FAST RECONSTRUCTION OF CT IMAGES FROM PARSIMONIOUS ANGULAR MEASUREMENTS VIA COMPRESSED SENSING∗

N. S. AYBAT† AND A. CHAKRABORTY‡

Abstract. Computed Tomography (CT) is one of the most popular diagnostic tools available to medical professionals. However, its diagnostic power comes at the cost of significant radiation exposure to the patient. Modifying the tube current has been the primary hardware-based dose reduction mechanism; however, the image quality is also dependent on the tube current strength and the standard deviation of image noise is inversely proportional to the square root of the tube current level. On the other hand, for a fixed X-ray tube current level, since the amount of radiation exposure increases linearly with the number of angular measurements taken, reducing the number of measurements while ensuring the same reconstruction quality is of immense value in this field. The essence of this paper is to design a measurement and an iterative reconstruction scheme that makes Compressed Sensing (CS) measurement paradigm practically applicable to CT imaging. Given a target object, continuous in the spatial domain, our method can faithfully reconstruct the image using 4 to 8 times less measurements than that mandated by the Nyquist criterion. Each iteration of the method has the same worst case computational complexity as fast fourier transform. Moreover, our method only requires acquiring measurements at specifically chosen projection angles; hence, it can be readily used with the existing parallel beam tomography machines.

1. Introduction. Computed Tomography (CT) is one of the most powerful diagnostic tools at the disposal of clinicians today. Since it’s inception in the 1970’s, CT has seen an explosive growth in its use not just for localization and quantification of a variety of ailments for symptomatic patients, but also as a preventive mechanism to screen asymptomatic patients; both adults and increasingly even children. However, despite it’s benefits, CT scan exposes the patient to ionizing radiation, of which amount is a function of the image resolution and quality desired. In fact, CT has contributed disproportionately to radiation exposure. A study done in Britain for instance noted that while CT constitutes 4% of the diagnostic procedures, it accounts for 40% of the radiation exposure [34]. A more recent study done in the US suggests that 1.5-2% of all cancers in the US may be attributed to CT radiation [3].

It is not surprising then that reducing CT radiation exposure without compromising the image quality is a major area of research. This problem is being addressed from from two different perspectives: hardware innovations and algorithmic improvements in image reconstruction. Since the radiation dose is directly proportional to the tube current, modifying the tube current has been the primary hardware-based dose reduction mechanism. However, the image quality is also dependent on the tube current strength; and the standard deviation of image noise is inversely proportional to the square root of the tube current level. Hence, reducing the radiation dose by a factor of four by decreasing the tube current level, comes at the cost of an increase in the image noise by a factor of two [4, 9].

A further significant reduction can be achieved by reducing the amount of raw CT data required to reconstruct the images through the use of advanced iterative image reconstruction algorithms that exploit the characteristics of the imaged volume, e.g., smoothness of the boundaries of internal organs, as opposed to the traditional filtered back projection algorithm (FBP) [18]. These methods are also more robust to missing and/or noisy data since they use prior models of the imaged volume.

We approach the dose reduction problem from the second perspective. Our goal is to reconstruct a good quality image from fewer angular projections than what is required by the Nyquist sampling criterion for perfect reconstruction. Normally this should result in streaking artifacts when the traditional FBP algorithm [18] for CT image reconstruction is used. However, using the method proposed in this paper, an almost artifact free reconstruction was achieved for well-known test images using a compressed sensing (CS) formulation of the CT reconstruction problem through the use of partial pseudo-polar Fourier sensing matrix. Our studies show that on the test problems radiation dose can be reduced by 75% to 87.5% without perceptible difference in the image quality. Moreover, one can combine our proposed method with hardware dose reduction

∗This work has been done when the first author was a PhD student at Columbia University
†Industrial Engineering Department, The Pennsylvania State University. Email: nsa10@psu.edu
‡Siemens Corporate Research. Email: amit.chakraborty@siemens.com

1
strategies such as tube current reduction to achieve further reduction in the radiation exposure.

In Section 2 we briefly discuss other CS based image reconstruction methods for CT and summarize our contribution. There are also other algorithmic approaches that aim to reduce raw CT data collection. The first seminal work in this direction is [30] using half-scan fan-beam projections (thus cutting data acquisition by half) which was improved further by [28, 29]. This was further extended for exact reconstruction of an interior region such as a specific organ, especially for applications such as cardiac CT with prior analytic information of the region of interest [11, 21, 37, 36, 38]. However, obtaining a priori analytic information about a region is quite complicated except for a few sub-structures that are either filled with air or blood, thereby limiting the use of these methods for reducing raw projection data in CT image reconstruction.

The rest of the paper is organized as follows. We first briefly introduce the reconstruction problem in Section 1.1, and in Section 1.2 we give a short introduction of the compressed sensing framework used. In Section 2, we briefly discuss previous CS based methodologies for CT imaging that are relevant to our work and summarize our contribution. In Section 3, we model CT projection data obtained via the compressed sensing framework using partial pseudo-polar Fourier transform. In Section 4, we briefly describe the compressed sensing optimization algorithm used in the paper without any details. Next, in Section 5 we present some experimental results to test the performance of the proposed algorithm on the reconstruction of Shepp-Logan head phantom; and Section 6 has some concluding comments.

1.1. CT Image Reconstruction. The 2D target object is modeled as a two-dimensional distribution of X-ray attenuation coefficients and is represented by a two-dimensional function $f(x, y)$ in the standard Cartesian coordinate system. In the rest of this paper we will use object and $f$ interchangeably. We assume that the object is bounded in the spatial domain, i.e. $f(x, y) = 0$ when $|x| > B$ or $|y| > B$ for some $B > 0$. Let $\theta \in [0, \pi]$ be a projection angle in radians. Then given a point $(\theta, t)$ in polar coordinates of the spatial-domain, the trajectory of an X-ray beam passing through the point $(\theta, t)$ is represented by $l_{\theta, t} := \{(x, y) : x \cos \theta + y \sin \theta = t\}$, as shown in Figure 1.1.

Let $\Omega_1$ be a finite set of projection angles and $P_\theta(t) := \int_{l_{\theta, t}} f(x, y) \, ds$ denote the total attenuation of the X-ray beam along the line $l_{\theta, t}$. For a fixed $\theta \in \Omega_1$, $P_\theta := \{P_\theta(t) : -\infty < t < +\infty\}$ is the set of parallel projection data of the object corresponding to the projection angle $\theta$. In conventional monochromatic X-ray tomography [20], the interaction of X-rays with the substance of the object causes the projection data $P_\theta(t)$ and it can be approximated by the logarithm of the ratio of entering to the leaving number of photons.

Throughout this paper parallel ray projections will be used. Parallel ray projection data can be easily collected. After fixing a constant projection angle $\theta \in \Omega_1$, the X-ray source and X-ray detector are moved together along parallel lines on the opposite sides of the target object to acquire the data $P_\theta$. Then repeating this step for a set of projection angles in $\Omega_1$, a set of parallel ray projection data, i.e. $\{P_\theta : \theta \in \Omega_1\}$ can be obtained.

Since for a fixed level of X-ray tube current, the total radiation exposure of the target object is proportional to the number of view angles in $\Omega_1$, we would like to minimize this number. Compressed Sensing Theory (CS) [6, 12] shows that, if the signal to be reconstructed has a sparse representation in some basis, then it can be perfectly reconstructed by sampling at a significantly smaller rate than the Nyquist rate. Thus, this theory provides us with the possibility of reducing the number of view angles provided that the attenuation coefficient distribution $f$ of the target object has a sparse representation in some basis. Although real medical images are rarely sparse, transformations of the image via a wavelet transform and/or the total variation operator can result in representations of $f$ having sparse coefficients.

1.2. Compressed Sensing. Although most real-life signals do not have sparse representations, many of them are compressible, i.e., the magnitude of the (sorted) transform coefficients decays rapidly. Therefore, compressible signals can be well approximated by the relatively few large transform coefficients. Classical compression techniques such as JPEG2000 for image compression exploits this property after the signal is completely acquired. Compressed sensing (CS) [6, 12] is a new area of research which has gained enormous popularity due to its ability to perfectly reconstruct a signal in $\mathbb{R}^n$ using only $O(\log(n))$ samples provided
that the signal has a sparse representation in some transform domain. Hence, the CS methodology allows one to compute approximations to compressible signals without first acquiring the signal. CS is being increasingly adopted in a variety of applications (see [7] for a review of the theory and applications).

Let the vector $\bar{f} \in \mathbb{R}^n$ represent a signal. We assume that there exists a transform $\Phi$, e.g., Fourier or wavelet transform, such that $\bar{\gamma} := \Phi \bar{f} \in \mathbb{R}^n$ is sparse, i.e., the number of non-zeros in $\bar{\gamma}$ is much smaller than $n$, denoted by $\|\bar{\gamma}\|_0 \ll n$.

The CS process consists of two steps: sensing and decoding. The sensing step encodes $\bar{f}$ into a vector $b \in \mathbb{R}^m$, $m \ll n$; via a series of linear measurements $b = R\bar{f}$ for some linear transform $R \in \mathbb{R}^{m \times n}$ (the exact physical nature of these measurements are context dependent). Since $\bar{f} = \Phi^{-1} \bar{\gamma}$, it follows that $b = A\bar{\gamma}$ for $A := R\Phi^{-1}$. Note that $R$ is certainly not invertible since $m < n$ and the entire $\bar{f}$ is never acquired. The decoding step recovers $\bar{\gamma}$ (and thus, $\bar{f}$) from $b$. Since $\bar{\gamma}$ is sparse, it can be recovered as the sparsest solution to the underdetermined system of equations $A\gamma = b$ when $A$ satisfies certain conditions, i.e., $\bar{\gamma}$ is the solution to the $\ell_0$-problem

$$\min\{\|\gamma\|_0 : A\gamma = b\}.$$ 

However, this problem is NP-hard [27] for general data, which makes this model impractical for nearly all real applications. The $\ell_1$-problem

$$\min\{\|\gamma\|_1 : A\gamma = b\}, \quad (1.1)$$

is a tractable convex approximation of the $\ell_0$-problem. The optimal solution $\gamma^*$ of (1.1) is also optimal for the $\ell_0$-problem under certain conditions on $A$ (see [13, 14, 15] for details). Ideally, we would like to take the least possible number of measurements, that is, $m$ being equal to $s := \|\bar{\gamma}\|_0$. However, we must pay the price for not knowing the locations of the non-zeros in $\bar{\gamma}$ (there are $n$ choose $s$ possibilities!) while still asking for the perfect reconstruction of the most sparse $\bar{\gamma}$, which satisfies $A\bar{\gamma} = b$. It was shown in [6, 32] that, when $R$ is Gaussian or partial Fourier, i.e., the entries of $R$ are randomly selected using a Gaussian pdf or $m$ rows of $R$ are randomly selected from the rows of $n \times n$ Fourier matrix, then (1.1) can recover $\bar{\gamma}$ (with a very high probability) from $m = O(s \log(n/s))$ or $O(s \log(n)^4)$ linear measurements, respectively. In the case of CT, it has been shown that using compressed sensing the perfect reconstruction of a subregion of interest is still possible from an incomplete dataset of projection values if we only know the Radon transform of the subregion of interest [38]. Finally, once $\bar{\gamma}$ is recovered, $\bar{f}$ becomes available through the inverse transform, $\bar{f} = \Phi^{-1} \bar{\gamma}$.

Fig. 1.1: An object and its parallel projections [18, 31].
The exact recovery of $\bar{f}$ by (1.1) requires that $\bar{\gamma}$ be sparse and $b = A\bar{\gamma}$, exactly. In practice, especially in imaging, it is often the case that $\bar{f}$ is compressible, i.e., $|\bar{\gamma}_i|$ decays rapidly, but none of the components are exactly zero. In some applications, it is also the case that the measurement vector $b$ is contaminated with noise $e$, i.e., $b = A\bar{\gamma} + e$. In both cases, a more appropriate model is given by

$$\min_{\gamma} \{ \|\gamma\|_1 : \|A\gamma - b\|_2^2 \leq \sigma^2 \},$$

(1.2)

where $\sigma > 0$ is a tunable parameter. It was shown in [5] that the solution $\gamma^*$ of (1.2) satisfies

$$\|\gamma^* - \bar{\gamma}\|_2 \leq O\left(\sigma + \frac{1}{\sqrt{s}}\|\bar{\gamma} - \bar{\gamma}(s)\|_1\right),$$

(1.3)

where $\bar{\gamma}(s)$ is the $s$-approximation to $\bar{\gamma}$ by zeroing out the $n-s$ smallest magnitude components. Clearly, if $\bar{\gamma}$ is exactly $s$-sparse and $\sigma$ is set to 0, then (1.3) reduces to the exact recovery of $\bar{\gamma}$, i.e., $\gamma^* = \bar{\gamma}$.

Solving problem (1.2) is equivalent to solving the simpler Lagrange relaxation problem

$$\min_{\gamma} \mu \|\gamma\|_1 + \frac{1}{2}\|A\gamma - b\|_2^2,$$

(1.4)

for an appropriately chosen $\mu > 0$. Given the data $A, b$, and $\sigma$ in (1.2), there exist practical ways (see [16]) to determine an appropriate $\mu$ in (1.4).

2. Previous CS based methods for CT imaging and our contribution. Chen et al. show that the number of measurements and the associated radiation in 4D CT can be significantly reduced by using a sparsifying gradient transformation and prior image, exploiting the spatial-temporal correlations in dynamic CT images [10] (See also [22]). In their study, each image slice is sequentially scanned in order to capture the dynamics of the imaged object. Although each scan is down sampled by 32 times, they make sure that the union of dynamic data satisfies Nyquist sampling criterion for each image slice. They first reconstruct the prior image of each slice utilizing all the dynamic data using FBP algorithm. Subsequently, in their model, they exploit the fact that the difference between phased and the prior images of each slice is sparse. Further applications of this approach has been shown in [23]. However, acquiring prior information is always a challenge. This approach is susceptible to motion artifacts and inconsistent projections due to utilization of the dynamic data to reconstruct the prior.

Candès and Romberg [8] show that a digital image $f^d$ sampled on the Cartesian grid can be reconstructed using partial Fourier coefficients. Assuming that the wavelet representation $\Phi f^d$ is sparse, they show that $f^d$ can be recovered by solving

$$\min_{\gamma} \text{TV}(\Phi^{-1}\gamma) + \beta\|\gamma\|_1,$$

(2.1)

$$\text{s.t. } A\gamma = b,$$

for an appropriate $\beta > 0$, where $A := P\Phi^{-1}$, $P$ is the partial Fourier matrix, $\Phi^{-1}$ is the inverse discrete wavelet transform matrix, and $\text{TV}(\cdot)$ is the total variation norm (for definition, see Section 4). $P$ is called the partial Fourier matrix, because $P$ is formed by only those rows corresponding to a subset of frequencies. In particular, Candès and Romberg [8] used a star shaped subset for reconstruction of the Shepp-Logan phantom. In their experimental setup, they take Fourier measurements along 22 equally equiangular lines through the origin in the frequency domain to mimic a parallel beam CT machine taking 22 equiangular projections of the object. It is shown in [8] that solving (2.1) exactly reconstructs the digital Shepp-Logan head phantom. However, Candès and Romberg warn the reader to be cautious about their result: “In most scenarios, these are samples of a continuous-space image that do not lie on a Cartesian grid, making the situation somewhat more complicated. We are considering a simplified version here, where we are given samples of digital images that lie on the usual discrete Fourier grid” [8].

Candès and Romberg [8] assumed that the object of interest is already in the discrete-space and that the star-shaped Fourier coefficient measurements also lie on discrete Cartesian coordinates in the frequency domain.
However, the Fourier Slice Theorem [19] implies that those Fourier measurements do not lie on discrete cartesian grid but on a polar grid of the frequency domain.

As far as we are aware of the current literature on the application of CS to CT imaging, either the computational complexity of the reconstruction algorithm is not properly analyzed, e.g., [38], or one of the following two simplifying assumptions are made in order to model the image reconstruction problem and develop an efficient algorithm to solve it:

1. In the first approach, e.g. [10, 17, 22, 24, 35], the Radon transform is discretized to obtain a measurement matrix \( A \) that approximates the Radon transform. This approach implicitly assumes that the attenuation coefficients of the target object is a piecewise constant function in the spatial domain. In practice, when matrix \( A \), which only approximates the Radon transform, is used many times in an iterative reconstruction algorithm leads to a significant accumulation of approximation errors.

2. In the second approach, e.g., [8], one assumes that the Fourier coefficients of the target object are available at certain points on the Cartesian grid. However, this assumption is not entirely realistic since one cannot compute the Fourier coefficients on the Cartesian grid using the projection data available from the parallel beam tomography machine. The Fourier slice theorem implies that the Fourier coefficients are only available on the polar grid. Therefore, the measurement setup in this approach is not applicable in practice, i.e., choosing \( A \) as a partial Fourier matrix.

In addition, since the computation time increases rapidly as a function of the image size, computational complexity aspect of the basic operations used in a reconstruction algorithm becomes vital, especially if one wants to reconstruct images with high resolution. We briefly discuss this issue in Section 5.

In our paper, we assume a much more realistic scenario:

1. The object of interest is continuous in the spatial domain, i.e., we do not assume a discrete attenuation function.

2. We directly work with the Radon transform of the object provided by any parallel beam CT machine. In particular, we do not assume that the Fourier coefficients are available on the Cartesian grid.

In summary, the essence of our paper is to design a scheme that makes CS measurement paradigm practically applicable to Computed Tomography. We believe that the reason why CS has not been as widely used in CT imaging as compared to in Magnetic Resonance Imaging (MRI) is because of the difficulty of designing a signal acquisition scheme such that the measurements \( b \) can be written as \( b \approx A\gamma \) for a matrix \( A \) which allows fast multiplication without introducing too much error.

The contribution of this paper to the literature on the applications of CS to CT imaging can be summarized as follows: Given a target object, continuous in the spatial domain, our measurement and reconstruction scheme can faithfully reconstruct the image using 4 to 8 times less measurements than that mandated by the Nyquist criterion. Each iteration of the scheme has the same worst case computational complexity as fast fourier transform (FFT). Moreover, our scheme only requires acquiring measurements at specifically chosen projection angles; hence, it can be readily used with the existing parallel beam tomography machines.

3. CT Image Reconstruction using CS Framework. The essence of our proposed method is as follows. First, we need to decide on how to select a measurement matrix \( R \) consistent with the attenuation process that generates \( P_\theta \) when an X-ray traverses through the target object (see Section 1.1 for the definition of \( P_\theta \)). Moreover, \( R \) should be such that left and right vector multiplications with \( R \) are efficiently computable. The second step is to compute the measurement vector \( b \) corresponding to \( R \) using the acquired projection data \( P_\theta \) for all \( \theta \in \Omega_1 \). Once \( R \) and \( b \) are known, the resulting CS problem

\[
\min_\gamma \alpha \ TV(\Phi^{-1}\gamma) + \beta \|\gamma\|_1 + \frac{1}{2} \|R\Phi^{-1}\gamma - b\|_2^2.
\]

(3.1)

is solved using an appropriate optimization algorithm, where \( \alpha > 0 \) and \( \beta > 0 \) are model parameters.

Let \( F : \mathbb{R}^2 \rightarrow \mathbb{C} \) denote the Fourier transform of \( f \), \( F(u, v) := \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) e^{-2\pi i (ux + vy)} \, dx \, dy \). Let \( s : \)
\[ \mathbb{R}^2 \rightarrow \mathbb{R} \] denote the two dimensional sampling function, \( s(x, y) := T^2 \sum_{n_1 = -\infty}^{+\infty} \sum_{n_2 = -\infty}^{+\infty} \delta(x-n_1 T)\delta(y-n_2 T) \), where \( \delta(\cdot) \) is the Dirac delta function and \( T \) is the sampling period for both \( x \) and \( y \) axes. Then we define the sampled object \( f_s(x, y) \) to be \( f_s(x, y) := f(x, y) s(x, y) \). Let \( F_s \) denote the Fourier transform of the sampled object,

\[
F_s(u, v) := \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_s(x, y) \, e^{-2\pi i (ux+vy)} \, dx \, dy,
\]

\[
= T^2 \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(n_1 T, n_2 T) \, e^{-2\pi i (un_1 + vn_2) T} \, dx \, dy,
\]

\[
= T^2 \sum_{n_1 = [-\frac{B}{T}]}^{+\frac{B}{T}} \sum_{n_2 = [-\frac{B}{T}]}^{+\frac{B}{T}} f(n_1 T, n_2 T) \, e^{-2\pi i (un_1 + vn_2) T}.
\]

It can be easily shown that \( F_s(u, v) = \sum_{k_1 = -\infty}^{+\infty} \sum_{k_2 = -\infty}^{+\infty} F(u - \frac{k_1}{T}, v - \frac{k_2}{T}) \). Since the support of the object, \( f \), is bounded in the spatial domain, the support of \( F \) is not bounded. However, we assume that there exists a cutoff frequency \( W \) such that \( \|F(u, v)\| \approx 0 \), when \( |u| > W \) or \( |v| > W \) and

\[
\hat{f}(x, y) := \int_{-W}^{W} \int_{-W}^{W} F(u, v) \, e^{2\pi i (ux+vy)} \, du \, dv \approx \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} F(u, v) \, e^{2\pi i (ux+vy)} \, du \, dv = f(x, y),
\]

in the sense that \( \hat{f}(x, y) \) is a sufficiently good approximation of \( f(x, y) \) for all practical purposes. Assuming that such a bound \( W \) exists, the aliasing effects will be small and \( F_s(u, v) \approx F(u, v) \) for \( |u| \leq W \) and \( |v| \leq W \) whenever the sampling period \( T < \frac{1}{2W} \).

Let \( S_0(w) = \int_{-\infty}^{+\infty} P_0(t) e^{-2\pi i wt} dt \) be the Fourier transform of \( P_0(t) = \int_{l_\theta,t} f(x, y) \, ds \), where \( l_{\theta,t} = \{(x, y) : x \cos \theta + y \sin \theta = t\} \) is the trajectory of an X-ray beam when the projection angle is equal to \( \theta \). The two-dimensional polar Fourier transform \( \mathcal{F}(w, \theta) \) of \( f \) is defined as \( \mathcal{F}(w, \theta) = F(w \cos \theta, w \sin \theta) \). Then, the Fourier Slice Theorem [19] implies that \( S_0(w) = \mathcal{F}(w, \theta) \). However, in practice, instead of the entire polar profile \( P_0 = \{P_0(t) : -\infty < t < +\infty\} \), parallel beam tomography only measures the total attenuation values of a finite set of X-ray beams,

\[
\mathcal{P}_\theta := \left\{ P_0(n\tau) : n \in \mathbb{Z}, \left\lfloor -\frac{\sqrt{2} B}{\tau} \right\rfloor \leq n \leq \left\lceil \frac{\sqrt{2} B}{\tau} \right\rceil \right\},
\]

where \( \tau \) is the distance between the adjacent X-ray detectors, and \( B > 0 \) is a bound such that \( f(x, y) = 0 \) when \( |x| > B \) or \( |y| > B \). The lower and upper bounds in the definition of \( \mathcal{P}_\theta \) follows from the fact that the object, \( f \), lies in a ball centered at the origin with radius \( \sqrt{2} B \).

Let

\[
\hat{S}_\theta(w) := \tau \sum_{n=\left\lfloor -\frac{\sqrt{2} B}{\tau} \right\rfloor}^{\left\lceil \frac{\sqrt{2} B}{\tau} \right\rceil} P_0(n\tau) \, e^{-2\pi i wn\tau}.
\]  

(3.2)

Since \( \hat{S}_\theta(w) = \sum_{k=-\infty}^{+\infty} S_\theta(w - \frac{k}{\tau}) \), it follows that \( \hat{S}_\theta(w) \rightarrow S_\theta(w) \) for all \( w \in \mathbb{R} \) such that \( |w| \leq W \) as \( \tau \searrow 0 \). Moreover, for all for \( |u| \leq W \) and \( |v| \leq W \), we also have \( F_s(u, v) \rightarrow F(u, v) \) as \( T \searrow 0 \). Hence, for \( \tau > 0 \) and \( T < \frac{1}{2W} \) sufficiently small, it follows that

\[
\hat{S}_\theta(w) \approx S_\theta(w) = \mathcal{F}(w, \theta) = F(w \cos \theta, w \sin \theta) \approx F_s(w \cos \theta, w \sin \theta),
\]

whenever \( |u| \leq W \) and \( |v| \leq W \).

Sampling period of the machine \( \tau > 0 \) is assumed to be fixed and chosen such that \( \hat{S}_\theta \) sufficiently approximates \( S_\theta \) for all practical purposes. Using a database of past images, for different type of CT images e.g. head,
We now describe how to select $\Omega$. For a finite set $\Omega = \{\theta_1, \theta_2, \ldots, \theta_{N_1}\}$ of projection angles, we use the optimization algorithm [25] described in Section 4. For each $\theta \in \Omega$, $\hat{S}_\theta(w)$ can be computed offline for all $-W \leq w \leq W$ and for all $\theta \in \Omega_1$. For each $\theta \in \Omega_1$, let $\hat{S}_\theta(w)$ be computed for all $w \in \Omega_2 = \{w_1, w_2, \ldots, w_{N_2}\}$ such that $|w_i| \leq W$ for all $1 \leq i \leq N_2$. Then each point in $\Omega = \Omega_1 \times \Omega_2$ is a polar coordinate representation of points in two-dimensional frequency domain of $f$. The set of observations $\{\hat{S}_\theta(w) : (\theta, w) \in \Omega\}$ can be used to reconstruct digital image $f^d$ of the object $f$. Since for small $\tau > 0$, $F_s(w \cos \theta, w \sin \theta) \approx \hat{S}_\theta(w)$ when $|u| \leq W$ and $|v| \leq W$, then for some $e_{\theta,w} \in \mathbb{R}$ close to zero, $F_s(w \cos \theta, w \sin \theta) = \hat{S}_\theta(w) + e_{\theta,w}$ for all $(\theta, w) \in \Omega$, i.e.

$$T^2 \sum_{n_1=\lfloor \frac{B}{T} \rfloor}^{\lfloor \frac{B}{T} \rfloor} \sum_{n_2=\lfloor \frac{B}{T} \rfloor}^{\lfloor \frac{B}{T} \rfloor} f^{d}_{n_1,n_2} e^{-2\pi i (n_1 \cos \theta + n_2 \sin \theta) w T} = \hat{S}_\theta(w) + e_{\theta,w} \quad \forall (\theta, w) \in \Omega$$

This set of equations can be written in matrix form: $R f^d = b + e$, where $R \in \mathbb{C}^{n_2^2} \times n_2^2$, $f^d \in \mathbb{R}^{n_2^2}$, $b \in \mathbb{C}^{n_2^2}$ and elements of $b$ are $\hat{S}_\theta(w)$ for $(\theta, w) \in \Omega$, and $e \in \mathbb{C}^{n_2^2}$ is a vector of approximation errors.

We solve (3.1) using a first-order algorithm. The bottleneck operation in the first-order algorithms are matrix-vector multiplications in the form of $Rx$ and $R^*y$, where $R^*$ denotes the adjoint matrix of $R$. Moreover, the matrix $R$ is typically so large and dense that it cannot in practice be stored in memory. Therefore, in order for the algorithm to be fast and memory efficient, $Rx$ and $R^*y$ must be computed efficiently for all $x \in \mathbb{R}^{n_2^2}$ and $y \in \mathbb{R}^{n_2^2}$ without storing $R$ explicitly. The efficiency of these matrix vector multiplications depends critically on the structure of matrix $R$, which is determined by the sampling set $\Omega$. The cardinality of $\Omega_1$, i.e., the number of projections, determines the radiation exposure of the patient and the time needed for tomography; hence, we want the cardinality of $\Omega_1$ to be small. Then for a fixed cardinality of $\Omega_1$, the elements of $\Omega_1$ and $\Omega_2$ can be chosen in such a way to satisfy our needs. Once a good sensing matrix $R$ is chosen, we solve (3.1) using the optimization algorithm [25] described in Section 4.

We now describe how to select $\Omega_1$ and $\Omega_2$. Without loss of generality choose $T$ and $B$ such that $n_d = \frac{2B}{T}$ and

$$F_s(u,v) = T^2 \sum_{n_1=-\lfloor \frac{B}{T} \rfloor}^{\lfloor \frac{B}{T} \rfloor} \sum_{n_2=-\lfloor \frac{B}{T} \rfloor}^{\lfloor \frac{B}{T} \rfloor} f(n_1 T, n_2 T) e^{-2\pi i (u n_1 + v n_2) T} = T^2 \sum_{i_1=0}^{n_d-1} \sum_{i_2=0}^{n_d-1} f_{i_1,i_2} e^{-2\pi i (u (i_1 - \frac{n_d}{2}) + v (i_2 - \frac{n_d}{2})) T}.$$  

Given $(u,v)$, let $w(u,v) = \sqrt{u^2 + v^2}$ and $\theta(u,v) = \tan^{-1}(v,u)$, where $\tan^{-1}(v,u) : \mathbb{R}^2 \to [-\pi, \pi]$ gives the angle coordinate of the polar representation of the cartesian point $(u,v)$. Then

$$e_{(u,v)} + \hat{S}_\theta(u,v) = F_s(u,v) = T^2 \sum_{i_1=0}^{n_d-1} \sum_{i_2=0}^{n_d-1} f_{i_1,i_2} e^{-2\pi i (u (i_1 - \frac{n_d}{2}) + v (i_2 - \frac{n_d}{2})) T}. \quad (3.3)$$

Thus, when $(u,v) \in \mathbb{R}^2$ is fixed, corresponding row in the $R$ matrix formed according to Equation 3.3.

Moreover, $F_s(u,v)$ is periodic in both directions with a period $\frac{1}{T} = \frac{\frac{2B}{T}}{\frac{2B}{T}}$. Hence, it is sufficient to measure Fourier coefficients of the object for $\{(u,v) : u \in [-\frac{n_d}{2}, \frac{n_d}{2}], v \in [-\frac{n_d}{2}, \frac{n_d}{2}]\}$. We define two sets, which together form $\Omega$:

$$V = \{ (\epsilon_x, \epsilon_y) : \epsilon_y = \frac{l}{4B}, 0 \leq l \leq n_d; \epsilon_x = \epsilon_y \frac{2m}{n_d}, -\frac{n_d}{2} \leq m < \frac{n_d}{2} \}, \quad (3.4)$$

$$H = \{ (\epsilon_x, \epsilon_y) : \epsilon_x = \frac{l}{4B}, 0 \leq l \leq n_d; \epsilon_y = \epsilon_x \frac{2m}{n_d}, -\frac{n_d}{2} \leq m < \frac{n_d}{2} \}. \quad (3.5)$$
Theorem 3.1. Let \( R \in \mathbb{C}^{\Omega \times n_d^2} \) be the pseudo-polar Fourier matrix formed according to (3.3) using \( \Omega = V \cup H \), where \( V \) and \( H \) are given in (3.4) and (3.5), respectively. Then for all \( x \in \mathbb{R}^{n_d \times n_d} \) and \( y \in \mathbb{R}^{|\Omega|} \), \( Rx \) and \( R^*y \) can be computed in \( O(80n_d^2 \log(n_d)) \) arithmetic operations.

Proof. Please see Appendix B for the proof. \( \square \)

Note that computational complexity of the operations we will be using in the first-order reconstruction algorithm is comparable to that of FFT and Fractional Fourier Transform [2], which are \( O(5n \log(n)) \) and \( O(30n \log(n)) \), respectively, for vectors in \( \mathbb{R}^n \).

Remark 3.1. Properly choosing the measurement matrix \( A \) speeds up the reconstruction process significantly.

To give reader a rough estimate, the method in [38] requires 60 minutes on a PC with 4.0 GB memory and 3.16 GHz CPU to reconstruct the interior of a 256 × 256 Shepp-Logan image; whereas our method can reconstruct the entire 256 × 256 Shepp-Logan phantom in 326, 294 and 265 seconds using 256, 128 and 64 projection angles, respectively, on a PC with 3.0 GB memory and 2.0 GHz CPU!

4. Reconstruction Algorithm. In Section 3, CT projection constraints were formulated as \( Rf^d = b + \epsilon \), where \( f^d \) is the digital image of the continuous target object \( f \), \( R \) is the Partial Pseudo Polar Fourier matrix formed using (3.3) for all \( (u, v) \in \Omega \subset V \cup H \) (each \( (u, v) \in \Omega \) corresponds to a row in \( R \) ), \( b \in \mathbb{R}^{|\Omega|} \) is the measurement vector such that for each \( (u, v) \in \Omega \) the corresponding element of \( b \) is equal to \( \hat{S}_{\theta(u,v)} (w(u,v)) \) as in (3.3) and \( \epsilon \approx 0 \) is the measurement noise vector.

Let \( \Phi \) be the discrete wavelet transform matrix. Since empirical evidence suggests that wavelet transforms result in sparse representations for natural images, we assume that \( f^d \) has a sparse wavelet representation, i.e., \( \gamma^* := \Phi f^d \) is sparse. Since different anatomical structures are supposed to show uniform characteristics, we also exploit the fact that CT images of internal organs are expected to be piecewise continuous, i.e., \( \Phi^{-1} \gamma^* = f^d \) is expected to have small total variation [33], where \( \Phi^{-1} \) is the inverse wavelet transform. Hence,
we minimize a weighted combination of the $\ell_1$-norm of the wavelet coefficients, $\|\gamma\|_1$, the total variation norm of the digital image, $TV(\Phi^{-1}\gamma)$, and the $\ell_2$-norm square of the measurement error $R\Phi^{-1}\gamma - b$, i.e., for some parameters $\alpha > 0$ and $\beta > 0$, we solve the following optimization problem

$$\min_{\gamma} \alpha TV(\Phi^{-1}\gamma) + \beta \|\gamma\|_1 + \frac{1}{2} \|R\Phi^{-1}\gamma - b\|_2^2. \quad (4.1)$$

In this paper, we solve (4.1) using the algorithm developed by Ma et al. in [25], which is described below. Throughout this section $\Phi$ represents the single level discrete 1-D Haar wavelet transform and $A = R\Phi^{-1}$.

Total variation $TV(f^d)$ of the digital image, $f^d$, is defined as $TV(f^d) = \sum_i \sum_j ((\nabla_1 f^d_{ij})^2 + (\nabla_2 f^d_{ij})^2)^{1/2}$, where $\nabla_1$ and $\nabla_2$ denote the forward finite difference operators on the first and second coordinates, respectively. Let $f^d \in \mathbb{R}^{n_1 \times n_2}$ represent 2D image of $n_1 \times n_2$ pixels. Then $L = (\nabla_1, \nabla_2) : \mathbb{R}^{n_1 \times n_2} \to \mathbb{R}^{n_1 \times n_2} \times \mathbb{R}^{n_1 \times n_2}$ denotes the discrete finite difference operator and its suboperator $L_{ij}(f^d) = (\nabla_1 f^d_{ij}, \nabla_2 f^d_{ij})$ is such that $TV(f^d) = \sum_{ij} \|L_{ij}(f^d)\|_2$. Finally, let $h(.) = \frac{1}{2} \|A - b\|_2$. With this new notation, (4.1) can be written as

$$\min_{\gamma} \alpha \sum_{ij} \|L_{ij}(\Phi^{-1}\gamma)\|_2 + \beta \|\gamma\|_1 + h(\gamma). \quad (4.2)$$

In [25], Ma et al. suggest using the following fixed point iterations to solve (4.2):

$$s^{(k+1)} = \gamma^{(k)} - \tau_1 \left(\alpha \Phi \sum_{ij} L_{ij}^* (y^{(k)}) + \nabla h(\gamma^{(k)})\right),$$

$$t_{ij}^{(k+1)} = y^{(k)} + \tau_2 L_{ij} (\Phi^{-1}\gamma^{(k)}),$$

$$\gamma^{(k+1)} = \text{sign} \left(s^{(k)}\right) \odot \max \left\{0, |s^{(k)}| - \tau_1 \beta\right\},$$

$$y_{ij}^{(k+1)} = \min \left\{\frac{1}{\tau_2}, \|t_{ij}^{(k)}\|_2 / \|t_{ij}^{(k)}\|_2\right\} t_{ij}^{(k)}$$

for $k \geq 0$, starting from a set of initial points $\gamma^{(0)}$, $y^{(0)}$, $s^{(0)}$, $t^{(0)}$, where sign : $\mathbb{R}^n \to \mathbb{R}^n$ such that

$$(\text{sign}(x))_i = \begin{cases} 
+1, & \text{if } x_i > 0; \\
0, & \text{if } x_i = 0; \\
-1, & \text{if } x_i < 0.
\end{cases}$$

and $\odot$ is a componentwise multiplication operator. See [25] for a justification of the 4-step iterative method and its convergence proofs.

5. Results. In the following we experimentally demonstrate the potential of the proposed method.

5.1. Shepp-Logan Head Phantom Test Object. We used a variant of Shepp-Logan head phantom to test the quality of the reconstruction using the proposed method. The variant of Shepp-Logan head phantom we used consisted of ten ellipses as described in Table 5.1. $I$ is the additive intensity value of the ellipse, $a$ is the length of the horizontal semi-axis of the ellipse, $b$ is the length of the vertical semi-axis of the ellipse, $x_0$ is the x-coordinate of the center of the ellipse, $y_0$ is the y-coordinate of the center of the ellipse and $\phi$ is the angle (in degrees) between the horizontal semi-axis of the ellipse and the x-axis of the image.

Suppose the target object is an ellipse: $f(x, y) = \begin{cases} 
I & x^2/a^2 + y^2/b^2 \leq 1; \\
0 & x^2/a^2 + y^2/b^2 > 1,
\end{cases}$

then it is easy to show that projection $P_b(t)$ of the target object is given by

$$P_b(t) = \begin{cases} 
2I \frac{ab}{c(\theta)} \sqrt{c^2(\theta) - t^2} & |t| \leq c(\theta), \\
0 & |t| > c(\theta),
\end{cases} \quad (5.1)$$
where \( c(\theta) = a^2 \cos^2(\theta) + b^2 \sin^2(\theta) \). Moreover, if the target object is centered at \((x_0, y_0)\) and rotated by \( \phi \) counterclockwise, then the new projections, \( P_{\phi}'(t) \), are obtained using the projections of the centered and unrotated ellipse, \( P_{\phi}(t) \), as follows:

\[
P_{\phi}'(t) = P_{\phi}(t - \sqrt{x_0^2 + y_0^2} \cos \left( \tan^{-1} \left( \frac{y_0}{x_0} \right) - \phi \right)).
\]

Thus using equations (5.1) and (5.2), projection values for Shepp-Logan head phantom can be calculated.

Since \( P_{\phi}(t) \) is known, \( \hat{P}_{\phi}(w) \) can be calculated using equation (3.2) for any \( |w| \leq W \). We chose \( W = 128 \), i.e., \( \|F(u, v)\| \approx 0 \), when \( |u| > 128 \) or \( |v| > 128 \), and we set \( T = \frac{1}{2W} = \frac{1}{256} \). Since \( B = 1 \) for the Shepp-Logan head phantom, size of the reconstructed image is chosen to be \( 512 \times 512 \), i.e. \( n_d = \frac{2B}{W} = 512 \).

### 5.2. Noise-Free Reconstruction Results.

According to the setup described in Section 3, the number of projections angles should be equal to \( |\Omega_1| = 2n_d = 1024 \). In order to test the reconstruction abilities of the proposed method with undersampled measurements, we subsampled angles from \( \Omega_1 \) with a factor \( \Delta \): For a given \( \Delta \), the set of projection angles used is \( \tilde{\Omega}_1 := \tilde{\Omega}_1^V \cup \tilde{\Omega}_1^W \subset \Omega_1 \) were chosen as follows:

\[
\tilde{\Omega}_1^V = \{ \tan^{-1}(n_d, 2m) | m = -n_d/2 + k\Delta, \ 0 \leq k < \frac{n_d}{\Delta} \},
\]

\[
\tilde{\Omega}_1^W = \{ \tan^{-1}(2m, n_d) | m = n_d/2 - k\Delta, \ 0 \leq k < \frac{n_d}{\Delta} \},
\]

\[
\tilde{\Omega}_1 = \tilde{\Omega}_1^V \cup \tilde{\Omega}_1^W.
\]

We used 4 different values of \( \Delta \): 8, 16, 32, 64, which is equivalent to using 128, 64, 32 and 16 projection angles for the reconstruction. Figure 5.1 displays the variant of the Shepp-Logan image corresponding to the parameters in Table 5.1. Figure 5.2(a), Figure 5.2(b), Figure 5.2(c) and Figure 5.2(d) are reconstructed using the proposed method with CT data coming from equations (5.1) and (5.2). Figure 5.2(a) also recovers the phantom almost perfectly, except for small artifacts only within one ellipse. When 32 or 16 projection angles are used, the edges of the reconstructed images, Figure 5.2(b) and Figure 5.2(d), are not as sharp as previous reconstructions.

In order to compare our method with the commonly used Filtered Back Projection (FBP) [20], we have also included figures reconstructed by FBP using undersampled projection data. Figure 5.3(a) and Figure 5.3(b) are FBP reconstructions using 64 and 32 projection angles, respectively. These experimental results clearly show that our method should be preferred to FBP when undersampled data is used, because all the images reconstructed by our technique are free of aliasing artifacts and have all the desired features of the Shepp-Logan phantom. We define the relative reconstruction error by \( \|f^* - f_d^d\|_F / \|f_d^d\|_F \), where \( f^* = \Phi^{-1}y^* \) and \( y^* \) are the digital image and corresponding wavelet coefficients produced by our method, \( f_d^d \) is the digitized original Shepp-Logan Phantom image and \( \|.\|_F \) denotes the Frobenius norm. In Table 5.2, we list relative reconstruction errors for different number of projection angles used for reconstruction, i.e., \( |\Omega_1| \).

---

### Table 5.1: A Variant of Shepp-Logan Head Phantom Parameters

<table>
<thead>
<tr>
<th>Ellipse#</th>
<th>a</th>
<th>b</th>
<th>x₀</th>
<th>y₀</th>
<th>φ</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.69</td>
<td>.92</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>.8</td>
<td>.6624</td>
<td>.8740</td>
<td>0</td>
<td>.0184</td>
</tr>
<tr>
<td>3</td>
<td>.2</td>
<td>1100</td>
<td>.3100</td>
<td>.22</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>.2</td>
<td>1600</td>
<td>.4100</td>
<td>-2.2</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>.1</td>
<td>2100</td>
<td>.2500</td>
<td>0</td>
<td>.35</td>
</tr>
<tr>
<td>6</td>
<td>.1</td>
<td>.0460</td>
<td>.0460</td>
<td>0</td>
<td>.1</td>
</tr>
<tr>
<td>7</td>
<td>.1</td>
<td>.0460</td>
<td>.0460</td>
<td>0</td>
<td>.1</td>
</tr>
<tr>
<td>8</td>
<td>.1</td>
<td>.0460</td>
<td>.0280</td>
<td>-0.08</td>
<td>-605</td>
</tr>
<tr>
<td>9</td>
<td>.1</td>
<td>.0230</td>
<td>.0230</td>
<td>0</td>
<td>-606</td>
</tr>
<tr>
<td>10</td>
<td>.1</td>
<td>.0230</td>
<td>.0460</td>
<td>.06</td>
<td>-605</td>
</tr>
</tbody>
</table>

---

**Fig. 5.1:** Shepp-Logan Head Phantom

**Table 5.1:** A Variant of Shepp-Logan Head Phantom Parameters
Table 5.2: Relative reconstruction error results for CT data without noise

| $|\hat{\Omega}_1|$ | 128 | 64 | 32 | 16 |
|-----------------|------|----|----|----|
| $\|f^* - f^\theta\|_F/\|f^\theta\|_F$ | 0.1113 | 0.1214 | 0.1453 | 0.2296 |

Fig. 5.2: $512 \times 512$ Shepp-Logan Head Phantom reconstructed from CT data without noise using $|\hat{\Omega}_1|$ projection angles

Fig. 5.3: Shepp-Logan Head Phantom Filtered Back Projection reconstruction from CT data without noise using $|\hat{\Omega}_1|$ projection angles

5.3. Noisy Reconstruction Results. In this section we tested the robustness of our modeling technique to the noise in CT data. We generated observations according to $P^{\text{noisy}}_{\theta}(t) = P_{\theta}(t) + \eta_{\theta,t}$, where $P_{\theta}(t)$ is defined in (5.2), and $\eta_{\theta,t} \sim \text{Normal}(0,1)$. We assumed that the measurement noise set, $\{\eta_{\theta,t} : \theta \in \hat{\Omega}_1, \forall t\}$, consists of independent Normal random variables with standard deviation $\sigma(\theta,t)$.

The standard deviation of the measurement error was assumed to be proportional to the X-ray attenuation. Let $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ denote the X-ray attenuation function of the target object, $P_{\theta}(t)$ denote the radon transform (line integral) value of $f$ corresponding to the line defined by $(\theta,t)$ and $P_{\theta}^{m}(t) := \ln\left(\frac{N_0}{N_{\text{out}}(\theta,t)}\right)$ denote the measured value (see [20]), where $N_0$ is the average number of photons sent from each source within a given time when X-ray tube current is $A_0$ and $N_{\text{out}}(\theta,t)$ is the number of photons reaching the detector through the object from the beam of photons corresponding to $(\theta,t)$. The noise model widely accepted in the literature of X-ray imaging suggests that $E[P_{\theta}^{m}(t)] = P_{\theta}(t)$ and $\text{Var}[P_{\theta}^{m}(t)] = \sigma_{P_{\theta}^{m}(t)}^2$, where $E[.]$ and $\text{Var}[.]$ denote the expectation and variance operators, respectively (see [26, 20]). Hence, in order to create the noisy projection data we assumed the following relation: $P_{\theta}^{m}(t) = P_{\theta}(t) + \eta_{\theta,t}$ for all $\theta$ and $t$, where $\eta_{\theta,t} \sim \text{Normal}(0,1)$ and

$$\sigma(\theta,t) = \xi \sqrt{\frac{e^{P_{\theta}(t)}}{N_0}},$$

when the X-ray tube current is $\frac{A_0}{\xi^2}$ (see [4, 9]). Here, the parameter $\xi$ controls the signal-to-noise ratio (SNR) of the CT data.

Given $\xi > 0$, $\eta_{\theta,t}$ has a standard deviation $\sigma(\theta,t)$ as given in (5.6). Figure 5.4 displays reconstructions using
the proposed method with 128 projection angles, i.e. $\Delta = 8$, for different $\xi$ values and Table 5.3 lists relative reconstruction errors for changing $\xi$ values. Figure 5.5(a) plots relative reconstruction error versus log $\xi$. Clearly the proposed solution is robust to significant levels of noise.

<table>
<thead>
<tr>
<th>Figure #</th>
<th>5.4(a)</th>
<th>5.4(b)</th>
<th>5.4(c)</th>
<th>5.4(d)</th>
<th>5.4(e)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\xi$</td>
<td>$1 \times 10^{-3}$</td>
<td>$5 \times 10^{-3}$</td>
<td>$1 \times 10^{-2}$</td>
<td>$5 \times 10^{-2}$</td>
<td>$1 \times 10^{-1}$</td>
</tr>
<tr>
<td>$|f^* - f^d|_F/|f^d|_F$</td>
<td>0.1134</td>
<td>0.1143</td>
<td>0.1147</td>
<td>0.1441</td>
<td>0.2596</td>
</tr>
</tbody>
</table>

**Table 5.3**: Relative reconstruction error results for noisy CT data, $\sigma(\theta, t) = \xi \sqrt{\frac{P_\delta(t)}{N_0}}$, $|\hat{\Omega}_1| = 128$ ($\Delta = 8$)

Fig. 5.4: Noisy CT data reconstruction using 128 projection angles, $\sigma(\theta, t) = \xi \sqrt{\frac{P_\delta(t)}{N_0}}$

Fig. 5.5: Relative reconstruction error versus log $\xi$
5.4. Tube Current Reduction vs Undersampled Projections. An alternative way to reduce the total radiation exposure is to reduce the current in the X-ray tube. In the rest of this section, we compare the following two alternatives:

1) Acquire all necessary projections but reduce the tube current. Then, reconstruct the image from the complete data set using denoising.
2) Acquire fewer projections than necessary according to the proposed method in this paper. Then, reconstruct the image using the proposed TV-$\ell_1$ norm minimization, i.e. Problem 4.1.

In the first case, the decrease in the tube current will reduce the SNR, but the reconstruction problem will not be under-determined. In the second case, the reconstruction problem is under-determined, and it may be expected that the SNR will also decrease due to fewer photons collected. However, CS paradigm suggests that when the measurement matrix $A = R\Phi^{-1}$ satisfies certain regularity conditions, and $\gamma^d = \Phi f^d$ is sparse or $f^d$ is piecewise-constant; solving $\min\{||\gamma||_1 : A\gamma = b\}$ or $\min\{TV(\Phi^{-1}\gamma) : A\gamma = b\}$, respectively, will recover $\gamma^d$ exactly even though the $A\gamma = b$ system is under-determined, i.e. $m << n$. While we could not show that the matrix $A$ proposed in this paper satisfies the regularity conditions, the results in our numerical experiments suggest that $A$ performs well in practice. In order to compare the two methods on the same grounds, we will reconstruct the Shepp-Logan phantom from the noisy projection data while keeping the radiation exposure constant.

Let $A_1$, $A_2$ be the two different tube current levels at which we collect the measurements $P^{m_1}_\theta(t)$ and $P^{m_2}_\theta(t)$, respectively. We assume that the measurement noise is given by the model in Section 5.3, i.e., for $i \in \{1, 2\}$,

$$P^{m_i}_\theta(t) = P_\theta(t) + \eta_{\theta,t}$$

for all $\theta$ and $t$, where $\eta_{\theta,t} \sim \sqrt{\frac{\sigma_{\theta,t}}{N_i}} \text{Normal}(0,1)$ and $N_i$ is the average number of photons sent from each source within a given time when X-ray tube current is set at $A_i$. Let $\sigma_1(\theta,t)$, $\sigma_2(\theta,t)$ be the corresponding standard deviations of $P^{m_1}_\theta(t)$ and $P^{m_2}_\theta(t)$, respectively. Then we have the following relation: $\frac{\sigma_1(\theta,t)}{\sigma_2(\theta,t)} = \sqrt{\frac{A_2}{A_1}}$ (see [4, 9]).

First, we generated the projection data $P^{m_1}_\theta(t)$ using only 64 projection angles for the $512 \times 512$ the Shepp-Logan phantom. Second, the projection data $P^{m_2}_\theta(t)$ for all 512 angles is generated; however, in order to keep the radiation amount constant, the tube current was decreased by a factor of 8. Thus, if the average number of photons sent in the first case (64 projection angles) is $N_1$, then it should be $\frac{N_1}{8}$ in the second case, where 512 projection angles were used. Then we have the following data model:

$$P^{m_1}_\theta(t) = P_\theta(t) + \eta_{\theta,t},$$
$$P^{m_2}_\theta(t) = P_\theta(t) + 2\sqrt{2} \eta_{\theta,t},$$

where $\eta_{\theta,t} \sim \sqrt{\frac{\sigma_{\theta,t}}{N_i}} \text{Normal}(0,1)$.

Next, we first reconstructed $512 \times 512$ the Shepp-Logan phantom using $P^{m_1}_\theta(t)$ consisting of 64 projection angles data with the proposed method in this paper and later using $P^{m_2}_\theta(t)$ consisting of 512 projection angle data both with filtered backward projection (FBP) and the proposed method in this paper. In Figure 5.6, we display the 3 reconstructions and the $512 \times 512$ original Shepp-Logan phantom. The relative error for the reconstructions in Figure 5.6(a) and Figure 5.6(b) are 18% and 13%, respectively. Figure 5.6(c) is the reconstruction via Filtered Backward Projection (FBP) using all 512 angle projection data. Furthermore, we have also applied Rudin-Osher-Fatemi (ROF) Denoising [33] to the image in Figure 5.6(c), which gave Figure 5.6(d). Clearly, our proposed method gives the closest reconstruction to the original phantom shown in Figure 5.1.
(a) Reconstruction using 512 angles data with the proposed method (tube current A/8)

(b) Reconstruction using 64 angles data with the proposed method (tube current A)

(c) Reconstruction using 512 angles data with FBP (tube current A/8)

(d) ROF Denoising applied to image in Figure 5.6(c)

Fig. 5.6: Reconstruction of 512 × 512 Shepp-Logan Phantom
5.5. **Reconstruction using randomly chosen projection angles.** Given the undersampling factor \( \Delta \), instead of uniformly choosing the projection angles as in equations (5.3), (5.4) and (5.5), this time we selected \( \frac{2n_d}{\Delta} \) angles randomly among \( 2n_d \) angles defined in (3.4) and (3.5), with equal probability. Since the random set is still a subset of the projection angles in (3.4) and (3.5), right and left multiplication with the associated random measurement matrix \( A \) can still be done efficiently.

We set \( \Delta = 16 \), i.e., we used 64 random projection angles instead of 512, and replicated the experiment 5 times to reconstruct \( 512 \times 512 \) Shepp-Logan phantom. Randomly choosing projection angles can lead to large unsampled regions in the frequency domain, possibly leading to bad reconstruction. However, given that \( \frac{2n_d}{\Delta} \) angles would be selected, choosing them with a constant angular increment of \( \frac{\pi \Delta}{2n_d} \) (i.e. equiangular rays) minimizes the maximum unsampled area in the frequency domain. The results in Table 5.4, Figure 5.7 and Figure 5.8, also support this hypothesis in terms of reconstruction quality. Reconstructions of \( 512 \times 512 \) Shepp-Logan Phantom using 64 equiangular rays with noise free and noisy data shown in Figure 5.8 have relative reconstruction errors of 0.1214 and 0.1231, respectively.

<table>
<thead>
<tr>
<th>Run No :</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure # :</td>
<td>5.7(a)</td>
<td>5.7(b)</td>
<td>5.7(c)</td>
<td>5.7(d)</td>
<td>5.7(e)</td>
</tr>
<tr>
<td>( | f^{sol} - f^d |_2 / | f^d |_2 ) :</td>
<td>0.1465</td>
<td>0.1873</td>
<td>0.1345</td>
<td>0.1507</td>
<td>0.1842</td>
</tr>
</tbody>
</table>

**Table 5.4:** Relative errors for reconstructions with randomly chosen projection angles

![Fig. 5.7: Reconstruction of 512 x 512 Shepp-Logan Phantom using 64 random projection angles with noiseless data](image)

![Fig. 5.8: Reconstruction of 512 x 512 Shepp-Logan Phantom](image)

6. **Conclusion.** CT is one of the most important and widely used diagnostic tools available to modern medicine. Compressed sensing (CS) is one of the most exciting recent advances in computational science. In
this paper, we integrate the two to propose an efficient reconstruction framework that can significantly reduce the radiation exposure without compromising the image quality.

CS has found more application in the magnetic resonance imaging (MRI) than in CT imaging. The reason is that applying the CS framework to CT is much more challenging. In this paper we bridge the gap by formulating the CT reconstruction problem from limited projection data as a CS problem via the use of partial pseudo-polar Fourier transform. The effectiveness of the proposed method, both in terms of computation time and quality of reconstruction, is evident in the numerical tests we conducted using a variant of the Shepp-Logan phantom. Moreover, one can combine the proposed method with other dose reduction strategies such as ECG modulated tube current to achieve further radiation reduction.

Appendix A. Fractional Fourier Transform.

Definition: Given a signal, \( x \in \mathbb{C}^N \) and arbitrary \( \alpha \), Fractional Fourier Transform of \( x \) is given by

\[
X_k = \sum_{n=0}^{N-1} x_n e^{-2\pi i \frac{k \alpha n}{N}} \quad \text{for} \quad 0 \leq k < N.
\]

Evaluation of Fractional Fourier Transform on \( x \in \mathbb{C}^N \) can be done in \( O(30N \log(N)) \) operations [2]. Since \( 2kn = k^2 + n^2 - (k - n)^2 \),

\[
X_k = \sum_{n=0}^{N-1} x_n e^{-2\pi i \frac{k \alpha n}{N}} = \sum_{n=0}^{N-1} x_n e^{-\pi i \frac{k^2 + n^2 - (k - n)^2}{2N} \alpha} = e^{-\pi i \frac{k^2}{2N} \alpha} \sum_{n=0}^{N-1} x_n e^{-\pi i \frac{n^2}{2N} \alpha} e^{\pi i \frac{(k-n)^2}{2} \alpha}
\]

Define \( z_n = e^{\pi i \frac{n^2}{2N} \alpha} \) and \( y_n = x_n z_k^* \) for \( 0 \leq n < N \), where \((.)^*\) denotes the complex conjugate. Then

\[
X_k = z_k^* \sum_{n=0}^{N-1} y_n z_{k-n}
\]

Thus Fractional Fourier Transform of \( x \) can be computed by multiplying \( x \) with \( z^* \) elementwise and taking the convolution of the resulting vector with \( z \), then finally by multiplying by \( z^* \). In order to calculate the convolution efficiently, Finite Fourier Transform (FFT) is used. Since FFT computes circular convolutions, \( x \) and \( z \) sequences are modified, to get rid of the circular effects. First, \( x \) is zero padded to length \( 2N \) to obtain \( \otimes \). Second, \( z \) is extended to \( 2N \) as follows to satisfy periodicity requirement for FFT:

\[
z_n = \begin{cases} 
e^{-\pi i \frac{n^2}{2N} \alpha}, & 0 \leq n < N; \\ e^{-\pi i \frac{(2N-n)^2}{2N} \alpha}, & N \leq n < 2N. \end{cases}
\]

Now the above convolution can be efficiently computed using fast Fourier transform by computing the FFT of \( \otimes \) and \( \otimes \), then multiplying them elementwise and taking the inverse FFT of the resulting vector. Finally, the first \( N \) components of the resulting \( 2N \) vector are taken and multiplied by \( z^* \) to compute \( X \).

Since we are using 2 FFT of \( 2N \) vectors and a final inverse FFT of the resulting vector in \( \mathbb{C}^{2N} \), Fractional Fourier Transform takes \( O(30N \log(N)) \) operations, since FFT of \( N \) dimensional vector takes \( O(5N \log(N)) \) operations.

Appendix B. Proof of Theorem 3.1.

Proof. Let \( x \in \mathbb{R}^{n_d \times n_d} \). Below, it is shown that \( R_V x \) can be computed efficiently and similar arguments are also valid for \( R_H x \). For set \( V \), let the inner product of \( x \) with the row of \( R_V \) corresponding to \( \epsilon_x = \frac{2m}{4Bn_d} \) and \( \epsilon_y = \frac{1}{4B} \) be \( X_{m,l} \). Then,

\[
X_{m,l} = T^2 \sum_{i_1=0}^{n_d-1} \sum_{i_2=0}^{n_d-1} x_{i_1,i_2} e^{-2\pi i \frac{2mL}{4Bn_d} (i_1 - \frac{n_d}{2}) + \frac{i_1}{4B} (i_2 - \frac{n_d}{2}) + \frac{2\pi i}{n_d}}
\]

\[
= T^2 \sum_{i_1=0}^{n_d-1} \sum_{i_2=0}^{n_d-1} x_{i_1,i_2} e^{-i \left( \frac{2\pi m}{n_d} i_1 + \frac{\pi}{n_d} i_2 \right) + \pi (\frac{m}{n_d} + \frac{1}{2})}.
\]
Define $\hat{x}$ such that
\[
\hat{x}_{i_1,l} = \sum_{i_2=0}^{n_d-1} x_{i_1,i_2} e^{-2\pi i \frac{i_1 i_2}{n_d}} = \sum_{i_2=0}^{n_d-1} x_{i_1,i_2} e^{-2\pi i \frac{i_1 l}{n_d}},
\]
(B.1)
where $x_{i_1,i_2} = \begin{cases} x_{i_1,i_2} & 0 \leq i_2 < n_d \\ 0 & n_d \leq i_2 < 2n_d \end{cases}$.

By zero padding $x$, for a given $i_1$ ($1 \leq i_1 < n_d$) we can use FFT for (B.1) to compute $\hat{x}_{i_1,l}$ for $0 \leq l \leq n_d$, efficiently. Complexity bound of this computation is $O(5(2n_d \log(2n_d))$ for each $0 \leq i_1 < n_d - 1$. Thus $O(10n_d^2 \log(n_d))$ operations are needed in total to compute $\hat{x}$. Now, $X_{m,l}$ can be written in terms of $\hat{x}$:
\[
X_{m,l} = T^2 e^{i\pi l (\frac{m}{n_d} + \frac{1}{2})} \sum_{i_1=0}^{n_d-1} \hat{x}_{i_1,l} e^{-2\pi i \frac{m+l}{n_d}(\frac{i_1}{n_d})},
\]
Given $l$ ($0 \leq l \leq n_d$), in order to calculate $X_{m,l}$ efficiently for $-n_d \leq m < \frac{n_d}{2}$, Fractional Fourier Transform is used. Complexity bound of this computation is $O(30n_d \log(n_d))$ for each $0 \leq l \leq n_d$. Thus $O(40n_d^2 \log(n_d))$ operations are needed in total to compute $R_V^* x$. For the set $H$ same arguments above still hold. Therefore, $R_x$ can be computed using arithmetic operations in the order of $80n_d^2 \log(n_d)$.

Now given $y = \begin{bmatrix} y_V \\ y_H \end{bmatrix}$, where $y_V \in \mathbb{C}^{[\frac{n}{d}]}$ and $y_H \in \mathbb{C}^{[\frac{n}{d}]}$, adjoint multiplication $R^* y = R_{V^*} y_V + R_{H^*} y_H$ can be computed efficiently. Below, it will be shown that $R_{V^*} y_V$ can be computed efficiently and similar arguments are also true for $R_{H^*} y_H$. For set $V$, let the inner product of $y_V$ with the row of $R^* V$ corresponding to $\epsilon_x = \frac{2im}{2n_d}$ and $\epsilon_y = \frac{l}{n_d}$ be $Y_{i_1,i_2}$. Then,
\[
Y_{i_1,i_2} = T^2 \sum_{m=-\frac{n_d}{2}}^{\frac{n_d}{2}-1} \sum_{l=0}^{n_d-1} y_{m,l} e^{2\pi i \frac{l}{n_d}(i_1 - \frac{n_d}{2}) + \frac{i}{2n_d}(i_2 - \frac{n_d}{2})},
\]
\[
= T^2 \sum_{l=0}^{n_d-1} e^{2\pi i \frac{l}{n_d}(i_2 - \frac{n_d}{2})} \sum_{m=-\frac{n_d}{2}}^{\frac{n_d}{2}-1} y_{m,l} e^{2\pi i \frac{l}{n_d}(i_1 - \frac{n_d}{2})}.
\]
Let
\[
\hat{y}_{i_1,l} = \sum_{m=-\frac{n_d}{2}}^{\frac{n_d}{2}-1} y_{m,l} e^{-2\pi i \frac{m}{n_d}(i_1 - \frac{n_d}{2})(-\frac{l}{n_d})},
\]
(B.2)
Given $l$ ($0 \leq l \leq n_d$), we can use Fractional Fourier Transform for (B.2) to compute $\hat{y}_{i_1,l}$ for $0 \leq i_1 < n_d$, efficiently. The complexity bound for this operation is $O(30n_d \log(n_d))$ for each $0 \leq l \leq n_d$. Thus $O(30n_d^2 \log(n_d))$ operations are needed in total to compute $\hat{y}$. Now $Y_{i_1,i_2}$ can be written in terms of $\hat{y}$:
\[
Y_{i_1,i_2} = T^2 \sum_{l=0}^{n_d-1} (\hat{y}_{i_1,l} e^{\frac{2\pi i l}{n_d}}) e^{2\pi i \frac{l}{2n_d} \frac{i_1}{n_d}} = T^2 \sum_{l=0}^{2n_d-1} (\hat{y}_{i_1,l} e^{-\pi i \frac{l}{n_d}}) e^{2\pi i \frac{l}{2n_d} \frac{i_1}{n_d}}
\]
(B.3)
where $\hat{y}_{i_1,l} = \begin{cases} \hat{y}_{i_1,l} & 0 \leq l \leq n_d \\ 0 & n_d \leq l < 2n_d \end{cases}$.

By zero padding $\hat{y}$, given $i_1$ ($0 \leq i_1 < n_d$), we can use inverse FFT for (B.3) to compute $Y_{i_1,i_2}$ for $0 \leq i_2 < n_d$, efficiently. This can be done in $O(5(2n_d \log(2n_d)))$ time for each $0 \leq i_1 < n_d$. Thus $O(40n_d^2 \log(n_d))$ operations are needed in total to compute $R_V^* y_V$. Moreover, same arguments above still hold for $R_{H^*} y_H$. Therefore, $R^* y$ can be computed within the order of $80n_d^2 \log(n_d)$ arithmetic operations. \[\square\]
REFERENCES


