RECONSTRUCTION OF CT IMAGES FROM PARSIMONIOUS ANGULAR MEASUREMENTS VIA COMPRESSED SENSING

N. S. AYBAT* AND A. CHAKRABORTY†

Abstract. Computed Tomography is one of the most popular diagnostic tools available to medical professionals. However, its diagnostic power comes at a cost to the patient- significant radiation exposure. The amount of radiation exposure is a function of the number of angular measurements necessary to successfully reconstruct the imaged volume. Compressed sensing on the other hand is a technique that allows one to reconstruct signals from a very limited set of samples provided that the target signal is compressible in a transform domain. Combining the two gives clinicians the benefits of CT while at the same time limiting the risk posed by excessive radiation. In this work we formulate the computed tomography reconstruction problem within a compressed sensing framework using partial pseudo-polar Fourier transform. Simulated results indicate that the number of angular projections and hence the associated radiation can be cut by a factor of 4 or more without noticeable loss of image quality.

1. Introduction. Computed Tomography (CT) is one of the most powerful diagnostic tools that clinicians have today. Since the inception of CT scans in the 1970's, it 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 for adults and increasingly even for children. It has found diagnostic use for such diverse problems as brain tumors, strokes, pulmonary embolisms, renal stones, abdominal aortic aneurysm etc. Consequently, by some estimates the number of CT scans have increased from a mere 2.8 million scans in 1981 [17] to 62 million in 2006 [18], including 4 million on children.

Over the last 40 years the scanning mechanism has evolved significantly to improve the resolution and the scanning time. However, the basic principle has remained the same. In a conventional CT setup, the X-ray source and the detector is housed in a circular band (also called gantry) which moves circularly around a patient who lies on a motorized table. Up to 64 rows of detectors are used, which gather a series of X-ray images at various angles as the gantry moves around the patient. Each complete movement of the gantry represents a slice of the body. The motorized table is moved to cover the third dimension in order to construct a three dimensional image of the organ of interest. The detectors detect the attenuated x-ray signal after it passes through the organ of interest. This 3D data is then used to reconstruct the images of internal human organs in a non-invasive manner.

Any CT scan exposes the patient to radiation - the precise amount being dependent on the image resolution and quality desired, size of the volume scanned, the patient build and the scanning protocol used. Further, it has been noted that CT requires larger radiation doses than conventional x-ray imaging procedures [18]. For example for an abdominal exam CT requires 50 times more radiation compared to conventional x-ray. Hence, the phenomenal popularity of CT has come at a cost of increased radiation and the associated risks. In fact, CT has contributed disproportionately to the problem of radiation exposure. A study done in Britain for instance noted that while CT constitutes 4% of the diagnostic procedures, they account for 40% of the radiation exposure [19].

It is not surprising then that reducing the radiation exposure without compromising the image quality is a major area of research. Several techniques are being explored include hardware adjustments as well as improved image reconstruction algorithms. We approach this problem from the latter perspective. Our goal is to reconstruct the image using a far fewer number of angular projections than what is mandated by Nyquist's criteria for a perfect reconstruction. Normally this should result in streaking artifacts when using the traditional back projection algorithm for CT image reconstruction [2]. However, we show that using a compressed sensing formulation of the CT reconstruction problem through the use of partial pseudo-polar Fourier formulation, one can still achieve near perfect reconstruction.

In the rest of this paper, we fist briefly introduce the reconstruction problem, followed by a short introduction

*IEOR Department, Columbia University. Email: nsa2106@columbia.edu
†Siemens Corporate Research. Email: amit.chakraborty@siemens.com

of the compressed sensing framework. In Section 4, we show how to reconstruct CT image within the compressed sensing framework using partial pseudo-polar Fourier transform. In Section 5, we briefly describe the compressed sensing optimization algorithm used here without any details. Finally, we present some results to test the performance of the proposed algorithm and then we conclude.

2. CT Image Reconstruction. Let $f(x, y)$ denote the x-ray attenuation coefficient distribution of a 2D target object and $x\ cos(\theta) + \ y\ sin(\theta) = \ t_1$ represent the trajectory of an x-ray beam, defined as in Figure 2.1. The x-ray attenuation of that beam is represented as $P_\theta(t_1)$. The method that will be described in this paper uses projection data of the target object. In conventional x-ray tomography [1], these projections represent the total amount of attenuation of x-rays while passing through the target object. In other words, the interaction of x-rays with the substance of the object causes the projection data, which can be approximated as a line integral of the x-ray attenuation coefficient of the object. In the case of monochromatic x-ray photons, a line integral of the x-ray attenuation coefficient can be approximated with the log of the ratio of the entering to the leaving photons. Throughout the paper parallel ray projections will

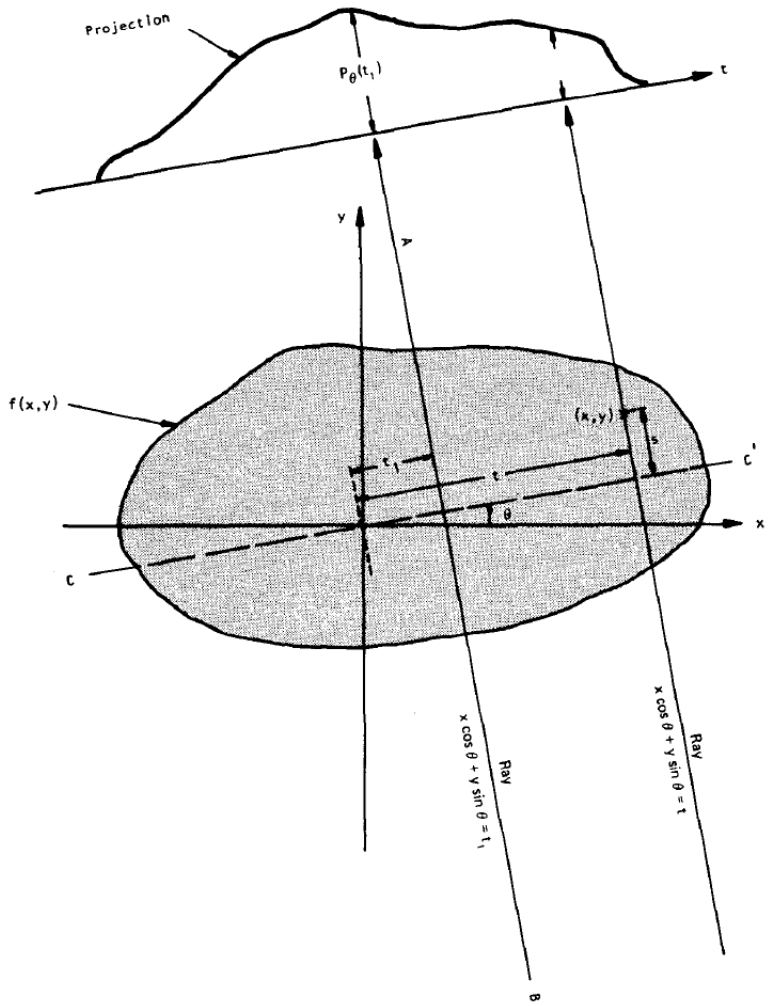


FIG. 2.1. *An object and its projection are shown for an angle of θ. [2]*

be used. Parallel ray projection data can be easily collected. After fixing a constant projection angle, the x-ray source and x-ray detector are moved together along parallel lines on the opposite sides of the target

object. Then repeating this step for a set of projection angles, a set of parallel ray projection data can be obtained.

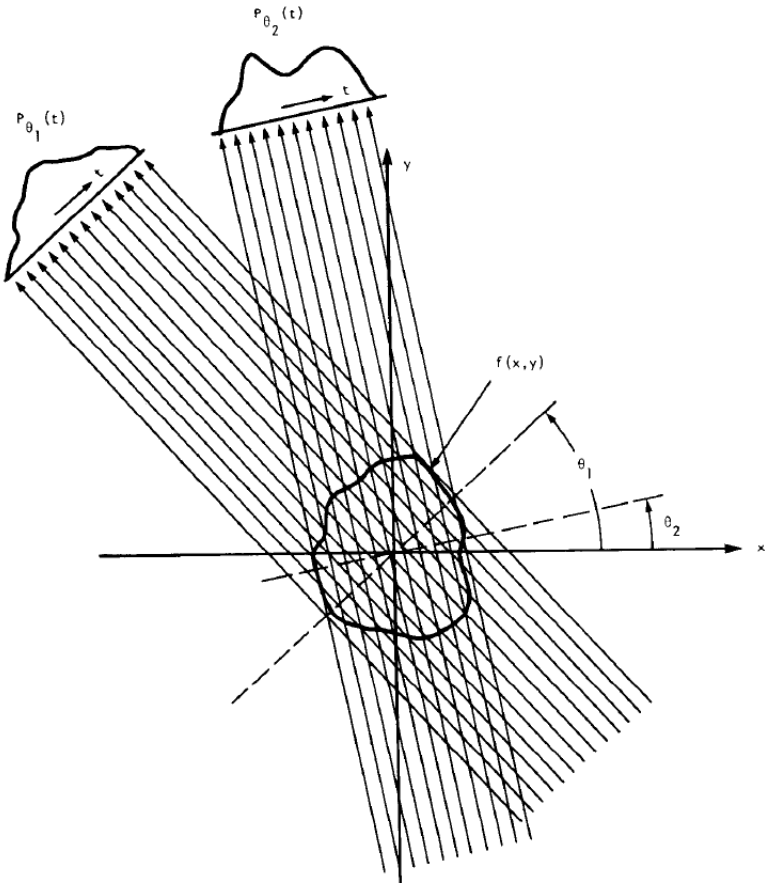


FIG. 2.2. *Parallel projections are taken by measuring a set of parallel rays for a number of different angles. [3]*

In standard CT imaging practice for a perfect reconstruction, the total number of projections of the target object must satisfy the Shannon-Nyquist Sampling Theorem. Since the total radiation exposure of the target object and the time needed for imaging is proportional to the number of view angles, we would like to minimize this number. Compressed Sensing Theory (CS) [5, 6] showes 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 than the Nyquist Rate. Thus, this theory provides us with the possibility of reducing the number of view angles provided the target object is sparse 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 sparse images.

3. Compressed Sensing in General and Imaging. Compressed sensing [5, 6] is an interesting new area of research which has gained enormous popularity of late due to its ability to perfectly reconstruct a *signal* from *a limited number of samples* provided that the signal is sparse in a transform domain. It is being increasingly adopted in a variety of applications (see [12] for a review of the theory and applications). As noted earlier, in the case of CT images, being able to reconstruct images with high quality using a *limited set* of the projection angles translates to less radiation for the patient.

Compressed sensing in general is based upon the fact that most signals are compressible. Let the vector $\bar{u}$ represent a signal. A compression technique (such as JPEG2000 for image compression) compresses the signal by finding some transform Φ (e.g., Fourier or wavelet transform) such that $\Phi\bar{u} = \bar{x}$ is (approximately) sparse,

and saving the locations and values of the nonzero entries of $\bar{x}$ (or those with relatively large magnitudes). To recover $\bar{u}$, one simply applies the inverse transform to $\bar{x}$ to obtain $\bar{u} = \Phi^{-1}\bar{x}$. For ease of notation, we let

$$k = \|\bar{x}\|_0 := \text{the number of non-zeros in } \bar{x},$$
$$n := \text{dimension}(\bar{x}).$$

The entire process of compressed sensing consists of three steps: encoding, sensing, and decoding. In the first step, $\bar{u}$ is encoded into a *smaller* vector $b = R\bar{u}$ of a size $m < n$ by a linear transform R, which is certainly not invertible since $m < n$. Clearly, b contains less information than $\bar{u}$, so it is a *compression* of $\bar{u}$. Since $\bar{u} = \Phi^{-1}\bar{x}$, $b = A\bar{x}$ (for $A = R\Phi^{-1}$) is also a compression of $\bar{x}$. In many applications of compressed sensing, the linear transform R is not calculated by a computer but obtained by certain physical or digital means. The second step is obtaining the measurement b from the imaging system. The third step is to recover $\bar{x}$ (and thus, $\bar{u}$) from b. Since $\bar{x}$ is sparse, it can be recovered as the sparsest solution to the underdetermined system of equations $Ax = b$ unless there exists another *even sparser* solution to these equations. This gives rise to the ℓ_0-problem

$$\min_x \{\|x\|_0 : Ax = b\}.$$

However, this problem is NP-hard [16] for general data, impractical for nearly all real applications. Hence, it is more realistic to solve a computationally tractable alternative such as the ℓ_1-problem

$$\min_x \{\|x\|_1 : Ax = b\}, \tag{3.1}$$

which has also been known to yield sparse solutions under some conditions (see [13, 14, 15] for details). Ideally, we would like to take the least possible number of measurements, that is, m being equal to k. However, we must pay a price for not knowing the locations of the non-zeros in $\bar{x}$ (there are n choose k possibilities!) while still asking for perfect reconstructions of most sparse $\bar{x}$. It was shown in [5, 21] that, when R is Gaussian random and partial Fourier, i.e. we randomly select using a Gaussian pdf a set of Fourier coefficients, (3.1) can recover $\bar{x}$ (technically, with a high probability) from b of a size $m = O(k\log(n/k))$ and $O(k\log(n)^4)$, respectively. In the case of Computed Tomography, it has been shown that perfect reconstruction is possible from an incomplete dataset if we have a known subregion inside a region of interest [20]. Finally, once $\bar{x}$ is recovered, $\bar{u}$ becomes available through $\bar{u} = \Phi^{-1}\bar{x}$.

The exact recovery of $\bar{u}$ by (3.1) requires that $\bar{x}$ be sparse and $b = A\bar{x}$, exactly. In practice, especially in imaging, it is often the case that $\bar{x}$ is only approximately sparse, that is, $\bar{x}$ contains a small number of components with magnitudes significantly larger than those of the rest, which are not necessarily zero. In some applications, it is also the case that the measurements b are contaminated with noise $\mathbf{n}$, i.e., $b = A\bar{x}+\mathbf{n}$. In both cases, an appropriate relaxation of $b = Ax$ should be considered. The topic of compressed sensing stability studies how accurately a model can recover signals under the above conditions.

There are stability results established for model (3.1) and its extension

$$\min_x \{\|x\|_1 : \|Ax - b\|_2^2 \le \sigma^2\}. \tag{3.2}$$

For (3.2), it was shown in [22] that with similar measurements as in the exact recovery case, its solution x^* satisfies

$$\|x^* - \bar{x}\|_2 \le O(\sigma + k^{-1/2}\|\bar{x} - \bar{x}(k)\|_1), \tag{3.3}$$

where $\bar{x}$ is approximately k-sparse and $\bar{x}(k)$ is the k-approximation to $\bar{x}$ by zeroing the $n - k$ insignificant components. Clearly, if $\bar{x}$ is exactly k-sparse and σ is set to 0, then (3.3) reduces to the exact recovery of $x^* = \bar{x}$.

Solving problem (3.2) is equivalent to solving the simpler Lagrange relaxation problem

$$\min_x \mu\|x\|_1 + \frac{1}{2}\|Ax - b\|_2^2. \tag{3.4}$$

This equivalence is in the sense that solving one of the two problems will determine the parameter in the other such that both give the same solution [23]. Given the data A, b, and σ in (3.2), there exist practical ways (see [24], for example) to estimate an appropriate μ in (3.4).

In the following, we formulate the compressed sensing problem for computed tomography as (3.2) through the appropriate choice of A and b.

4. CT Image Reconstruction using CS Framework. The essence of our proposed method is to formulate the CT image reconstruction problem within a compressed sensing framework which involves two main steps. First, we need to compute the measurement matrix A (in (3.2)) from the mapping that represents the attenuation and summation process that generates P_θ when an x-ray traverses through an object. The second step is to derive the measurement b (in (3.2)) from the projections P_θ defined in Section 3 obtained at different angles θ. Once A and b are known, the resulting CS problem (5.1), which can be solved using any optimization algorithm appropriate for (3.2), will be solved using the algorithm developed by Ma et. al [8].

The imaged object is modeled as a two-dimensional distribution of X-ray attenuation coefficient and is represented by a two-dimensional function $f(x, y)$ in the standard x-y coordinate system. In the rest of this paper we still use object and $f(x, y)$ interchangeably. We assume the object to be bounded in the spatial domain, i.e. $f(x, y) = 0$ when $|x| > B$ or $|y| > B$. A line integral of the object corresponds to the total attenuation of an X-ray beam as it travels through the object in a straight line. Given a projection angle $0 \leq \theta < \pi$ in radian, if we define $l_{\theta,t} = \{(x, y) : x\ cos\theta + y\ sin\theta = t\}$ to be the path traveled by a beam of X-ray, then $P_\theta(t) = \int_{l_{\theta,t}} f(x, y)\ ds$ denotes the total attenuation of the X-ray beam along the line $l_{\theta,t}$. For a fixed θ, $P_\theta = \{P_\theta(t)| - \infty < t < +\infty\}$ defines a parallel projection of the object f corresponding to the projection angle θ.

Let $F : \Re^2 \to \mathbb{C}$ denote the Fourier transform of f, i.e. $F(u, v) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x, y)\ e^{-2\pi i(ux+vy)}dx\ dy$. Let $s : \Re^2 \to \mathbb{R}$ denote the two dimensional sampling function: $s(x, y) = T^2 \sum_{n_1=-\infty}^{+\infty}\sum_{n_2=-\infty}^{+\infty} \delta(x-n_1T)\delta(y-n_2T)$, where δ 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,

$$F_s(u, v) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_s(x, y)\ e^{-2\pi i(ux+vy)}dx\ dy, \tag{4.1}$$

$$= T^2 \sum_{n_1=-\infty}^{+\infty}\sum_{n_2=-\infty}^{+\infty} f(n_1T, n_2T)\ e^{-2\pi i(un_1+vn_2)T}, \tag{4.2}$$

$$= T^2 \sum_{n_1=\lfloor\frac{-B}{T}\rfloor}^{\lceil\frac{B}{T}\rceil}\sum_{n_2=\lfloor\frac{-B}{T}\rfloor}^{\lceil\frac{B}{T}\rceil} f(n_1T, n_2T)\ e^{-2\pi i(un_1+vn_2)T}. \tag{4.3}$$

It can be easily shown that $F_s(u, v) = \sum_{k1=-\infty}^{+\infty}\sum_{k2=-\infty}^{+\infty} F(u - \frac{k_1}{T}, v - \frac{k_2}{T})$. Since the support of the object is bounded in the spatial domain, the support of F is not bounded. However, it can be assumed that the frequency response is band-limited, i.e. there exists a cutoff frequency W such that $\|F(u, v)\| \approx 0$, when $|u| > W$ or $|w| > W$. We assume that $f(x, y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} F(u, v)\ e^{2\pi i(ux+vy)}du\ dv \approx \int_{-W}^{W}\int_{-W}^{W} F(u, v)\ e^{2\pi i(ux+vy)}du\ dv = \hat{f}(x, y)$, 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 that sampling period $T < \frac{1}{2W}$, i.e. $F_s(u, v) \longrightarrow F(u, v)$ for all for $|u| \leq W$ and $|v| \leq W$ as $T \searrow 0$.

Let $S_\theta : \Re \to \mathbb{C}$ be the Fourier transform of P_θ, where $P_\theta(t) = \int_{l_{\theta,t}} f(x, y)\ ds$ and $l_{\theta,t} = \{(x, y) : x\ cos\theta + y\ sin\theta = t\}$ is the path traveled by a beam of X-ray, when the projection angle is equal to θ. $S_\theta(w) = \int_{-\infty}^{+\infty} P_\theta(t)e^{-2\pi iwt}dt$. The two-dimensional polar Fourier transform $\tilde{F}(w, \theta)$ of f is defined in the usual manner, $\tilde{F}(w, \theta) = F(w\cos\theta, w\sin\theta)$. Then, the Fourier Slice Theorem implies that $S_\theta(w) = \tilde{F}(w, \theta)$.

However, in practice instead of the entire polar profile $P_\theta = \{P_\theta(t) | -\infty < t < +\infty\}$, parallel beam tomography only measures the total attenuation values of a finite set of X-ray beams, $\hat{P}_\theta = \{P_\theta(n\tau) |\ n \in \mathbb{Z},\ \lfloor \frac{-\sqrt{2}\ B}{\tau} \rfloor \leq n \leq \lceil \frac{\sqrt{2}\ B}{\tau} \rceil\}$, where τ is the sampling period of the tomography machine, i.e. distance between the adjacent x-ray detectors. Let

$$\hat{S}_\theta(w) = \tau \sum_{n=\lfloor \frac{-\sqrt{2}\ B}{\tau} \rfloor}^{\lceil \frac{\sqrt{2}\ B}{\tau} \rceil} P_\theta(n\tau)\ e^{-2\pi i w n \tau}. \tag{4.4}$$

Since $\hat{S}_\theta(w) = \sum_{k=-\infty}^{+\infty} S_\theta(w - \frac{k}{\tau})$, it is true that $\hat{S}_\theta(w) \longrightarrow S_\theta(w)$ for all $w \in \Re$ such that $|w| \leq W$ as $\tau \searrow 0$. It was also assumed that $F_s(u,v) \approx F(u,v)$ for $|u| \leq W$ and $|v| \leq W$. Then for $\tau > 0$ and $T < \frac{1}{2W}$ small, we have $\hat{S}_\theta(w) \approx S_\theta(w) = \tilde{F}(w,\theta) = F(w\cos\theta, w\sin\theta) \approx F_s(w\cos\theta, w\sin\theta)$ 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_θ for all practical purposes. Using a database of past images, for different type of CT images e.g. head, heart, each corresponding cutoff frequency value e.g. W_{head}, W_{heart}, can be estimated. Therefore, assuming that an estimate of cutoff value, W, is known for the type of the object f, choosing an appropriate sampling period T such that $T < \frac{1}{2W}$, actually determines the resolution of the digital image, f^d, to be reconstructed, i.e. $f^d \in \Re^{\ (n_d+1)\ \times\ (n_d+1)}$, $n_d = 2\lceil \frac{B}{T} \rceil$. In the actual process, tomography machine evaluates $\hat{P}_\theta = \{P_\theta(n\tau) |\ n \in \mathbb{Z},\ \lfloor \frac{-\sqrt{2}\ B}{\tau} \rfloor \leq n \leq \lceil \frac{\sqrt{2}\ B}{\tau} \rceil\}$ for a finite set of θ values. Let $\Omega_1 = \{\theta_1,\ \theta_2,\ ...,\ \theta_{N_1}\}$ be the finite set of projection angles. Once $\hat{P}_\theta$ is known for all $\theta \in \Omega_1$, $\hat{S}_\theta(w)$ 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 $w \in \Omega_2 = \{w_1,\ w_2,\ ....,\ w_{N_2}\}$ 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 $\{\hat{S}_\theta(w) \mid (\theta, w) \in \Omega\ \}$ 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 $e_{\theta,w} \in \Re$ 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.

$$T^2 \sum_{n_1=\lfloor \frac{-B}{T} \rfloor}^{\lceil \frac{B}{T} \rceil} \sum_{n_2=\lfloor \frac{-B}{T} \rfloor}^{\lceil \frac{B}{T} \rceil} f^d_{n_1,n_2}\ e^{-2\pi i(n_1\cos\theta + n_2\sin\theta)wT} = \hat{S}_\theta(w) + e_{\theta,w} \qquad \forall (\theta, w) \in \Omega \tag{4.5}$$

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

In order to be able to use the optimization algorithm [8] given in Section 5, Rx and $R^H y$ must be computed efficiently for all $x \in \Re^{n_d^2}$ and $y \in \Re^{|\Omega|}$. The efficiency of these matrix vector multiplications depends critically on the structure of matrix R. On the other hand, as long as the cardinality of Ω_1 stays same, elements of Ω_1 and Ω_2 can be chosen to satisfy our needs (Since the cardinality of Ω_1 is equal to the number of projections used in reconstruction and that number determines the radiation exposure of the patient and the time needed for tomography, we want the cardinality of the set to be equal to a predetermined value, which is much less than the number of projections used today).

Without loss of generality choose T and B such that $n_d = \frac{2B}{T}$ and

$$F_s(u,v) = T^2 \sum_{n_1=-\frac{n_d}{2}}^{\frac{n_d}{2}-1} \sum_{n_2=-\frac{n_d}{2}}^{\frac{n_d}{2}-1} f(n_1T, n_2T)\ e^{-2\pi i(un_1 + vn_2)T}, \tag{4.6}$$

$$= T^2 \sum_{i_1=0}^{n_d-1} \sum_{i_2=0}^{n_d-1} f^d_{i_1,i_2}\ e^{-2\pi i[u(i_1 - \frac{n_d}{2}) + v(i_2 - \frac{n_d}{2})]T}. \tag{4.7}$$

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

$$e_{(u,v)}+\hat{S}_{\theta(u,v)}(w(u,v))=F_s(u,v)=T^2\sum_{i_1=0}^{n_d-1}\sum_{i_2=0}^{n_d-1}f^d_{i_1,i_2}\ e^{-2\pi i[u(i_1-\frac{n_d}{2})+v(i_2-\frac{n_d}{2})]T}. \tag{4.8}$$

Thus, when a $(u,v)\in\Re^2$ is fixed, corresponding row in the R matrix formed according to Equation (4.8).

Moreover, $F_s(u,v)$ is periodic in both directions with a period $\frac{1}{T}=\frac{n_d}{2B}$. Hence, it is sufficient to collect Fourier observations of the object in the set $\{(u,v)|\ u\in[-\frac{n_d}{4B},\frac{n_d}{4B}]\ ,\ v\in[-\frac{n_d}{4B},\frac{n_d}{4B}]\}$. Then we define two sets, which together form Ω :

$$V=\{(\epsilon_x,\epsilon_y)|\ \epsilon_y=\frac{l}{4B},\ 0\leq l\leq n_d;\ \epsilon_x=\epsilon_y\frac{2m}{n_d},\ -\frac{n_d}{2}\leq m<\frac{n_d}{2}\}, \tag{4.9}$$

$$H=\{(\epsilon_x,\epsilon_y)|\ \epsilon_x=\frac{l}{4B},\ 0\leq l\leq n_d;\ \epsilon_y=\epsilon_x\frac{2m}{n_d},\ -\frac{n_d}{2}<m\leq\frac{n_d}{2}\}. \tag{4.10}$$

The definition of the sets U and V directly follows from pseudo-Polar grid selection in [10]. R matrix is fixed since Ω is fixed and its structure is determined by the sets U and V. Half of the rows of R is determined by set V and the other half is determined by H. Thus, $R=\begin{bmatrix}R_V\\R_H\end{bmatrix}$. For all $(\epsilon_x,\epsilon_y)\in(VUH)$, note that $(\epsilon_x,\epsilon_y)\in\{(u,v)|\ v+u\geq 0\}$. Let

$$\overline{V}=\{(\epsilon_x,\epsilon_y)|\ \epsilon_y=\frac{l}{4B},\ -n_d\leq l<n_d;\ \epsilon_x=\epsilon_y\frac{2m}{n_d},\ -\frac{n_d}{2}\leq m<\frac{n_d}{2}\}, \tag{4.11}$$

$$\overline{H}=\{(\epsilon_x,\epsilon_y)|\ \epsilon_x=\frac{l}{4B},\ -n_d\leq l<n_d;\ \epsilon_y=\epsilon_x\frac{2m}{n_d},\ -\frac{n_d}{2}<m\leq\frac{n_d}{2}\}. \tag{4.12}$$

Since f^d only takes real values, it can be argued that observing Fourier coefficients in the set $(\overline{V}U\overline{H})$ is equivalent to observing those only in (VUH). Therefore, we will use the smaller set (VUH) of same information content.

The points in the set V lie on n_d different rays starting from the origin. The angle between any ray and $(1,0)$ vector is an element of the following set: $\Omega_1^V=\{\tan^{-1}(n_d,\ 2m)|\ -n_d/2\leq m<n_d/2\}$. Moreover, the points in the set U also lie on n_d different rays starting from the origin and the angle between any ray and $(1,0)$ vector is an element of the set $\Omega_1^U\{\tan^{-1}(2m,\ n_d)|\ -n_d/2<m\leq n_d/2\}$. Thus the finite set of angles that will be used for the CT scan is the union of both sets, i.e. $\Omega_1=\Omega_1^V\cup\Omega_1^U$, where $|\Omega_1|=2n_d$. Below it will be shown that when Ω is chosen as suggested above, then all necessary operations can be done efficiently.

Given $x\in\Re^{(n_d\times n_d)}$, Rx can be computed efficiently. Below, it will be shown that R_Vx can be done fast and similar arguments are also true for R_Hx. Computational complexity of FFT and Fractional Fourier Transform [9] on an n dimensional vector is $O(5nlogn)$ and $O(30nlogn)$, respectively.

For set V, let the inner product of x with the row of R_V corresponding to $\epsilon_x=\frac{2lm}{4Bn_d}$ and $\epsilon_y=\frac{l}{4B}$ be $X_{m,l}$. Then,

$$X_{m,l}=T^2\sum_{i_1=0}^{n_d-1}\sum_{i_2=0}^{n_d-1}x_{i_1,i_2}\ e^{-2\pi i[\frac{2lm}{4Bn_d}(i_1-\frac{n_d}{2})+\frac{l}{4B}(i_2-\frac{n_d}{2})]\frac{2B}{n_d}}, \tag{4.13}$$

$$=T^2\sum_{i_1=0}^{n_d-1}\sum_{i_2=0}^{n_d-1}x_{i_1,i_2}\ e^{-i[\frac{2\pi lm}{n_d^2}i_1+\frac{\pi l}{n_d}i_2]}\ e^{i\pi l(\frac{m}{n_d}+\frac{1}{2})}. \tag{4.14}$$

Define $\hat{x}$,

$$\hat{x}_{i_1,l}=\sum_{i_2=0}^{n_d-1}x_{i_1,i_2}\ e^{-2\pi i\frac{i_2l}{2n_d}}=\sum_{i_2=0}^{2n_d-1}x^o_{i_1,i_2}\ e^{-2\pi i\frac{i_2l}{2n_d}}, \tag{4.15}$$

where $x^o_{i_1,i_2} = \begin{cases} x_{i_1,i_2} & 0 \le i_2 < n_d \\ 0 & n_d \le i_2 < 2n_d \end{cases}$.

By zero padding x, for a given i_1 ($1 \le i_1 < n_d$) we can use FFT for (4.15) to compute $\hat{x}_{i_1,l}$ for $0 \le l \le n_d$, efficiently. Complexity bound of this computation is $O(5(2n_d)\log(2n_d))$ for each $0 \le i_1 \le n_d - 1$. 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 $\hat{x}$:

$$X_{m,l} = T^2 \; e^{i\pi l(\frac{m}{n_d}+\frac{1}{2})} \sum_{i_1=0}^{n_d-1} \hat{x}_{i_1,l} \; e^{-2\pi i \frac{mi_1}{n_d}(\frac{l}{n_d})}$$

Given l ($0 \le l \le n_d$), in order to calculate $X_{m,l}$ efficiently for $-\frac{n_d}{2} \le m < \frac{n_d}{2}$, Fractional Fourier Transform is used. Complexity bound of this computation is $O(30n_d \log(n_d))$ for each $0 \le l \le n_d$. 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, Rx 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 $y_V \in \mathbb{C}^{\frac{|\Omega|}{2}}$ and $y_H \in \mathbb{C}^{\frac{|\Omega|}{2}}$, 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 done fast 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 $\epsilon_x = \frac{2lm}{4Bn_d}$ and $\epsilon_y = \frac{l}{4B}$ be Y_{i_1,i_2}. Then,

$$Y_{i_1,i_2} = T^2 \sum_{m=-\frac{n_d}{2}}^{\frac{n_d}{2}-1} \sum_{l=0}^{n_d} y_{m,l} \; e^{2\pi i[\frac{lm}{Bn_d^2}(i_1-\frac{n_d}{2})+\frac{l}{2n_d}(i_2-\frac{n_d}{2})]}, \tag{4.16}$$

$$= T^2 \sum_{l=0}^{n_d} e^{2\pi i \; \frac{l}{2n_d}(i_2-\frac{n_d}{2})} \sum_{m=-\frac{n_d}{2}}^{\frac{n_d}{2}-1} y_{m,l} \; e^{2\pi i \frac{lm}{Bn_d^2}(i_1-\frac{n_d}{2})}. \tag{4.17}$$

Let

$$\hat{y}_{i_1,l} = \sum_{m=-\frac{n_d}{2}}^{\frac{n_d}{2}-1} y_{m,l} \; e^{-2\pi i \frac{m}{n_d}(i_1-\frac{n_d}{2})(-\frac{l}{n_d})}. \tag{4.18}$$

Given l ($0 \le l \le n_d$), we can use Fractional Fourier Transform for (4.18) to compute $\hat{y}_{i_1,l}$ for $0 \le i_1 < n_d$, efficiently. The complexity bound for this operation is $O(30n_d \log(n_d))$ for each $0 \le l \le n_d$. 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 $\hat{y}$:

$$\begin{aligned} Y_{i_1,i_2} &= T^2 \sum_{l=0}^{n_d} \hat{y}_{i_1,l} \; e^{2\pi i \; \frac{l}{2n_d}(i_2-\frac{n_d}{2})} \\ &= T^2 \sum_{l=0}^{2n_d-1} (\hat{y}^o_{i_1,l} \; e^{-\pi i \frac{l}{2}}) e^{2\pi i \; \frac{li_2}{2n_d}} \end{aligned} \tag{4.19}$$

where $\hat{y}^o_{i_1,l} = \begin{cases} \hat{y}_{i_1,l} & 0 \le l \le n_d \\ 0 & n_d < l < 2n_d \end{cases}$.

By zero padding $\hat{y}$, given i_1 ($0 \le i_1 < n_d$), we can use inverse FFT for (4.19) to compute Y_{i_1,i_2} for $0 \le i_2 < n_d$, efficiently. This can be done in $O(5(2n_d)\log(2n_d))$ time for each $0 \le i_1 < n_d$. 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.

5. Reconstruction Algorithm. There are several methods available for solving the CS problem defined in (3.4) [5, 6]. We choose to use the method developed by Ma et. al in [8]. They formulated the reconstruction problem using partial Fourier matrix R and wavelet transform Φ, ($A = R\Phi^*$). Total variation $TV(f)$ of the signal 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 ∇_1 and ∇_2 denote the forward finite difference operators on the first and second coordinates, respectively. They also added a TV term to the objective function in problem (3.4), and proposed an efficient algorithm for solving:

$$\min_x \alpha TV(\Phi^* x) + \beta \|x\|_1 + \frac{1}{2}\|Ax - b\|_2^2, \tag{5.1}$$

where α and β are positive parameters.

In this paper, we will solve Problem 5.1 using the algorithm described below. R is the Partial Pseudo Polar Fourier matrix formed using equation 4.8 for all $(u, v) \in V \bigcup U$ (each $(u, v) \in V \bigcup U$ corresponds to a row in R). b is the measurement vector and for each $(u, v) \in V \bigcup U$ corresponding element of b is equal to $\hat{S}_{\theta(u,v)}(w(u, v))$ as in equation 4.8. Moreover, Φ is the single level discrete 1-D wavelet transform.

Let $f^d \in \Re^{n_1 \times n_2}$ represent 2D image of $n_1 \times n_2$ pixels, $L = (\nabla_1, \nabla_2) : \Re^{n_1 \times n_2} \rightarrow \Re^{n_1 \times n_2} \times \Re^{n_1 \times n_2}$ denote the discrete finite difference operators and its suboperator $L_{ij}(f^d) = (\nabla_1 f^d_{ij}, \nabla_2 f^d_{ij})$, $f(.) = \|.\|_2 : \Re^2 \rightarrow \Re$ such that $TV(f^d) = \sum_{ij} f(L_{ij}(f^d))$, $g(.) = \|.\|_1$, $h(.) = \frac{1}{2}\|A. - b\|_2$ and $\Psi = \Phi^*$, which is the adjoint operator of Φ, for any orthonormal transform Φ. With the new notation, (5.1) can be written as

$$\min_x \alpha \sum_{ij} f(L_{ij}(\Psi x)) + \ \beta g(x) + h(x). \tag{5.2}$$

In [8], Ma et. al suggest using the following fixed point iterations to solve (5.2):

$$s^{(k+1)} = x^{(k)} - \tau_1(\alpha\Phi \sum_{ij} L^*_{ij}(y^{(k)}) + \nabla_x h(x^{(k)})) \tag{5.3}$$

$$t^{(k+1)}_{ij} = y^{(k)} + \tau_2 L_{ij}(\Psi x^{(k)}) \tag{5.4}$$

$$x^{(k+1)} = sign(s^{(k)}) \max\{0, |s| - \tau_1\beta\} \tag{5.5}$$

$$y^{(k+1)}_{ij} = min\{\frac{1}{\tau_2}, \|t^{(k)}_{ij}\|_2\} \frac{t^{(k)}_{ij}}{\|t^{(k)}_{ij}\|_2} \tag{5.6}$$

for $k = 0, 1, ...$, starting from a set of initial points $x^{(0)}$, $y^{(0)}$, $s^{(0)}$, $t^{(0)}$.

See [8] for a justification of the 4-step iterative method and its convergence proofs.

6. Results. In order to test the quality of the solutions obtained a variant of Shepp-Logan head phantom is used. The variant of Shepp-Logan head phantom consists of ten ellipses:

In Table 6.1, 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 α 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:

$$f(x, y) = \begin{cases} I & \frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1, \\ 0 & \frac{x^2}{a^2} + \frac{y^2}{b^2} > 1, \end{cases} \tag{6.1}$$

then it is easy to show that projections of the target object can be represented in closed form:

$$P_\theta(t) = \begin{cases} 2I\frac{ab}{c(\theta)}\sqrt{c^2(\theta) - t^2} & |t| \leq c(\theta), \\ 0 & |t| > c(\theta), \end{cases} \tag{6.2}$$

Ellipse#	**I**	**a**	**b**	x_0	y_0	α
1	1	.69	.92	0	0	0
2	-.8	.6624	.8740	0	-.0184	0
3	-.2	.1100	.3100	..22	0	-18
4	-.2	.1600	.4100	-.22	0	18
5	.1	.2100	.2500	0	.35	0
6	.1	.0460	.0460	0	.1	0
7	.1	.0460	.0460	0	-.1	0
8	.1	.0460	.0230	-.08	-.605	0
9	.1	.0230	.0230	0	-.606	0
10	.1	.0230	.0460	..06	-.605	0

TABLE 6.1
A Variant of Shepp-Logan Head Phantom Parameters

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 α counterclockwise, then the new projections, $P_\theta^{'}(t)$ are obtained using the projections of the centered and unrotated ellipse, $P_\theta(t)$, as follows:

$$P_\theta^{'}(t) = P_{\theta-\alpha}(t - s\cos(\gamma - \theta)), \tag{6.3}$$

where $s = \sqrt{x_0^2 + y_0^2}$ and $\gamma = tan^{-1}(\frac{y_0}{x_0})$. Thus using equations (6.2) and (6.3), projection values for Shepp-Logan head phantom can be calculated.

Since $P_\theta(t)$ is known, $\hat{S}_\theta(w)$ can be calculated using equation (4.4) for any $|w| \leq W$. For the Shepp-Logan head phantom cutoff frequency, W, can be chosen to be 256 i.e. $\|F(u,v)\| \approx 0$, when $|u| > 256$ or $|w| > 256$. For this reason, size of the reconstructed image is chosen to be 512×512, i.e. $n_d = 512$. According to the setup described in section 4, the number of angles used should be $2n_d = 1024$, i.e. $|\Omega_1| = 2n_d = 1024$. However, in order to test the reconstruction abilities of the proposed method for the case of undersampled measurements, one out of Δ angles from Ω_1 is used. During the numerical studies we have used 4 different values of Δ: 8, 16, 32, 64 which is equivalent to using 128, 64, 32 and 16 projection angles for the reconstruction. For a given Δ the set of projection angles used is Ω_1, which is the union of Ω_1^V and Ω_1^U, where

$$\Omega_1^V = \{\tan^{-1}(n_d,\ 2m)|\ m = -n_d/2 + k\Delta,\ 0 \leq k < \frac{n_d}{\Delta}\}, \tag{6.4}$$

$$\Omega_1^U = \{\tan^{-1}(2m,\ n_d)|\ m = n_d/2 - k\Delta,\ 0 \leq k < \frac{n_d}{\Delta}\}, \tag{6.5}$$

$$\Omega_1 = \Omega_1^V \cup \Omega_1^U. \tag{6.6}$$

Figure 6.2 is a variant of the Shepp-Logan image, of which parameters are given in Table 6.1. Figure 6.3, Figure 6.4, Figure 6.5 and Figure 6.6 are reconstructed using CT data coming from equations (6.2) and (6.3). Figure 6.3 recovers all the important characteristic features of the Shepp-Logan Head phantom, i.e. Figure 6.2. Figure 6.4 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 6.5 and Figure 6.6, are not as sharp as previous reconstructions. Moreover, in order to compare our method with the commonly used Filtered Back Projection (FBP) [1], we have also included figures reconstructed by FBP using undersampled projection data. Figure 6.7 and Figure 6.8 are FBP reconstructions using 64 and 32 projection angles, respectively. According to the experimental results, because all the images reconstructed by our technique is free of aliasing artifacts and have all the desired features of the Shepp-Logan phantom in contrast to the reconstructions using FBP, Figure 6.7 and Figure 6.8, our technique is preferable to FBP when undersampled projection data is used. In Table 6.2, relative reconstruction error values are given for different number of projection angles used. Relative reconstruction error is defined by $\|x^{sol} - x^*\|_2/\|x^*\|_2$, where x^{sol} is the image produced by our method and x^* is the digitized Shepp-Logan Phantom image.

On the other hand to test the robustness of our modeling technique to the noisy CT measurement data, two

Projection # :	128	64	32	16
$\|x^{sol} - x^*\|_2/\|x^*\|_2$:	0.1113	0.1214	0.1453	0.2296

TABLE 6.2
Relative reconstruction error results for CT data without measurement error

types of noisy CT data is generated. Now instead of $P_\theta^{'}(t)$ in equation 6.3, $P_\theta^{noisy}(t)$ is used to generate the observations, where $P_\theta^{noisy}(t) = P_\theta^{'}(t) + \eta_{\theta,t}$, and $\eta_{\theta,t} \sim \sigma(\theta,t)\ Normal(0,1)$. We have assumed that the measurement noise set, $\{\eta_{\theta,t} : \theta \in \Omega_1\ \forall\ t\}$, consists of independent normal random variables with variance $(\sigma(\theta,t))^2$.

In the first case, we have assumed a constant variance i.e. $\sigma(\theta,t) = \sigma\ \forall \theta \in \Omega_1$ and $\forall\ t$. Define σ_ξ,

$$\sigma_\xi = \frac{\xi}{|\Omega_1|(2\lceil \frac{\sqrt{2}\ B}{\tau} \rceil + 1)} \sum_{\theta \in \Omega_1} \sum_{n=\lfloor \frac{-\sqrt{2}\ B}{\tau} \rfloor}^{\lceil \frac{\sqrt{2}\ B}{\tau} \rceil} P_\theta(n\tau), \tag{6.7}$$

where τ is the sampling period of the tomography machine, i.e. distance between the adjacent x-ray detectors. In equation (6.7), if $\xi = 1$, then σ_ξ is equal to the average of all the x-ray attenuations are measured. Hence, given ξ, for all $\theta \in \Omega_1$ and for all t, $\eta_{\theta,t}$ has a constant standard deviation σ_ξ as given in Equation (6.7), which can be interpreted as $100\xi\%$ of the average x-ray attenuation measured. Figure 6.9 contains the reconstructions of the phantom for different ξ values and Table 6.3 contains relative reconstruction errors for changing ξ values. Figure 6.1(a) plots relative reconstruction error versus $\log \xi$.

Figure # :	6.9(a)	6.9(b)	6.9(c)	6.9(d)	6.9(e)
ξ :	1×10^{-3}	5×10^{-3}	1×10^{-2}	5×10^{-2}	1×10^{-1}
$\|x^{sol} - x^*\|_2/\|x^*\|_2$:	0.1129	0.1137	0.1135	0.1369	0.2406

TABLE 6.3
Relative reconstruction error results for noisy CT data, $\sigma(\theta,t) = \sigma_\xi$

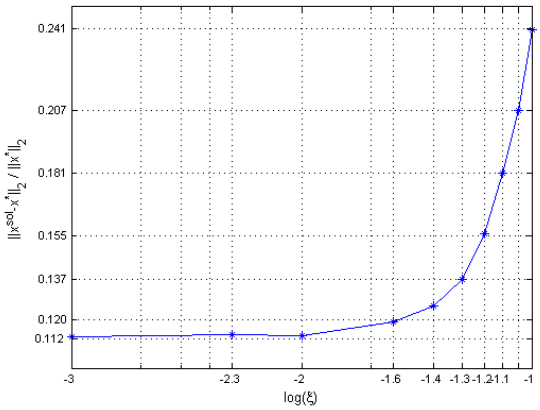
In the second case, the standard deviation of the measurement error is assumed to be proportional to the x-ray attenuation:

$$\sigma(\theta,t) = \xi\ P_\theta(t). \tag{6.8}$$

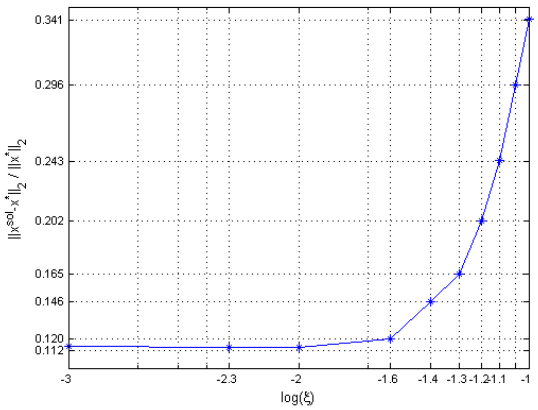
Given ξ, $\eta_{\theta,t}$ has a standard deviation $\sigma(\theta,t)$ as given in equation (6.8), which can be interpreted as $100\xi\%$ of the x-ray attenuation measured for θ and t. Figure 6.10 contains the reconstructions of the phantom for different ξ values and Table 6.4 contains relative reconstruction errors for changing ξ values. Figure 6.1(b) plots relative reconstruction error versus $\log \xi$. Clearly the proposed solution is robust to significant levels of noise.

Figure # :	6.10(a)	6.10(b)	6.10(c)	6.10(d)	6.10(e)
ξ :	1×10^{-3}	5×10^{-3}	1×10^{-2}	5×10^{-2}	1×10^{-1}
$\|x^{sol} - x^*\|_2/\|x^*\|_2$:	0.1151	0.1143	0.1142	0.1650	0.3412

TABLE 6.4
Relative reconstruction error results for noisy CT data, $\sigma(\theta,t) = \xi P_\theta(t)$



(a) $\sigma(\theta, t) = \sigma_\xi$



(b) $\sigma(\theta, t) = \xi P_\theta(t)$

FIG. 6.1. *Relative reconstruction error versus* $\log \xi$

FIG. 6.2. *Shepp-Logan Head Phantom*

FIG. 6.3. *Shepp-Logan Head Phantom reconstructed from CT data without noise using 128 projection angles*

FIG. 6.4. *Shepp-Logan Head Phantom reconstructed from CT data without noise using 64 projection angles*

FIG. 6.5. *Shepp-Logan Head Phantom reconstructed from CT data without noise using 32 projection angles*

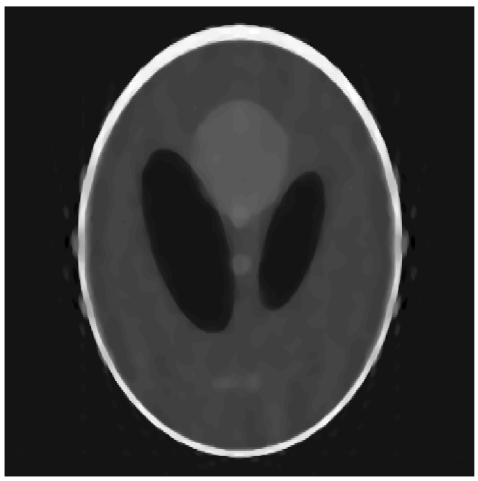


FIG. 6.6. *Shepp-Logan Head Phantom reconstructed from CT data without noise using 16 projection angles*

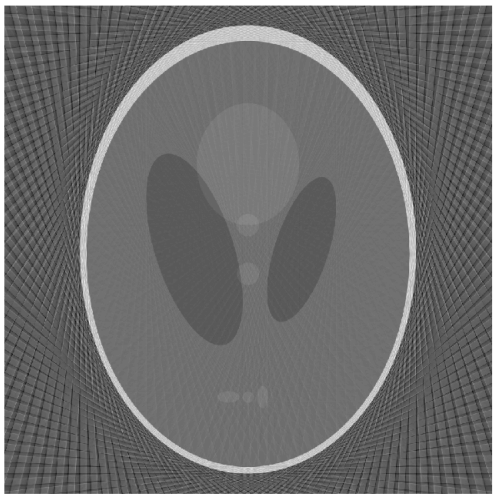


FIG. 6.7. *Shepp-Logan Head Phantom Filtered Back Projection reconstruction from CT data without noise using 64 projection angles*

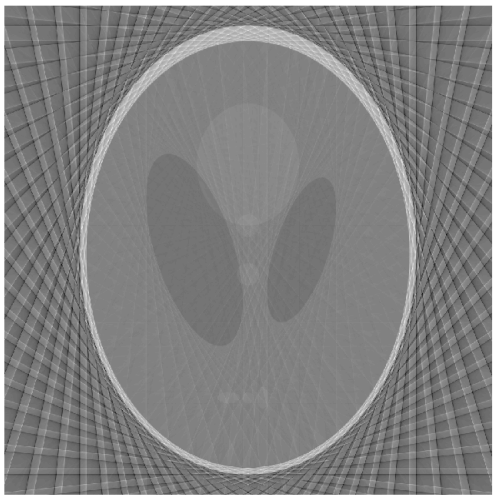
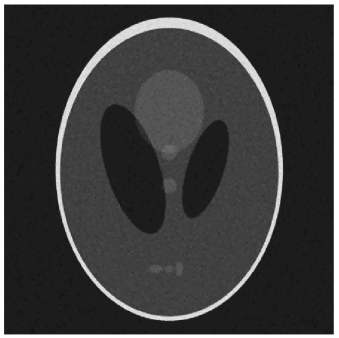


FIG. 6.8. *Shepp-Logan Head Phantom Filtered Back Projection reconstruction from CT data without noise using 32 projection angles*

(a) $\xi = 1 \times 10^{-3}$

(b) $\xi = 5 \times 10^{-3}$

(c) $\xi = 1 \times 10^{-2}$



(d) $\xi = 5 \times 10^{-2}$

(e) $\xi = 1 \times 10^{-1}$

FIG. 6.9. *Noisy CT data reconstruction using 128 projection angles,* $\sigma(\theta, t) = \sigma_\xi$

(a) $\xi = 1 \times 10^{-3}$

(b) $\xi = 5 \times 10^{-3}$

(c) $\xi = 1 \times 10^{-2}$

(d) $\xi = 5 \times 10^{-2}$

(e) $\xi = 1 \times 10^{-1}$

FIG. 6.10. *Noisy CT data reconstruction using 128 projection angles,* $\sigma(\theta, t) = \xi P_\theta(t)$

7. Conclusion. Computed Tomography is one of the most important and widely used diagnostic tools available to modern medicine. Compressed Sensing on the other hand, is one of the most exciting advances in computational science. In this paper our effort has been to integrate the two to develop an elegant framework that can spread the usefulness of CT to a larger section of the population and for even more diagnostic procedures through a significant reduction in the number of projections necessary to reconstruct the imaged volume and hence exposing the individual to a proportionally lower amount of radiation.

Compressed sensing has traditionally found more application in the magnetic resonance imaging (MRI) than in computed tomography imaging. The reason being that applying the framework to CT is much much challenging. In this paper we bridge the gap by formulating the CT reconstruction problem from limited projection data as one of compressed sensing via the use of partial pseudo-polar Fourier transform.

In order to test our formulation, we use a variant of the Shepp-Logan phantom where we can simulate reconstruction using various subsets of the projection data both for the traditional Filtered Back-Projection method as well as the proposed method. The effectiveness of the proposed method is evident. As demonstrated by the results (visual and the tables), the quality of the results using the proposed method is very high even when one uses significantly lower number of projection angles (64 instead of 512 used originally). The FBP method under such circumstances produces strong and visible streaking. Also, we note that the proposed method is computationally efficient as evidenced by the computational complexity calculations.

However, while we are excited by the opportunities this method presents, there are several avenues that we would like to explore to improve the technique. Currently, the formulation works for fan-bean projections, however a large number of CT instruments use cone-beam instead. The formulation needs adaptation to make it work under that environment. Also, the basic optimization framework used for compressed sensing continues to evolve, and hence there is room for improvement both, in terms of the accuracy and speed of reconstruction.

Appendix A. Fractional Fourier Transform. Definition: Given a signal, $x \in \mathbb{C}^N$ and arbitrary α, Fractional Fourier Transform of x is given by

$$X_k = \sum_{n=0}^{N-1} x_n \; e^{-2\pi i \frac{kn}{N}\alpha} \qquad \text{for } 0 \le k < N. \tag{A.1}$$

Evaluation of Fractional Fourier Transform on $x \in \mathbb{C}^N$ can be done in $O(30N\log(N))$ operations [9]. Since $2kn = k^2 + n^2 - (k-n)^2$,

$$X_k = \sum_{n=0}^{N-1} x_n \; e^{-2\pi i \frac{kn}{N}\alpha} \tag{A.2}$$

$$= \sum_{n=0}^{N-1} x_n \; e^{-\pi i \frac{[k^2+n^2-(k-n)^2]}{N}\alpha} \tag{A.3}$$

$$= e^{-\pi i \frac{k^2}{N}\alpha} \sum_{n=0}^{N-1} x_n \; e^{-\pi i \frac{n^2}{N}\alpha} \; e^{\pi i \frac{(k-n)^2}{N}\alpha} \tag{A.4}$$

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

$$X_k = z_k^* \sum_{n=0}^{N-1} y_n \; z_{k-n} \tag{A.5}$$

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 $\overline{x}$. Second, z is extended to $2N$ as follows to satisfy periodicity requirement for FFT:

$$\overline{z}_n = \begin{cases} e^{\pi i \frac{n^2}{N}\alpha}, & 0 \le n < N; \\ e^{\pi i \frac{(2N-n)^2}{N}\alpha}, & N \le n < 2N. \end{cases}$$

Now the above convolution can be efficiently computed using fast Fourier transform by computing the FFT of $\overline{x}$ and $\overline{z}$, 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. Oversampling. Let $Q \in \mathbb{Z}_+$ be an even number that determines the over sampling ratio. Then we define two sets, which together form Ω :

$$V = \{(\epsilon_x, \epsilon_y)|\ \epsilon_y = \frac{l}{4BQ},\ 0 \le l \le Qn_d;\ \epsilon_x = \epsilon_y \frac{2m}{n_d},\ -\frac{n_d}{2} \le m < \frac{n_d}{2}\} \tag{B.1}$$

$$H = \{(\epsilon_x, \epsilon_y)|\ \epsilon_x = \frac{l}{4BQ},\ 0 \le l \le Qn_d;\ \epsilon_y = \epsilon_x \frac{2m}{n_d},\ -\frac{n_d}{2} < m \le \frac{n_d}{2}\} \tag{B.2}$$

The definition of the sets U and V directly follows from pseudo-Polar grid selection in [10]. R matrix is fixed since Ω is fixed and its structure is determined by the sets U and V. Half of the rows of R is determined by set V and the other half is determined by H. Thus, $R = \begin{bmatrix} R_V \\ R_H \end{bmatrix}$. For all $(\epsilon_x, \epsilon_y) \in (VUH)$, note that $(\epsilon_x, \epsilon_y) \in \{(u, v)|\ v + u \ge 0\}$. Let

$$\overline{V} = \{(\epsilon_x, \epsilon_y)|\ \epsilon_y = \frac{l}{4BQ},\ -Qn_d \le l \le Qn_d;\ \epsilon_x = \epsilon_y \frac{2m}{n_d},\ -\frac{n_d}{2} \le m < \frac{n_d}{2}\} \tag{B.3}$$

$$\overline{H} = \{(\epsilon_x, \epsilon_y)|\ \epsilon_x = \frac{l}{4BQ},\ -Qn_d \le l \le Qn_d;\ \epsilon_y = \epsilon_x \frac{2m}{n_d},\ -\frac{n_d}{2} < m \le \frac{n_d}{2}\} \tag{B.4}$$

It can be argued that observing Fourier coefficients in the set $(\overline{V}U\overline{H})$ is equivalent to observing those only in (VUH) since f^d takes real values. Therefore, we will use the smaller set (VUH) of same information content.

Given $x \in \Re^{(n_d \times n_d)}$, Rx can be computed efficiently. Below, it will be shown that $R_V x$ can be done fast and similar arguments are also true for $R_H x$. Computational complexity of FFT and fractional Fourier transform [9] on an n dimensional vector is $O(5nlogn)$ and $O(30nlogn)$, respectively.

For set V, given $\epsilon_x = \frac{2lm}{4BQn_d}$ and $\epsilon_y = \frac{l}{4BQ}$, $X_{m,l}$ is defined as follows:

$$X_{m,l} = T^2 \sum_{i_1=0}^{n_d-1} \sum_{i_2=0}^{n_d-1} x_{i_1,i_2}\ e^{-2\pi i[\frac{2lm}{4BQn_d}(i_1 - \frac{n_d}{2}) + \frac{l}{4BQ}(i_2 - \frac{n_d}{2})]\frac{2B}{n_d}}, \tag{B.5}$$

$$= T^2 \sum_{i_1=0}^{n_d-1} \sum_{i_2=0}^{n_d-1} x_{i_1,i_2}\ e^{-i[\frac{2\pi lm}{Qn_d^2} i_1 + \frac{\pi l}{Qn_d} i_2]}\ e^{\frac{i\pi l}{Q}(\frac{m}{n_d} + \frac{1}{2})}, \tag{B.6}$$

$$\hat{x}_{i_1,l} = \sum_{i_2=0}^{n_d-1} x_{i_1,i_2}\ e^{-2\pi i \frac{i_2 l}{2Qn_d}},$$

$$= \sum_{i_2=0}^{2Qn_d-1} x^o_{i_1,i_2}\ e^{-2\pi i \frac{i_2 l}{2Qn_d}}, \tag{B.7}$$

where $x^o_{i_1,i_2} = \begin{cases} x_{i_1,i_2} & 0 \le i_2 < n_d \\ 0 & n_d \le i_2 < 2Qn_d \end{cases}$.

By zero padding x, we can use FFT for (B.7) to calculate $\hat{x}_{i_1,l}$ for $0 \leq l \leq Qn_d$, efficiently. Complexity bound of this computation is $O(10Qn_d^2 \log n_d)$.

$$X_{m,l} = T^2 \; e^{\frac{i\pi l}{Q}(\frac{m}{n_d}+\frac{1}{2})} \sum_{i_1=0}^{n_d-1} \hat{x}_{i_1,l} \; e^{-2\pi i \frac{mi_1}{n_d}(\frac{l}{Qn_d})} \tag{B.8}$$

In order to calculate $X_{m,l}$ efficiently for $-\frac{n_d}{2} \leq m < \frac{n_d}{2}$, Fractional Fourier Transform is used. Complexity bound of this computation is $O(30Qn_d^2 \log n_d)$. For the set H same arguments above still hold. Therefore, Rx can be computed using arithmetic operations in the order of $80Qn_d^2 \log n_d$.

Now given $y = \begin{bmatrix} y_V \\ y_H \end{bmatrix}$, where $y_V \in \mathbb{C}^{\frac{|\Omega|}{2}}$ and $y_H \in \mathbb{C}^{\frac{|\Omega|}{2}}$, 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 done fast and similar arguments are also true for $R_H^* y_H$. For set V, given $\epsilon_x = \frac{2lm}{4BQn_d}$ and $\epsilon_y = \frac{l}{4BQ}$, Y_{i_1,i_2} is defined as follows:

$$Y_{i_1,i_2} = T^2 \sum_{m=-\frac{n_d}{2}}^{\frac{n_d}{2}-1} \sum_{l=0}^{Qn_d} y_{m,l} \; e^{2\pi i[\frac{lm}{BQn_d^2}(i_1-\frac{n_d}{2})+\frac{l}{2Qn_d}(i_2-\frac{n_d}{2})]} \tag{B.9}$$

$$= T^2 \sum_{l=0}^{Qn_d} e^{2\pi i \; \frac{l}{2Qn_d}(i_2-\frac{n_d}{2})} \sum_{m=-\frac{n_d}{2}}^{\frac{n_d}{2}-1} y_{m,l} \; e^{2\pi i \frac{lm}{BQn_d^2}(i_1-\frac{n_d}{2})} \tag{B.10}$$

$$\tag{B.11}$$

Inside summation can be computed using fractional Fourier transform for all $0 \leq l \leq Qn_d$.

$$\hat{y}_{i_1,l} = \sum_{m=-\frac{n_d}{2}}^{\frac{n_d}{2}-1} y_{m,l} \; e^{-2\pi i \frac{m}{n_d}(i_1-\frac{n_d}{2})(-\frac{l}{Qn_d})} \tag{B.12}$$

$$\tag{B.13}$$

The complexity bound for this operation is $O(30Qn_d^2 \log n_d)$.

Then we have

$$Y_{i_1,i_2} = T^2 \sum_{l=0}^{Qn_d} \hat{y}_{i_1,l} e^{2\pi i \; \frac{l}{2Qn_d}(i_2-\frac{n_d}{2})}$$

$$= T^2 \sum_{l=0}^{2Qn_d-1} (\hat{y}^o_{i_1,l} \; e^{-\pi i \frac{l}{2Q}}) e^{2\pi i \; \frac{li_2}{2Qn_d}} \tag{B.14}$$

$$\tag{B.15}$$

where $\hat{y}^o_{i_1,l} = \begin{cases} \hat{y}_{i_1,l} & 0 \leq l \leq Qn_d \\ 0 & Qn_d < l < 2Qn_d \end{cases}$.

By zero padding $\hat{y}$, we can use inverse FFT for (B.14) to calculate Y_{i_1,i_2} for $0 \leq i_2 \leq n_d$, efficiently. This can be done in $O(10Qn_d^2 \log n_d)$ time.

REFERENCES

[1] A. C. Kak and Malcolm Slaney, *Principles of Computerized Tomographic Imaging*, IEEE Press, 1988.
[2] A. C. Kak, "Computerized tomography with x-ray emmision and ultrasound sources", Proc. IEEE, vol. 67, pp. 1245-1272, 1979.
[3] A. Rosenfeld and A. C. Kak, *Digital Picture Processing, 2nd ed.*, New York, NY, Academic Press, 1982.
[4] The Mathworks, Inc., http://www.mathworks.com/products/matlab, 2005.
[5] E. Cands and T. Tao, "Near optimal signal recovery from random projections: universal encoding strategies" IEEE Transactions on Information Theory, vol. 52(1):54065425, 2006.
[6] D. Donoho, "Compressed sensing", IEEE Transactions on Information Theory, 52:12891306, 2006.
[7] E. J. Cands, "Compressive sampling", Proceedings of the International Congress of Mathematicians, Madrid, Spain, 2006.
[8] Shiqian Ma, Wotao Yin, Yin Zhang and Amit Chakraborty, "An Efficient Algorithm for Compressed MR Imaging using Total Variation and Wavelets", IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2008, 1-8, 2008.
[9] D. H. Bailey and P. N. Swarztrauber, "The fractional Fourier transform and applications", SIAM Review, vol. 33, no. 3 (Sept. 1991), p. 389 - 404
[10] A. Averbuch, R. Coifman, D.L. Donoho, M. Elad, and M. Israeli, "Fast and Accurate Polar Fourier Transform", the Applied and Computational Harmonic Analysis journal, Vol. 21(2), pp. 145-167, September 2006.
[11] K. T. Block, M. Uecker, and J. Frahm. Undersampled radial MRI with multiple coils. iterative image reconstruction using a total variation constraint. *Magnetic Resonance in Medicine*, 57:1086–1098, 2007.
[12] E. Candès and M. Wakin. An introduction to compressive sampling. *IEEE Signal Processing Magazine*, March:21–30, 2008.
[13] D. Donoho and M. Elad. Optimally sparse representation in general (nonorthogonal) dictionaries vis ℓ_1 minimization. *Proceedings of the National Academy of Sciences*, 100:2197–2202, 2003.
[14] D. Donoho and X. Huo. Uncertainty principles and ideal atomic decompositions. *IEEE Transactions on Information Theory*, 47:2845–2862, 2001.
[15] J. Fuchs. On sparse representations in arbitrary redundant bases. *IEEE Transactions on Information Theory*, 50.
[16] B. K. Natarajan. Sparse approximate solutions to linear systems. *SIAM Journal on Computing*, 24:227–234, 1995.
[17] R.G. Evens and F.A. Mettler. National CT use and radiation exposure: United States 1983 *American Journal of Radiology*, 144:1077-1081, 1985.
[18] D.J. Brenner and E.J. Hall. Computed Tomography - An Increasing Source of Radiation Exposure *New England Journal of Medicine*, 357:2277:84, 2006
[19] P.C. Shrimpton and S. Edyvean CT scanner dosiometry *British Journal of Radiology*, 71:1-3, 1998.
[20] H. Yu, and G. Wang. Compressed Sensisng based interior tomography *Physics in Medicine and Biology*, 54:2791-2805, 2009.
[21] M. Rudelson and R. Vershynin. Geometric approach to error correcting codes and reconstruction of signals. *International Mathematical Research Notices*, 64:4019–4041, 2005.
[22] E. Candès, J. Romberg, and T. Tao. Stable signal recovery from incomplete and inaccurate information. *Communications on Pure and Applied Mathematics*, 2005(59):1207–1233, 2005.
[23] Jorge Nocedal and Steven Wright. *Numerical Optimization*. Springer, New York, 2nd edition, 2006.
[24] E. Hale, W. Yin, and Y. Zhang. FPC: A fixed-point continuation method for ℓ_1-regularization, `http://www.caam.rice.edu/~optimization/`. 2007.