the index set is divided into $s$ subsets such that $\bigcup_{i=1}^s U_i = \{1, 2, \ldots, m\}$ and $U_i \cap U_j = \emptyset$ if $i \neq j$. The iterative procedure of the algorithm is presented as follows:

> **Input:** $\boldsymbol{x}^{(0)} \in \mathbb{R}^n_+$ and $U_1, \ldots, U_s$.
> **Output:** $\boldsymbol{x}^{(\bar{k})}$.
> 1   for $i \leftarrow 1$ to $m$ do
> 2     $\gamma_i = \sum_{j=1}^n r_{ij}$
> 3   end
> 4   for $j \leftarrow 1$ to $n$ do
> 5     $d_j^* = \sum_{i=1}^m r_{ij} \gamma_i \dfrac{(\rho_i - d_i)^2}{\rho_i}$
> 6   end
> 7   for $k \leftarrow 1$ to $\bar{k}$ do
> 8     $\boldsymbol{x}^{(k,0)} = \boldsymbol{x}^{(k)}$
> 9     for $l \leftarrow 1$ to $s$ do
> 10       for $j \leftarrow 1$ to $n$ do
> 11         $x_j^{(k,l)} = x_j^{(k,l-1)} - \dfrac{s \sum_{i \in U_l} r_{ij} h_i'\big((\boldsymbol{Rx}^{(k,l-1)})_i\big)}{d_j^*}$
> 12       end
> 13     end
> 14     $\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k,s)}$
> 15   end

The pre-computations of $\gamma_i$ and $d_j^*$ in lines 1–6 require the calculation of a projection and its adjoint once before we can start iterating algorithm OSTR. Partial versions of these

calculations are repeated in each sub-iteration of the same algorithm at lines 10–13. We will make use of NFFT techniques in order to compute these operators efficiently.

### 3.2. Fast iterative shrinkage thresholding algorithm

The fast iterative shrinkage thresholding algorithm (FISTA), proposed by Beck & Teboulle (2009a,b), is a modification of the iterative shrinkage thresholding algorithm (ISTA), which is used to solve linear inverse problems. ISTA is similar to the classical gradient method and, although it is known for solving large-scale problems in a simple and practical way, it converges slowly. FISTA retains the computational simplicity of ISTA but substantially improves its convergence rate. FISTA is presented in order to compare its results with OSTR. Notice that application of fast Radon operators to FISTA is immediate, but the algorithm is not as fast as incremental methods in the first iterations, which motivates our research.

The general model to be solved by FISTA is

$$\min_{\mathbf{x} \in \mathbb{R}^n} \Psi(\mathbf{x}) := \varpi(\mathbf{x}) + \varrho(\mathbf{x}), \tag{9}$$

where $\varrho : \mathbb{R}^n \to \mathbb{R}$ is a non-smooth convex function and $\varpi : \mathbb{R}^n \to \mathbb{R}$ is a smooth differentiable convex function with Lipschitz gradient. For our purposes, we will use $\varpi = -L$ [from (6)–(7)] and $\varrho = \chi_{\mathbb{R}^n_+}$, where for a given non-empty closed convex set $X$ the indicator function $\chi_X$ is given by

$$\chi_X(\mathbf{x}) := \begin{cases} 0 & \text{if } \mathbf{x} \in X, \\ \infty & \text{if } \mathbf{x} \notin X. \end{cases}$$

Given this, unconstrained minimization $\varpi(\mathbf{x}) + \chi_{\mathbb{R}^n_+}$ is equivalent to minimizing $\varpi(\mathbf{x})$ constrained to $\mathbf{x} \in \mathbb{R}^n_+$. The steps of the method for the case $\rho = \chi_{\mathbb{R}^n_+}$ are described below:

**Input:** $\boldsymbol{x}^{(0)} \in \mathbb{R}^n_+$, $T > 0$.
**Output:** $\boldsymbol{y}^{(\bar{k}+1)}$.
1   $\boldsymbol{y}^{(1)} = \boldsymbol{x}^{(0)}$, and $t^{(1)} = 1$
2   **for** $k \leftarrow 1$ **to** $\bar{k}$ **do**
3     $\boldsymbol{x}^{(k)} = \mathcal{P}_{\mathbb{R}^n_+}\left(\boldsymbol{y}^{(k)} - \dfrac{1}{T}\nabla\varpi(\boldsymbol{y}^{(k)})\right)$
4     $t^{(k+1)} = \dfrac{1 + \sqrt{1 + 4(t^{(k)})^2}}{2}$
5     $\boldsymbol{y}^{(k+1)} = \boldsymbol{x}^{(k)} + \left(\dfrac{t^{(k)}-1}{t^{(k+1)}}\right)\left(\boldsymbol{x}^{(k)} - \boldsymbol{x}^{(k-1)}\right)$
6   **end**

## 4. Non-equispaced fast Fourier transform

Non-equispaced fast Fourier transform (NFFT) algorithms are used to perform the non-equispaced discrete Fourier transform (NDFT) quickly, which is a generalization of the discrete Fourier transform (DFT) from equally spaced to arbitrary sampling points or spatial nodes. Most NFFT methods are based on the calculation of the fast Fourier transform (FFT) (Keiner *et al.*, 2009; Fourmont, 2003; Fessler & Sutton, 2003) in order to obtain the NDFT.

We briefly discuss the method here in order to obtain, coupling it with the Fourier slice theorem previously presented, a fast and efficient way to calculate the partial projections and backprojections that are used in the sub-iterations of OSEM, OSTR and other incremental methods. Thus, these computations can be incorporated in the execution of the aforementioned methods to achieve quality images under a reasonable computational effort even for very large image sizes.

We will discuss, for simplicity, the one-dimensional case, but the generalization for higher-dimensional domains follow the same principle (Keiner *et al.*, 2009; Fourmont, 2003; Fessler & Sutton, 2003). There are other related problems, such as computing the equispaced Fourier coefficients from non-equispaced data or the more general computation of non-equispaced Fourier coefficients from non-equispaced data. However, for tomography, the important case is to compute non-equispaced Fourier coefficients from equispaced data. That is, given samples $\psi_i = f(x_i)$ with $i \in \{-N/2, -N/2+1, \ldots, N/2-1\}$, one wishes to compute

$$\hat{\psi}_k = \sum_{i=-N/2}^{N/2-1} \psi_i \exp(-\iota\xi_k x_i), \tag{10}$$

$\xi_k \in [-\pi, \pi)$, $k \in \{-N/2, -N/2+1, \ldots, N-1\}$.

The main principle of the NFFT algorithm is to use FFTs (Cooley & Tukey, 1965) to find an approximation of the trigonometric function (10). FFTs are algorithms for the effective computation of the discrete Fourier transform at samples $\xi_k = 2\pi k/N$, $k \in \{-N/2, -N/2+1, \ldots, N/2-1\}$. In order to achieve non-equispaced DFTs from FFT-computable equispaced results, a key tool is the following (from Fourmont, 2003):

*Proposition 1.* Given $c > 0$, let $0 < \pi/c < \alpha$ and $\alpha < \pi(2 - 1/c)$. Let also $\omega : \mathbb{R} \to \mathbb{C}$ be continuous and piece-wise continuously differentiable in $[-\alpha, \alpha]$, vanishing outside $[-\alpha, \alpha]$ and non-zero in $[-\pi/c, \pi/c]$. Then

$$\exp(-\iota x\xi) = \frac{1}{\omega(x)\sqrt{2\pi}}\sum_{j\in\mathbb{Z}}\hat{\omega}(\xi - j)\exp(-\iota jx).$$

In the above statement, $\hat{\omega}$, for a given $\omega : \mathbb{R}^n \to \mathbb{C}$, is the continuous Fourier transform as given by (2). Returning to the computation of the NDFT using FFTs, we denote $\eta_k = (N\xi_k/2\pi) \in [-N/2, N/2-1]$ (not necessarily equispaced) and use *Proposition 1* with $x_i = 2\pi i/cN$ in formula (10) to obtain

$$\hat{\psi}_k = \sum_{i=-N/2}^{N/2-1} \psi_i \exp(-\iota\xi_k x_i) = \sum_{i=-N/2}^{N/2-1} \psi_i \exp\left[-\iota\left(\frac{2\pi i}{N}\right)\eta_k\right]$$

$$= \frac{1}{\sqrt{2\pi}}\sum_{i=-N/2}^{N/2-1}\frac{\psi_i}{\omega(2\pi i/cN)}\sum_{j\in\mathbb{Z}}\hat{\omega}(c\eta_k - j)\exp\left[-\iota\left(\frac{2\pi i}{cN}\right)j\right]$$

$$= \frac{1}{\sqrt{2\pi}}\sum_{j\in\mathbb{Z}}\hat{\omega}(c\eta_k - j)\sum_{i=-N/2}^{N/2-1}\frac{\psi_i}{\omega(2\pi i/cN)}\exp\left[-\iota\left(\frac{2\pi i}{cN}\right)j\right].$$

In this formula, the inner summation can be computed using FFTs of length $cN$. If the function $\omega$ is chosen so that its decay is sufficiently fast when away from 0, few FFTs are necessary for results with a good precision. For a discussion on appro-

priate window functions $\omega$, see Keiner *et al.* (2009) and Fourmont (2003).

In the present paper paper, we used the NFFT 3 free library to perform the above computations in order to obtain fast projection and backprojection operators and use them in iterations and subiterations of the methods FISTA and OSTR described above.

### 4.1. NFFT and the Fourier slice theorem

The tomographic dataset provided by the experimental setup at the LNLS (Brazilian Synchrotron Light Laboratory) has the following form,

$$\mathcal{R}[f](\theta_\kappa, t_\ell), \quad (\kappa, \ell) \in \{1, 2, \ldots, n_\theta\} \times \{1, 2, \ldots, n_t\}, \quad (11)$$

where

$$\theta_\kappa = \pi \frac{(\kappa - 1)}{n_\theta - 1} \quad \text{and} \quad t_\ell = -1 + 2 \frac{(\ell - 1)}{n_t - 1}. \quad (12)$$

Thus, the steps required to approximate the necessary samples of the discrete Radon transform represented by **Rx** are the following: (i) compute, using a bidimensional NFFT routine, the samples

$$\mathcal{F}\Big[\sum_j f_j x_j\Big]\Big[t_\ell \begin{pmatrix} \cos \theta_\kappa \\ \sin \theta_\kappa \end{pmatrix}\Big];$$

(ii) for each fixed $\kappa$, set

$$\hat{\boldsymbol{\psi}}_\ell^\kappa = \mathcal{F}\Big[\sum_j f_j x_j\Big]\Big[t_\ell \begin{pmatrix} \cos \theta_\kappa \\ \sin \theta_\kappa \end{pmatrix}\Big]$$

and compute the one-dimensional inverse FFT of each vector $\hat{\boldsymbol{\psi}}^\kappa$.

According to *Theorem 1*, vector $\hat{\boldsymbol{\psi}}^\kappa$ contains samples of the Fourier transform $\hat{p}_{\theta_\kappa}$ of $p_{\theta_\kappa}$. Therefore, after computing its inverse FFT, we get samples of $p_{\theta_i}$. This is obtained with a $O(n_\theta n_t \log n_t + N^2 \log N)$ computational complexity, where the image is supposed to have $N \times N$ pixels. Then, since in practice we use $n_t \approx n_\theta \approx N$, we actually use $O(N^2 \log N)$ flops to complete this operation. The adjoint operation can be obtained by computing the adjoint of each step, each of which is linear, in the reverse order of the direct transform.

## 5. Computational experimentation

In this section we present the numerical experiments that were performed in order to ascertain the effectiveness of the proposed methodology. We experiment testing two different techniques for computing **Rx** and its adjoint: the ray-tracing method of Han & You (1999) and the NFFT technique described in the previous section. Each of the two different forms of computing the discrete Radon transform was used in two iterative algorithms: OSTR and FISTA, as described in §3.1 and §3.2. We had, therefore, in principle four combinations of methodologies to evaluate. However, in some experiments, the number of subsets used with OSTR was varied in order to verify how this parameter affects the method's behavior.

The experiments used both synthetic and real-world data. Synthetic data were used to evaluate the speed, as the dataset size varies, and accuracy, under ideal and noisy data acquisition schemes, of the methods. Practical data were used to evaluate the algorithms in practical circumstances that appear in the LNLS tomographic beamline. The remainder of the present section details the data simulation and collection procedures, details the algorithmic parameters setup we have used and reports the results of the reconstruction methods.

In order to make a more comprehensive assessment of the computational characteristics of the proposed methodologies, two different hardware setups were used in the experiments: one off-the-shelf high-end laptop computer featuring 32 GB of RAM running an Intel Core i7-7700HQ CPU at clock speeds up to 3.4 GHz. We will denote this equipment as the i7 computer. The second computational hardware used was a dedicated node of a large cluster, running two Intel Xeon E5-2680v2 processors at 2.8 GHz with 128 GB of available RAM. This machine will be referred to as the Xeon computer. These two different hardware setups allowed us to evaluate algorithmic performance in both high-performance computing specialized equipment as well as in readily available consumer-grade computers.

### 5.1. Accuracy comparison experiments

In order to precisely evaluate the accuracy of the proposed method, we have simulated data acquisition from a known image in the following form. There is an analytical formula for the Radon transform of the indicator function of an ellipse. Because the Shepp–Logan phantom (see the left part of Fig. 1) is a sum of such functions, in this case $\mathcal{R}[f](\theta, t)$ can be exactly computed for any pair $(\theta, t)$. For this set of experiments we have used the i7 computer.

**5.1.1. Data simulation.** We have assumed an ideal constant flat-field of value $\varphi_i^\dagger = 23000$ photons pixel$^{-1}$ and again an ideal dark field of $d_i^\dagger = 400$ photons pixel$^{-1}$, and computed $\rho_i^\dagger = \varphi_i^\dagger \exp\{-\mathcal{R}[f](\theta_{\kappa_i}, t_{\ell_i})\} + d_i^\dagger$. Here, we defined $\kappa_i = \lfloor (i-1)/n_t \rfloor + 1$, where $\lfloor x \rfloor$ is the largest integer smaller than $x$, and $\ell_i = [(i-1)\%n_\theta] + 1$, where $x\%y$ is the remainder of the integer division of $x$ by $y$. The pairs $(\theta_\kappa, t_\ell)$ were as in (11)–(12) with $n_\theta = 512$ and $n_t = 2048$, so that $i \in \{1, 2, \ldots, n_t n_\theta\}$ reconstructed images had $N \times N$ pixels with $N = n_t$ and represented the square $[-1, 1]^2$. After computing ideal data, Poisson noise was simulated with means $\varphi_i \simeq \text{Poisson}(\varphi_i^\dagger)$, $d_i \simeq \text{Poisson}(d_i^\dagger)$ and $\rho_i \simeq \text{Poisson}(\rho_i^\dagger)$. These parameters were meant to mimic the acquisition conditions found at the real data experiment that we will describe later in the present section.

**5.1.2. Algorithmic parameters.** FISTA requires a step size to be determined. We have found that for this dataset $T = 5 \times 10^4$ worked well. The OSTR algorithm requires determination of the number and composition of subsets; we have used versions of OSTR with $n_s \in \{1, 2, 4, 8, 16, 32\}$ subsets. Each subset $S_\nu$ was composed of the data obtained at the pairs $(\theta_\kappa, t_\ell)$ with $(\kappa, \ell) \in \{\nu, \nu + n_s, \ldots, \nu + n_s \cdot (n_\theta/n_s - 1)\} \times \{1, 2, \ldots, n_t\}$, where $n_s$ is the number of subsets and
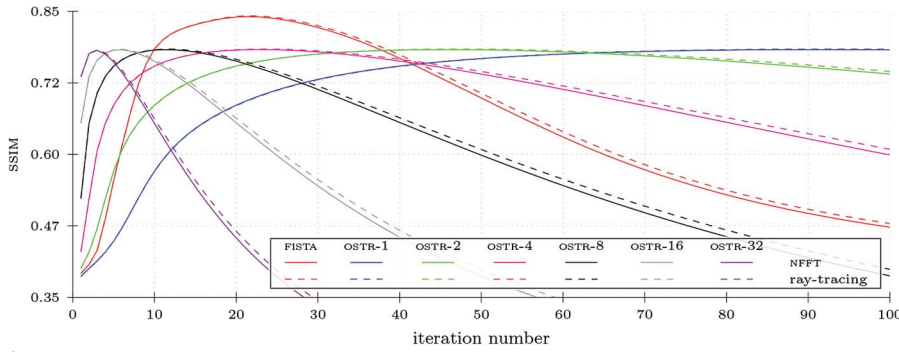
**Figure 2**
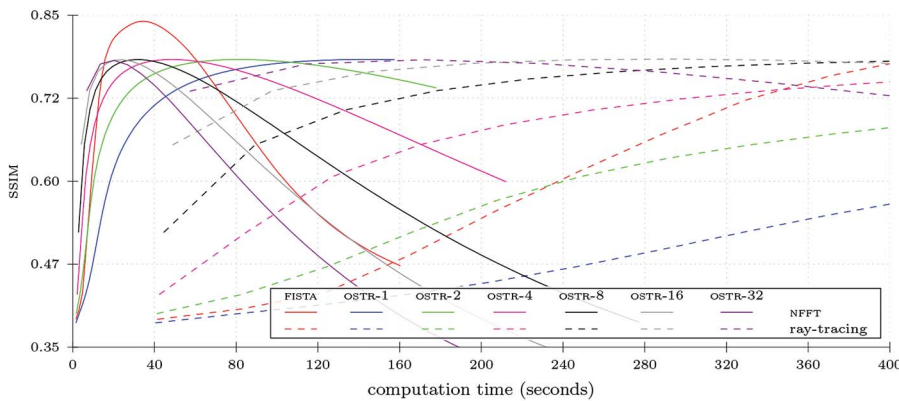Structural similarity per iteration number for a simulated noisy image.



**Figure 3**
Structural similarity per computation time for a simulated noisy image.

$\nu \in \{1, 2, \ldots, n_s\}$. We have denoted by OSTR-$n_s$ the algorithm OSTR using $n_s$ subsets. Both FISTA and OSTR require a starting image $\mathbf{x}^{(0)}$ to be determined. We have used the following formula for it, where $\mathbf{1}$ is the vector of appropriate dimension with all its components set to 1:

$$\mathbf{x}^{(0)} = \mathbf{1} \frac{\sum_{i=1}^{m} \mathbf{b}}{\sum_{i=1}^{m} (\mathbf{R1})_i},$$

where the vector $\mathbf{b}$ is the approximate Radon transform of the original image, which can be component-wise estimated as

$$b_i = \log\left(\frac{\varphi_i - d_i}{\rho_i - d_i}\right).$$

**5.1.3. Discussion.** We have measured each algorithm's reconstructed image accuracy using the well known structural similarity index measure (SSIM) (Wang *et al.*, 2004; Wang & Bovik, 2009). For this measure, a higher value is better. Figs. 2 and 3 show the values of the SSIM plotted *versus* iteration and computation time, respectively. Note that FISTA attains a higher peak SSIM value, but this is because this algorithm maintains non-negativity of the iterates, while OSTR does not. However, since images are displayed with negative values truncated to 0, the larger SSIM obtained by FISTA does not actually translate to better images for human viewing, as can be seen in Fig. 4.

A noticeable difference can be seen in Fig. 2 between the SSIM values of the iterates of methods that use NFFT-based methods compared with the equivalent algorithms that use ray-tracing operators. This, as can be seen from Fig. 5, is not due to a different iteration-wise convergence rate in terms of the objective function value but instead is due to the fact that the ray-tracing-based operator is more accurate than the NFFT-based one. However, while the effect of this difference may seem from the plots in Fig. 2 to be relevant, when we look at the images from the same iteration number of a ray-tracing-based method and compare with its NFFT-based counterpart, the difference is not visible. Fig. 6 shows this comparison for two of the methods.

If, on one hand, there is no relevant difference in images obtained in the same iteration number by NFFT-based methods when compared with ray-tracing-based methods, on the other hand, when we compare computational time to reach a given reconstructed image, the NFFT-based methods have a
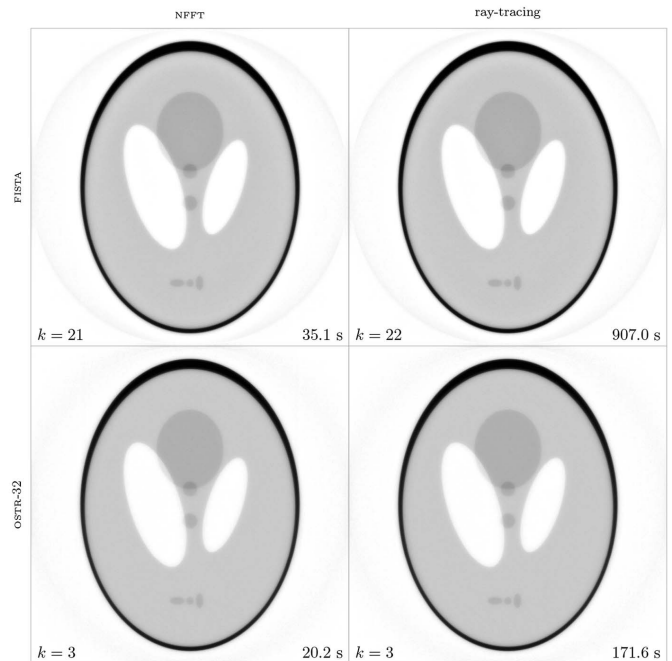


**Figure 4**
Best images obtained by FISTA and OSTR-32. Note that the NFFT-based versions are virtually indistinguishable from the ray-tracing-based ones. The number of iterations and computation time required for obtaining the reconstruction are displayed in the bottom-left and bottom-right corner, respectively, of each image.

**Table 1**
Iteration number and computation time for the best reconstruction for each algorithm.

|  | FISTA | OSTR-1 | OSTR-2 | OSTR-4 | OSTR-8 | OSTR-16 | OSTR-32 |
|---|---|---|---|---|---|---|---|
| NFFT | $k = 21$ | $k = 91$ | $k = 46$ | $k = 23$ | $k = 11$ | $k = 6$ | $k = 3$ |
| Ray-tracing | $k = 22$ | $k = 93$ | $k = 47$ | $k = 23$ | $k = 12$ | $k = 6$ | $k = 3$ |
| NFFT | 33.5 s | 143.4 s | 81.9 s | 48.8 s | 30.7 s | 24.6 s | 20.2 s |
| Ray-tracing | 907.0 s | 3820.6 s | 1956.2 s | 980.2 s | 536.6 s | 293.4 s | 171.6 s |

**Table 2**
Memory usage for each algorithm.

|  | FISTA | OSTR-1 | OSTR-2 | OSTR-4 | OSTR-8 | OSTR-16 | OSTR-32 |
|---|---|---|---|---|---|---|---|
| NFFT | 1558 MiB | 1557 MiB | 1821 MiB | 2387 MiB | 3346 MiB | 5315 MiB | 9278 MiB |
| Ray-tracing | 439 MiB | 440 MiB | 457 MiB | 449 MiB | 444 MiB | 443 MiB | 441 MiB |
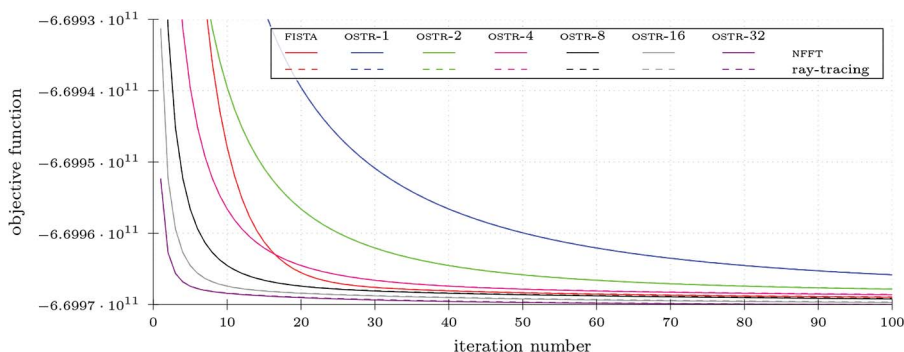


**Figure 5**
Objective function per iteration number for a simulated noisy image.
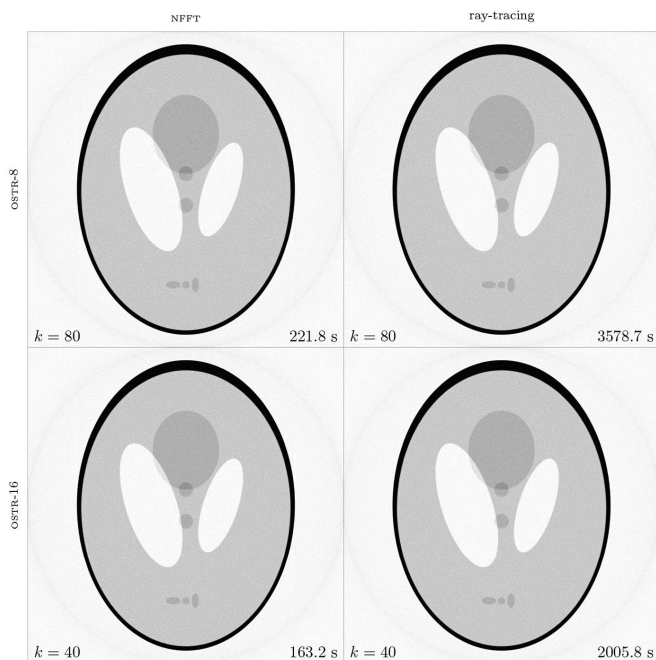


**Figure 6**
Images obtained at iterations 80 and 40 by OSTR-8 and OSTR-16, respectively.

major advantage over ray-tracing-based methods. Fig. 3 shows that most NFFT-based algorithms have already reached a satisfactory image reconstruction by the time that the ray-tracing-based methods take to compute the first iteration. Furthermore, as Fig. 5 shows, OSTR-$n_s$ performs better, iteration-wise, than FISTA if $n_s \geq 8$ and that such convergence is faster if $n_s$ is larger. Table 1 makes it clear that the iteration-wise speed-up provided by a larger number of subsets translates itself to a time-wise speed-up as well. As expected, for NFFT-based algorithms, the time-wise speed-up is not proportional to the number of subsets, as is approximately the case for ray-tracing-based methods. However, it is still advantageous to use several subsets with NFFT-based techniques.

One drawback of the NFFT-based algorithms, especially with a large number of subsets, is the high memory consumption for the computations. This is the case because the NFFT routines demand storage of the precomputed window functions used to perform the calculations, and these precomputations will result in different windows for each subset. Therefore, the memory usage grows in proportion to the number of subsets, not only to the image size. We have recorded the memory required to run each of the algorithms in Table 2, where it can be seen that the memory requirements for the algorithms using ray-tracing-based operators use essentially a constant amount of memory as the number of subsets is increased, while the NFFT-based methods use around 250 MiB of extra memory for each new subset in which the user splits the data.

## 5.2. Increasing data sizes

In this subsection we investigate how the algorithms behave as the dataset and reconstructed image sizes grow. We focus on memory usage and iteration time, since the convergence speed issues are quantitatively the same independent of problem size. That is, the OSTR-$n$ algorithm takes approximately twice the number of iterations to reach the same image as the OSTR-$2n$ and this behavior continues up to a reasonably large number of subsets (such as 32 in the case of $n_\theta = 512$).

Data generation was similar to that described in the previous subsection except that we have used $n_\theta = n_t = N$, the flat-field was set to 1, the dark-field was set to 0 and no noise was generated. We have tested $N \in \{1024, 2048, 3072, 4096, 5120, 6144\}$. The experiments described below were performed on the Xeon computer. For the results reported, FISTA stepsize is not important (we have

used $T = 2$) as image quality or convergence speed measurements are not the objective here, only the computation time per iteration. We have fixed the number of subsets and used OSTR-16 as representative of the OSTR family of algorithms. Starting image and composition of subsets were selected as in the previous section, only adapting the procedure for the new values of $n_\theta$, $n_t$ and $N$. The algorithms were run for seven iterations and the mean value of the measured computation times or the iterations was computed, disregarding the largest and smallest values. The deviation from the mean was small because the computational environment was well controlled. This could not be the case if other loads were present while reconstruction was being performed, which we have avoided.

**5.2.1. Discussion.** Fig. 7 shows that the computation time of the NFFT-method, as expected, becomes even more competitive as the dataset size grows because of the considerably lower asymptotic flops count required for each iteration of these techniques. On the other hand, as can be seen in Fig. 8, memory usage may become a bottleneck for larger image sizes or if a larger number of subsets, for more speed-up, is desired. However, for the image sizes tested, the amount of computation is still the main issue since, for example, for $6144 \times 6144$ pixel images reconstructed from $6144 \times 6144$ datasets, a maximum of around 47 GiB of memory was used, which is commonly available in current workstations. On the other hand, computation time for this size of dataset was over ten times smaller using NFFT-based operators than using ray-tracing-based operators, as can be seen in Fig. 7.

### 5.3. Reconstruction from real data

In the present section we discuss results obtained by applying the presented methodology to real data obtained at the IMX tomography beamline at the LNLS. Data dimensions were the same as those for the noisy phantom experimented with in §5.1, that is, $n_\theta = 512$, $n_t = N = 2048$ and reconstruction was again performed on the i7 computer.

These data were acquired with a mean flat-field count of 23359 photons pixel$^{-1}$ and a mean dark-field count of 401 photons pixel$^{-1}$. The imaged subject is an apple seed, within a field of view measuring 7.58 mm × 7.58 mm. Dark-field counts were measured before and after acquisition and linearly interpolated in order to obtain approximate counts for the number of events that would have occurred during acquisition. This is necessary because the LNLS storage ring is a second-generation ring and beam intensity decays with time.
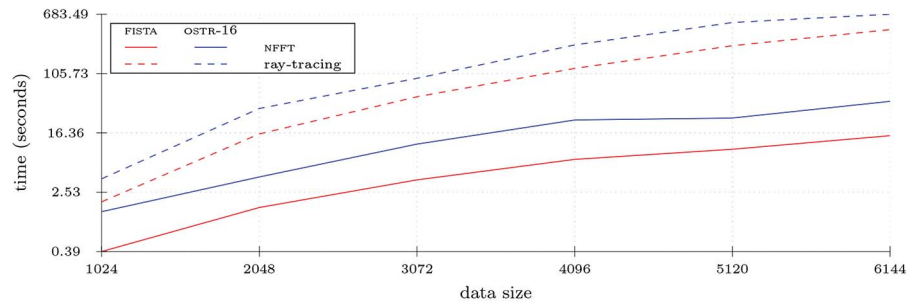


**Figure 7**
Mean iteration time per simulated data size. The vertical axis is on a logarithmic scale.
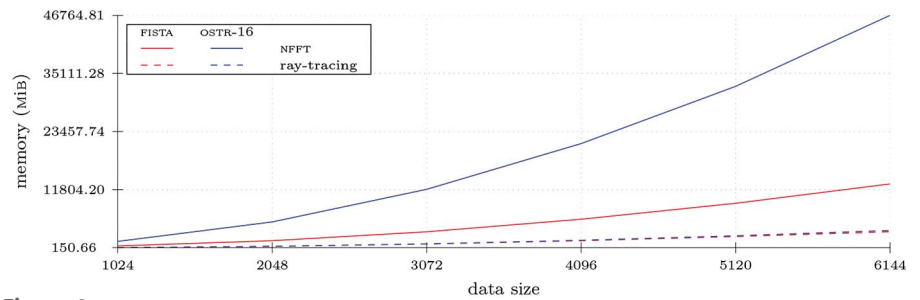


**Figure 8**
Memory usage per simulated data size.

The full polychromatic spectrum was used in the data collection, with no correction for eventual beam-hardening effects.

The resulting images are shown on Fig. 9, where we can confirm that the conclusions made from simulated testing still apply to real-world data. In particular, OSTR-32 converges iteration-wise approximately twice as fast as OSTR-16. Time-
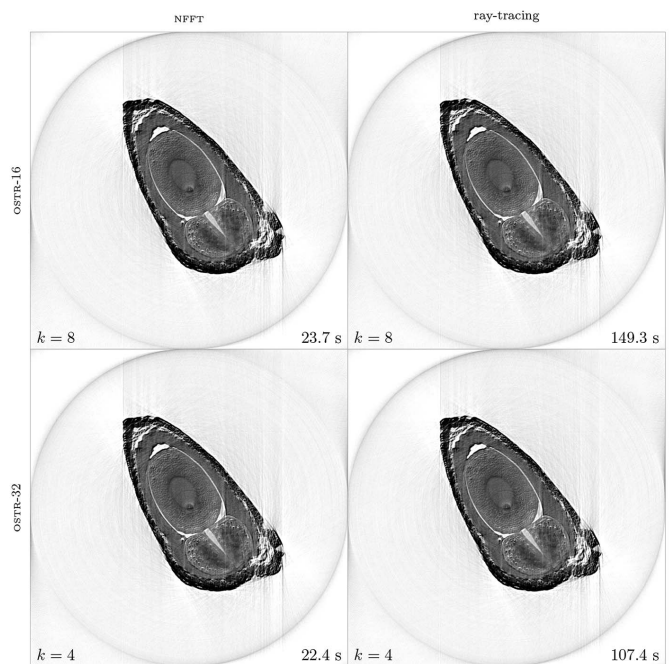


**Figure 9**
Images obtained by OSTR-16 and OSTR-32 from synchrotron-illuminated tomographic acquisition.

wise, however, we are starting to see some saturation for the NFFT-based algorithms and, indeed, although not shown, OSTR-48 spends as much time as OSTR-32 (but uses more memory) to obtain similar reconstructions if both are used with NFFT-based operators. This is because the NFFT-based methods have a large per-subset overhead. On the other hand, ray-tracing-based methods could benefit from further speed-up, but it would still not be competitive against the corresponding NFFT-based algorithms. Images are visually very similar and it is possible to say that there is no degradation caused because of the replacement of the ray-tracing by the NFFT.

## 6. Conclusions

We have proposed a combination of two different acceleration techniques for iterative methods for maximum-likelihood transmission tomography image reconstruction. We have applied NFFT-based Radon operators with incremental (ordered subsets) iterative techniques. The results are promising and the methods were successfully applied to both synthetic and real data, showing a good speed-up combined with uncompromising accuracy.

Current hardware for numerical computation focus on very large parallelism and we consider this trend to be a topic of future research. All of the codes used in the present paper were able to take advantage of the multi-core architecture of the hardware where they had been tried, but the general purpose GPUs currently on the market require specially crafted code and it remains to be seen if the advantage that NFFT-based algorithms present when running on CPUs remains valid when translated to GPUs.

Finally, there are other options of fast operators available, many of which have been applied to iterative algorithms [see, for example, Arcadu *et al.* (2016) for a comparison among many of the possibilities]. However, no other work in the literature, to our knowledge, has applied such methods to incremental iterative algorithms. In the present paper we have used the NFFT as implemented in the NFFT 3 library solely because of its availability and good documentation. Therefore, another direction for future work could focus on comparing the performance of other existing methods for fast Radon operators when applied to incremental methods.

## References

Andersson, F. (2005). *SIAM J. Appl. Math.* **65**, 818–837.
Arcadu, F., Nilchian, M., Studer, A., Stampanoni, M. & Marone, F. (2016). *IEEE Trans. Image Process.* **25**, 1207–1218.
Averbuch, A., Coifman, R. R., Donoho, D. L., Israeli, M., Shkolnisky, Y. & Sedelnikov, I. (2008). *SIAM J. Sci. Comput.* **30**, 785–803.
Beck, A. & Teboulle, M. (2009*a*). *IEEE Trans. Image Process.* **18**, 2419–2434.
Beck, A. & Teboulle, M. (2009*b*). *SIAM J. Imaging Sci.* **2**, 183–202.
Bracewell, R. N. (1965). *The Fourier Transform and its Applications.* New York: McGraw-Hill.
Browne, J. & De Pierro, A. R. (1996). *IEEE Trans. Med. Imaging*, **15**, 687–699.
Cooley, J. W. & Tukey, J. W. (1965). *Math. Comput.* **19**, 297–301.
De Pierro, A. R. & Yamagishi, M. E. B. (2001). *IEEE Trans. Med. Imaging*, **20**, 280–288.
Erdoğan, H. & Fessler, J. A. (1998). *Proc. IEEE Intl Conf. Image Process.* **2**, 680–684.
Erdoğan, H. & Fessler, J. A. (1999). *Phys. Med. Biol.* **44**, 2835–2851.
Fessler, J. A. & Sutton, B. P. (2003). *IEEE Trans. Signal Process.* **51**, 560–574.
Fourmont, K. (2003). *J. Fourier Anal. Appl.* **9**, 431–450.
George, A. & Bresler, Y. (2007). *SIAM J. Appl. Math.* **68**, 574–597.
Han, G., Liang, Z. & You, J. (1999). *IEEE Nucl. Sci. Symp.* **3**, 1515–1518.
Hawkes, P. W. (2006). *Advances in Imaging and Electron Physics*, Vol. 139. Elsevier Academic Press.
Helou, E. S. (2009). PhD thesis, Universidade Estadual de Campinas, Campinas, Brazil.
Helou, E. S., Censor, Y., Chen, T., Chern, I., De Pierro, A. R., Jiang, M. & Lu, H. H. (2014). *Inverse Probl.* **30**, 055003.
Helou, E. S. & De Pierro, A. R. (2005). *Inverse Probl.* **21**, 1905–1914.
Herman, G. T. (1980). *Image Reconstruction from Projections: the Fundamentals of Computerized Tomography.* New York: Academic Press.
Hudson, H. M. & Larkin, R. S. (1994). *IEEE Trans. Med. Imaging*, **13**, 601–609.
Kak, A. C. & Slaney, M. (1988). *Principles of Computerized Tomographic Imaging.* IEEE Press.
Keiner, J., Kunis, S. & Potts, D. (2009). *ACM Trans. Math. Softw.* **36**, 1–30.
Natterer, F. (1986). *The Mathematics of Computerized Tomography.* New York: Wiley.
Potts, D. & Steidl, G. (2000). *Proc. SPIE*, **4119**, 13.
Radon, J. (1986). *IEEE Trans. Med. Imaging*, **5**, 170–176.
Shepp, L. A. & Vardi, Y. (1982). *IEEE Trans. Med. Imaging*, **1**, 113–122.
Siddon, R. L. (1985). *Med. Phys.* **12**, 252–255.
Tanaka, E. & Kudo, H. (2003). *Phys. Med. Biol.* **48**, 1405–1422.
Thorson, J. (1978). *Stanford Explor. Proj.* **14**, 81–86.
Vardi, Y., Shepp, L. A. & Kaufman, L. (1985). *J. Am. Stat. Assoc.* **80**, 8–20.
Wang, Z., Bovik, A. C., Sheikh, H. R. & Simoncelli, E. P. (2004). *IEEE Trans. Image Process.* **13**, 600–612.
Wang, Z. & Bovik, A. C. (2009). *IEEE Signal Process. Mag.* **26**, 98–117.