Non-Convex Compressed Sensing CT Reconstruction Based on Tensor Discrete Fourier Slice Theorem*

Il Yong Chun[1], Ben Adcock[2], and Thomas M. Talavage[3]

***Abstract*— X-ray computed tomography (CT) scanners provide clinical value through high resolution and fast imaging. However, achievement of higher signal-to-noise ratios generally requires emission of more X-rays, resulting in greater dose delivered to the body of the patient. This is of concern, as higher dose leads to greater risk of cancer, particularly for those exposed at a younger age. Therefore, it is desirable to achieve comparable scan quality while limiting X-ray dose. One means to achieve this compound goal is the use of compressed sensing (CS). A novel framework is presented to combine CS theory with X-ray CT. According to the tensor discrete Fourier slice theorem, the 1-D DFT of discrete Radon transform data is exactly mapped on a Cartesian 2-D DFT grid. The nonuniform random density sampling of Fourier coefficients is made feasible by uniformly sampling projection angles at random. Application of the non-convex CS model further reduces the sufficient number of measurements by enhancing sparsity. The numerical results show that, with limited projection data, the non-convex CS model significantly improves reconstruction performance over the convex model.**

I. INTRODUCTION

We first consider the discrete Radon transform (DRT) of the discrete image $\mathbf{x} = x(m,n) \in \mathbb{R}^{N^2}$ on the Cartesian grid:

$$p_\theta(r) = \sum_{m=0}^{N-1}\sum_{m=0}^{N-1} a_{\theta,r}(m,n)x(m,n)$$

where $a_{\theta,r}(m,n) = 1$, if the r^{th} ray at projection angle θ intersects the center of pixel $x(m,n)$, and $a_{\theta,r}(m,n) = 0$, otherwise. Under the assumption that the resolution of the image is sufficiently high, this model is close to the line-based projection model using $a_{\theta,r}(m,n) =$ length of the r^{th} ray at angle θ intersecting $x(m,n)$. According to the tensor discrete Fourier slice theorem (T-DFST, see Proposition 2.1), if the projection angle and detector location are decided in a sophisticated way, the 1-D DFT of the DRT can be exactly mapped on a Cartesian 2-D DFT grid, a relation that cannot be achieved with the continuous Fourier slice theorem (CFST).

Compressed Sensing (CS) is an attractive theory to reconstruct images from few measurements. Assuming *1)* a partial sensing matrix $\mathbf{R}_\Omega\mathbf{\Phi} \in \mathbb{C}^{N'\times N'}$, in which a diagonal projection matrix $\mathbf{R}_\Omega$ has the m^{th} entry 1 if $m \in \Omega$ and 0 otherwise, $|\Omega| = M \ll N'$ is chosen uniformly at random, and $\mathbf{\Phi} \in \mathbb{C}^{N'\times N'}$ where $\{\phi_n\}_{n=1}^{N'}$ is an orthonormal basis of $\mathbb{C}^{N'}$; *2)* a sparsifying transform $\mathbf{\Psi} \in \mathbb{C}^{N'\times N'}$, where $\{\psi_n\}_{n=1}^{N'}$ is orthonormal basis of $\mathbb{C}^{N'}$; and *3)* $\mathbf{y} = \mathbf{R}_\Omega\mathbf{y}^0 \in \mathbb{C}^{N'}$ with full $\mathbf{y}^0$, then the s-sparse solution in basis $\mathbf{\Psi}$ ($\|\mathbf{\Psi z}\|_0 \triangleq |\text{supp}(\mathbf{\Psi z})| \leq s \ll N'$, where $\mathbf{x} \in \mathbb{C}^{N'}$) of $\mathbf{y} = \mathbf{R}_\Omega\mathbf{\Phi z}$ can be perfectly recovered with high probability by solving the following convex optimization problem:

$$\underset{\mathbf{z}}{\text{argmin}}\, \|\mathbf{\Psi z}\|_1, \text{ s.t. } \mathbf{y} = \mathbf{R}_\Omega\mathbf{\Phi z} \tag{1}$$

with sufficient number of measurements,

$$M \geq c\mu^2(\mathbf{U})N's\log(N') \tag{2}$$

for some constant c, where the mutual coherence (MC) is $\mu(\mathbf{U}) = \max_{m,n}|u_{m,n}| \in [1/\sqrt{N'}, 1]$ for $m,n = 1,...,N'$, and $\mathbf{U} = \mathbf{\Phi\Psi}^{-1}$ [1], [2]. If, for example $\mathbf{\Phi}$ = DFT and $\mathbf{\Psi}$ = Identity so that $\mu(U) = 1/\sqrt{N}$, then (2) states that compressed sensing requires an optimally small number of measurements, up to a log factor. However, if $\mathbf{\Psi}$ = discrete Haar transform (DHT), then the MC is high, $\mu(\mathbf{U}) = 1$, and (2) predicts a barrier in the performance of compressed sensing. To overcome this, one must sample according to a nonuniform density, as was recently explained in [3].

Even when the MC is low, the requirement (2) on the number of measurements may be too stringent. In practice, one can reduce this number by solving the following non-convex minimization problem:

$$\underset{\mathbf{z}}{\text{argmin}}\, \|\mathbf{\Psi z}\|_p^p, \text{ s.t. } \mathbf{y} = \mathbf{R}_\Omega\mathbf{\Phi z}, \tag{3}$$

with $l_p(p \in (0,1))$-quasi-norm. Several results in the literature show the advantage of this approach. For example, if $\mathbf{\Phi}$ has i.i.d. Gaussian entries and $\mathbf{\Psi} = \mathbf{I}$, then we require a number of measurements

$$M \geq C_1(p)s + pC_2(p)s\log(N'/s)$$

where the constants $C_1(p)$ and $C_2(p)$ decrease as $p \to 0$. In particular, the log factor in N' vanishes as $p \to 0$ [4], [5]. Although solving (3) remains an NP-hard problem [6], it has been demonstrated [4]-[6] that a local minimum can be computed, provided $\mathbf{\Psi z}$ decays quickly and M is sufficiently large. In [7], [8], the local minimum of (3) shows higher recovery accuracy than the global minimum from (1).

In our framework, based on T-DFST, nonuniform density random sampling of 2-D Fourier samples is possible by uniformly sampling projection angles at random. We applied

*This work is supported by the Indiana State Department of Health Spinal Cord and Brain Injury Research Fund.

[1]Il Yong Chun is with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907 USA (email: chuni@purdue.edu).

[2]Ben Adcock is with the Department of Mathematics, Purdue University, West Lafayette, IN 47907 USA (email: adcock@purdue.edu).

[3]Thomas M. Talavage is with the School of Electrical and Computer Engineering and Weldon School of Biomedical Engineering, Purdue University, West Lafayette, IN 47907 USA (email: tmt@purdue.edu).

978-1-4244-7929-0/14/$26.00 ©2014 IEEE

a non-convex CS model to further reduce the sufficient number of measurements.

II. METHODS

Let $\mathbf{X} = \{X(u,v) : u,v = 0,\ldots,N-1\} \in \mathbb{C}^{N^2}$ be the $N \times N$-point 2-D DFT of the image $\mathbf{x} = \{x(m,n) : m,n = 0,\ldots,N-1\} \in \mathbb{R}^{N^2}$. Let $\mathbf{P}_\theta = \{P_\theta(w) : w = 0,\ldots,N-1\} \in \mathbb{C}^N$ be the 1D N-point DFT of N-point DRT $\mathbf{p}_\theta = \{p_\theta(r) : r = 0,\ldots,N-1\} \in \mathbb{R}^N$ for each projection angle θ with the horizontal axis, where the set of θ is defined as $\{\theta = \arctan(v/u) : \ (u,v) \in J_{N,N}\}$. According to T-DFST (Proposition 2.1) and uniform random sampling of θ, the nonuniform random Fourier measurement $\mathbf{y} = \mathbf{R}_\Omega \mathbf{X} \in \mathbb{C}^{N^2}$ can be obtained from $\mathbf{P}_\theta$. Based on this framework, we have the following discrete CT system model:

$$\mathbf{y} = \mathbf{R}_\Omega \mathbf{\Phi} \mathbf{x} + \mathbf{n},$$

where a partial DFT matrix $\mathbf{R}_\Omega \mathbf{\Phi}$ with a DFT matrix $\mathbf{\Phi} \in \mathbb{C}^{N^2 \times N^2}$ and a diagonal projection matrix $\mathbf{R}_\Omega \in \mathbb{R}^{N^2 \times N^2}$ with m^{th} entry 1 if $m \in \Omega$ and 0 otherwise, where the nonuniform randomly sub-sampled $\Omega \subseteq \{1,\ldots,N^2\}$ and $|\Omega| = M \ll N^2$. Applying $l_p(p \in (0,1))$-quasi-norm on DHT, the proposed non-convex CS CT reconstruction model can be written as:

$$\mathbf{x}^* = \underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{\Psi x}\|_p^p + \|\mathbf{x}\|_{TV} \text{ s.t. } \|\mathbf{y} - \mathbf{R}_\Omega \mathbf{\Phi x}\| < \eta, \quad (4)$$

where DHT $\mathbf{\Psi} \in \mathbb{C}^{N^2 \times N^2}$; and the anisotropic total variation (TV) transform $\|\mathbf{x}\|_{TV} = \|\mathbf{G}_1 \mathbf{x}\|_1 + \|\mathbf{G}_2 \mathbf{x}\|_1$, where $\mathbf{G}_1 \in \mathbb{C}^{N^2 \times N^2}$ and $\mathbf{G}_2 \in \mathbb{C}^{N^2 \times N^2}$ denote horizontal and vertical direction gradient transform.

A. Tensor Discrete Fourier Slice Theorem

Proposition 2.1: (Tensor discrete Fourier slice theorem, T-DFST)

$$P_{(u,v)}(w) = X(wu \bmod N, wv \bmod N),$$

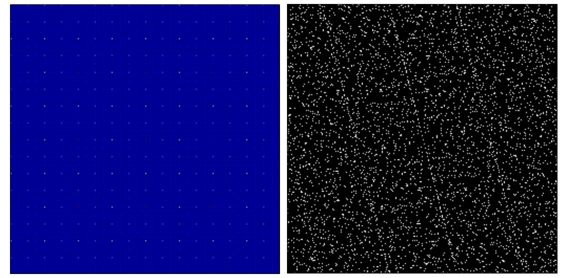
where $w = 0,\ldots,N-1$ (tensor representation of 2-D DFT in [9]).

Proof:

Let $a_{(u,v),r}(m,n) = \begin{cases} 1, & mu + nv = r \bmod N \\ 0, & \text{otherwise} \end{cases}$. Then,

$$\begin{aligned} P_{(u,v)}(w) &= \sum_{r=0}^{N-1} p_{(u,v)}(r) W_N^{wr} \\ &= \sum_{r=0}^{N-1} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} a_{(u,v),r}(m,n) x(m,n) W_N^{wr} \\ &= \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} x(m,n) \sum_{r=0}^{N-1} a_{(u,v),r}(m,n) W_N^{wr} \\ &= \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} x(m,n) W_N^{w(mu+nv)} \\ &= X(wu, wv) = X(wu \bmod N, wv \bmod N), \end{aligned}$$

where $X(u,v) = \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} x(m,n) W_N^{mu+nv}$ and $W_N = \exp(-j2\pi/N)$.



(a) (b)

Fig. 1. Mapping redundancy and random sampling in 2-D Fourier domain, based on T-DFST: (a) high mapping redundancy of T-DFST for $N = 256$ (brighter color means higher mapping redundancy) and (b) nonuniform random sampling pattern of 2-D DFT by uniform sampling of 17 angles at random for $N = 257$.

■

Note that $p_{(u,v)}(r)$, where $r = 0,\ldots,N-1$, is periodic with period N, i.e. $p_{(u,v)}(r) = p_{(u,v)}(r+N)$. The set of the projections is denoted by

$$T_{u,v} = \{(wu \bmod N, wv \bmod N) : \ w = 0,\ldots,N-1\},$$

since the signals carry the information about the 2-D DFT as Proposition 2.1 stated.

The set $J_{N,N}$ of frequency-points (u,v) should be selected in a way to cover the whole Cartesian lattice $L_{N,N} = \{(u,v) : u,v = 0,...,N-1\}$ with the minimum number of subsets $T_{u,v}$, i.e. $\bigcup_{(u,v)\in J_{N,N}} T_{u,v} = L_{N,N}$. The set $J_{N,N}$ contains $3N/2$ of (u,v) and can be defined as

$$\begin{aligned} J_{N,N} = &\{(1,v) : v = 0,\ldots,N-1\} \cup \\ &\{(2u,1) : u = 0,\ldots,N/2-1\}. \end{aligned} \quad (5)$$

The total number of projection measurements is $3N^2/2$ which exceeds the number of unknown pixels, N^2. There are many mutual intersections of the subsets $T_{u,v}$, where $(u,v) \in J_{N,N}$. In summary, if N is a power of 2, this mapping redundancy is inevitable and not suitable to reduce the number of projections. However, if N is a prime number, we can remove the high mapping redundancy. If N is prime, the cardinality of the irreducible set $J_{N,N}$ which covers Cartesian lattice $L_{N,N}$ is $N+1$. For example, it can be given by

$$J_{N,N} = \{(1,v) : v = 0,\ldots,N-1\} \cup \{(0,1)\} \text{ or}$$
$$J_{N,N} = \{(u,1) : u = 0,\ldots,N-1\} \cup \{(1,0)\}$$

Therefore, to calculate an $\mathbf{X}$ of size $N \times N$, it is sufficient to obtain $(N+1) \times N$ projection measurements, when N is prime. The mapping redundancy is graphically illustrated in Fig. 1(a); see details in [9].

B. Relaxation of Constrained Non-Convex Problem with Reweighted Constrained Convex Problem

Property 2.2: The constrained non-convex problem defined as

$$\mathbf{x}^* = \underset{\mathbf{x}\in\mathbb{C}^N}{\operatorname{argmin}} \|\mathbf{x}\|_p^p, \text{ s.t. } \textit{constraint}(\mathbf{x}), \quad (6)$$

for $0 < p < 1$, can be transformed into following reweighted constrained convex minimization problem:

$$\mathbf{x}^{(k+1)} = \underset{\mathbf{x}\in\mathbb{C}^N}{\operatorname{argmin}} \|\mathbf{Q}^{(k)}\mathbf{x}\|_1, \text{ s.t. } \textit{constraint}(\mathbf{x}), \quad (7)$$

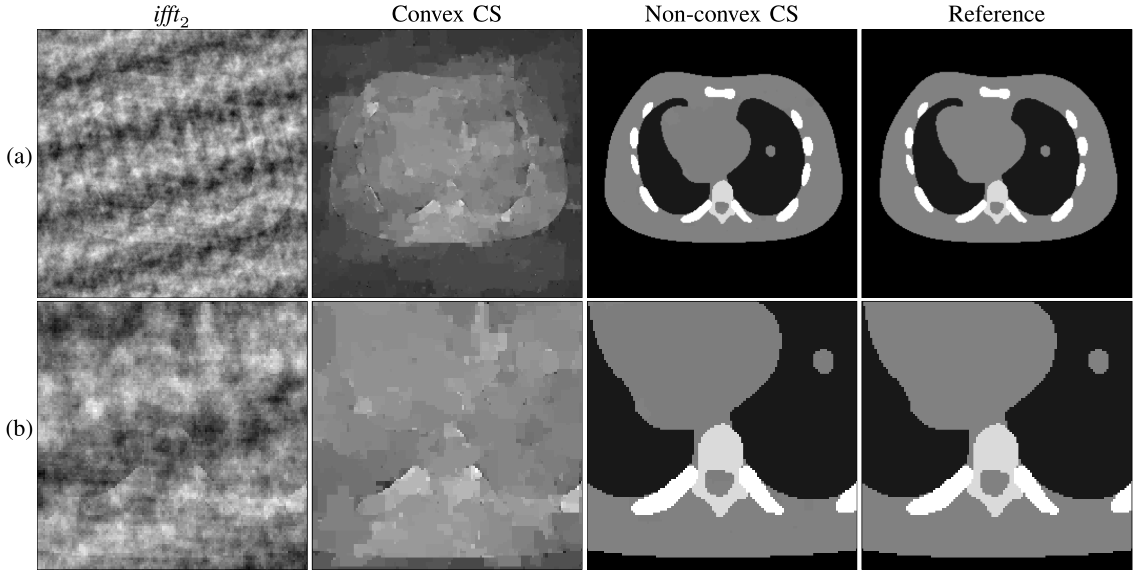


Fig. 2. Comparison of 257×257 reconstructed images from different CT reconstructions (∠ = 15 views): (a) whole images and (b) zoomed-in images

TABLE I
RECONSTRUCTION ACCURACY WITH DIFFERENT METHODS AND PROJECTION VIEWS

	SER_{dB}			$RMSE$ ($\times 10^{-1}$)		
∠	(a)	(b)	(c)	(a)	(b)	(c)
14	2.0164	4.8702	15.2234	2.6354	1.8973	0.5761
15	2.0194	5.0244	50.6299	2.6345	1.8636	0.0098
17	2.0239	5.3935	65.4033	2.6331	1.7861	0.0018

* CT reconstruction methods: (a) *ifft*$_2$, (b) convex CS, and (c) non-convex CS.

where $\mathbf{Q}^{(k)} = diag(\mathbf{q}^{(k)})$, in which $q_n^{(k)} = p(|x_n^{(k)}| + \epsilon)^{p-1}$, $n = 1, \ldots, N$, and $diag(\cdot)$ denotes the conversion of a vector into a diagonal matrix.

The brief derivation of Property 2.2 is described in [7] using the majorization minimization (MM) and approximation of $l_p(p \in (0,1))$-quasi-norm with Lipschitz continuity (i.e. $\|\mathbf{x}\|_p^p \approx l_{p,\epsilon}(\mathbf{x}) = \sum_{n=1}^{N}(|x_n| + \epsilon)^p$). Note that $\mathbf{x}^{(k+1)}$ of (7) converges to a local minimum of (6) if $\epsilon \to 0$.

According to Property 2.2, (4) can be transformed as

$$\mathbf{x}^{(k+1)} = \underset{\mathbf{x}}{\operatorname{argmin}} \; \|\mathbf{Q}^{(k)}\mathbf{\Psi}\mathbf{x}\|_1 + \|\mathbf{G}_1\mathbf{x}\|_1 + \|\mathbf{G}_2\mathbf{x}\|_1 \quad \text{s.t. } \|\mathbf{y} - \mathbf{R}_\Omega\mathbf{\Phi}\mathbf{x}\|_2^2 < \eta, \tag{8}$$

where $\mathbf{Q}^{(k)} = diag(\mathbf{q}^{(k)})$, with $q_n^{(k)} = p(|[\mathbf{\Psi}\mathbf{x}^{(k)}]_n| + \epsilon)^{p-1}$, $n = 1, \ldots, N^2$. Note that the reweighted l_1-norm minimization is expected to recover sparse signals with lower error than a reweighted l_2-norm minimization (e.g., FOCUSS) [10].

C. Reweighted Constrained l_1-Norm Minimization by SB

The Split Bregman (SB) method is known to exhibit rapid and efficient convergence for l_1-norm [11] minimization. Using a simplified Bregman iteration technique [12], (8) can be reduced to a sequence of unconstrained problems:

$$\mathbf{x}^{(k+1)} = \underset{\mathbf{x}^{(k)}}{\operatorname{argmin}} \|\mathbf{Q}^{(k)}\mathbf{\Psi}\mathbf{x}^{(k)}\|_1 + \|\mathbf{G}_1\mathbf{x}^{(k)}\|_1 + \|\mathbf{G}_2\mathbf{x}^{(k)}\|_1 + (\alpha/2)\|\mathbf{y}^{(k)} - \mathbf{R}_\Omega\mathbf{\Phi}\mathbf{x}^{(k)}\|_2^2; \tag{9}$$

$$\mathbf{y}^{(k+1)} = \mathbf{y}^{(k)} + \mathbf{y} - \mathbf{R}_\Omega\mathbf{\Phi}\mathbf{x}^{(k+1)},$$

After transforming (9) to a constrained problem (i.e. $\mathbf{d}_\Psi^{(k)} = \mathbf{\Psi}\mathbf{x}^{(k)}$, $\mathbf{d}_1^{(k)} = \mathbf{G}_1\mathbf{x}^{(k)}$, and $\mathbf{d}_2^{(k)} = \mathbf{G}_2\mathbf{x}^{(k)}$), (9) is equivalent to the following two phase algorithm, via SB:

$$\begin{aligned}
&(\mathbf{x}^{(k+1)}, \mathbf{d}_\Psi^{(k+1)}, \mathbf{d}_1^{(k+1)}, \mathbf{d}_2^{(k+1)}) \\
&= \underset{\mathbf{x}^{(k)}, \mathbf{d}_\Psi^{(k)}, \mathbf{d}_1^{(k)}, \mathbf{d}_2^{(k)}}{\operatorname{argmin}} \|\mathbf{Q}^{(k)}\mathbf{d}_\Psi^{(k)}\|_1 + \|\mathbf{d}_1^{(k)}\|_1 + \|\mathbf{d}_2^{(k)}\|_1 + \\
&\qquad (\alpha/2)\|\mathbf{y}^{(k)} - \mathbf{R}_\Omega\mathbf{\Phi}\mathbf{x}^{(k)}\|_2^2 + \\
&\qquad (\beta/2)\|\mathbf{d}_\Psi^{(k)} - \mathbf{\Psi}\mathbf{x}^{(k)} - \mathbf{b}_\Psi^{(k)}\|_2^2 \\
&\qquad (\gamma/2)\|\mathbf{d}_1^{(k)} - \mathbf{G}_1\mathbf{x}^{(k)} - \mathbf{b}_1^{(k)}\|_2^2 + \\
&\qquad (\gamma/2)\|\mathbf{d}_2^{(k)} - \mathbf{G}_2\mathbf{x}^{(k)} - \mathbf{b}_2^{(k)}\|_2^2; \\
&\mathbf{b}_\Psi^{(k+1)} = \mathbf{b}_\Psi^{(k)} + \mathbf{\Psi}\mathbf{x}^{(k+1)} - \mathbf{d}_\Psi^{(k+1)}, \\
&\mathbf{b}_1^{(k+1)} = \mathbf{b}_1^{(k)} + \mathbf{G}_1\mathbf{x}^{(k+1)} - \mathbf{d}_1^{(k+1)}, \\
&\mathbf{b}_2^{(k+1)} = \mathbf{b}_2^{(k)} + \mathbf{G}_2\mathbf{x}^{(k+1)} - \mathbf{d}_2^{(k+1)}.
\end{aligned} \tag{10}$$

Because l_1 and l_2 components are decomposed, we can solve (10) efficiently by minimizing it separately with respect to $\mathbf{x}^{(k)}$, $\mathbf{d}_\Psi^{(k)}$, $\mathbf{d}_1^{(k)}$, and $\mathbf{d}_2^{(k)}$. The $\mathbf{d}_\Psi^{(k+1)}$, $\mathbf{d}_1^{(k+1)}$, and $\mathbf{d}_2^{(k+1)}$ can be quickly solved by separability of norms and an element-wise soft-shrinkage operator:

$$\begin{aligned}
d_{\Psi,n}^{(k+1)} &= softshrink([\mathbf{\Psi}\mathbf{x}^{(k+1)}]_n + b_{\Psi,n}^{(k)}, q_n^{(k)}/\beta), \\
d_{1,n}^{(k+1)} &= softshrink([\mathbf{G}_1\mathbf{x}^{(k+1)}]_n + b_{1,n}^{(k)}, 1/\gamma), \\
d_{2,n}^{(k+1)} &= softshrink([\mathbf{G}_2\mathbf{x}^{(k+1)}]_n + b_{2,n}^{(k)}, 1/\gamma),
\end{aligned}$$

where $softshrink(x, \alpha) = (x/|x|)max(|x| - \alpha, 0)$ and $n = 1, \ldots, N^2$. Therefore the total reconstruction time depends on the computational cost to solve (10) with respect to $\mathbf{x}^{(k)}$:

$$\mathbf{x}^{(k+1)} = \mathbf{\Phi}^H\mathbf{\Lambda}^{-1}\mathbf{\Phi}\mathbf{h},$$

where diagonal matrix $\mathbf{\Lambda} = \alpha\mathbf{R}_\Omega^T\mathbf{R}_\Omega + \beta\mathbf{I} + \gamma\mathbf{\Phi}(\mathbf{G}_1^H\mathbf{G}_1 + \mathbf{G}_2^H\mathbf{G}_2)\mathbf{\Phi}^H$ and $\mathbf{h} = \alpha\mathbf{\Phi}^H\mathbf{R}_\Omega^T\mathbf{y}^{(k)} + \beta\mathbf{\Psi}^H(\mathbf{d}_\Psi^{(k)} - \mathbf{b}_\Psi^{(k)}) + \gamma\mathbf{G}_1^H(\mathbf{d}_1^{(k)} - \mathbf{b}_1^{(k)}) + \mathbf{G}_2^H(\mathbf{d}_2^{(k)} - \mathbf{b}_2^{(k)})$. Note that $\mathbf{G}_1^H\mathbf{G}_1 +$

$\mathbf{G}_2^H \mathbf{G}_2$ has circulant structure with periodic boundary condition. The main computation is only a pair of *fft*$_2$ and *ifft*$_2$. We should also note that $\mathbf{Q}^{(k)}$ is updated for every outer iteration.

III. SIMULATION RESULTS AND DISCUSSIONS

Reconstruction algorithms were tested on a 257×257 NCAT chest phantom image having intensity $\in [0, 1]$. The DRT-based sinogram was generated with $\angle \times 257$ parallel-beam scanner geometry and defined $J_{N,N}$, where uniform randomly sampled $\angle$ = 14, 15, and 17 views, which approximately correspond to 5.5%, 6.0%, and 6.5% of the 258 total projection angles. The ray spacing is variable with different angles. An additive noise is not considered. For the initial guess, $\mathbf{x}^{(0)} = \mathit{ifft}_2(\mathbf{y})$ was obtained. For both of the CS models, $\alpha = 1$, $\beta = 100$, $\gamma = 100$ are used. For non-convex CS, $p = 10^{-3}$ and $\epsilon = 5 \times 10^{-2}$ are additionally used. A DHT filter size of 4 was used. The stopping criterion can be calculated as $tol(k) = \|\mathbf{y} - \mathbf{R}_\Omega \boldsymbol{\Phi} \mathbf{x}^{(k)}\|_2^2 / \|\mathbf{y}\|_2^2$. The error minimization is evaluated with the following two measurements: $RMSE_{\log}(k) = \log_{10}\left(RMSE(\mathbf{x}^{true}, \mathbf{x}^{(k)})\right)$; $SER_{\text{dB}}(k) = 20 \log_{10}\left(\|\mathbf{x}^{true} - \mathbf{x}^{(k)}\|_2 / \|\mathbf{x}^{true}\|_2\right)$, where SER stands for signal to error ratio.

A. Reconstruction Accuracy

In terms of reconstruction accuracy, the non-convex CS reconstruction model outperforms the convex CS model, and accomplishes almost exact image reconstruction with only 15 projection angles (Fig. 2 and Table I).

B. Error Minimization Behavior

Fig. 3 shows that a larger number of measurements results in faster convergence and smaller error. Note that one cannot always expect to find a global minimum from $l_{p,\epsilon}(p, \epsilon \in (0, 1))$ minimization. Moreover, the stability and simplicity of the MM algorithm frequently comes at the price of slow convergence: Fig. 3 illustrates the "waterfall" convergence behavior of the non-convex minimization.

C. Practical Applicability of T-DFST

We should note that it is premature to apply this technique in practice. *1)* Actual imaging using the DRT-based geometry is tricky: non-uniform ray spacing across projections is difficult to implement. *2)* Although the DRT-based projection model can become close to a line-based model, it cannot match. Linear system solutions to this limitation are introduced in [13] and [14] to transform continuous Radon transform (CRT) data to DRT data. However, more CRTs are required at each angle to successfully estimate the DRTs (i.e. the actual sampling reduction here is over-estimated). *3)* The usage of both of the CFST and T-DFST for CT reconstruction is fragile to the practical case of Poisson noise.

IV. CONCLUSIONS

A non-convex CS CT reconstruction using T-DFST has been presented. The nonuniform randomness of Fourier samples is provided by uniform sampling of random projection angles. MM is used for non-convex optimization relaxation

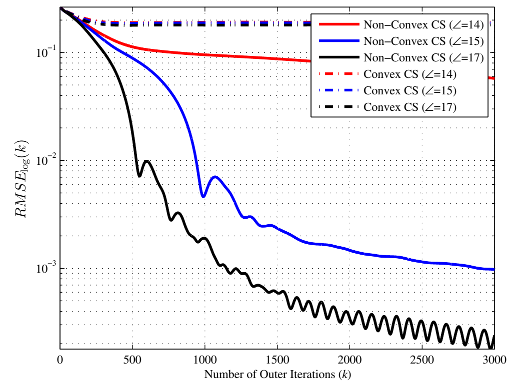


Fig. 3. Error minimization behavior for different CS models and projection angles

and SB is applied for efficient implementation. The method can achieve almost perfect image reconstruction from fewer measurements than a convex CS model. However, to apply this framework in practice, several limitations exist as we described above.

REFERENCES

[1] B. Adcock and A. C. Hansen, "Generalized sampling and infinite-dimensional compressed sensing," DAMTP Tech. Rep. NA2011/12, University of Cambridge, Available: http://www.math.purdue.edu/~adcock/Papers/BAACHGSCS.pdf, May 2013.

[2] E. J. Candes and Y. Plan, "A probabilistic and RIPless theory of compressed sensing," *IEEE Trans. Inform. Theory*, vol. 57, no. 11, pp. 7235 – 7254, Nov. 2011.

[3] B. Adcock, A. C. Hansen, C. Poon, and B. Roman, "Breaking the coherence barrier: a new theory for compressed sensing," *ArXiv preprint cs.IT/1302.0561v3*, Feb. 2014.

[4] R. Chartrand and V. Staneva, "Restricted isometry properties and nonconvex compressive sensing," *Inverse Probl.*, vol. 24, no. 3, p. 035020, Jun. 2008.

[5] Y. Shen and S. Li, "Restricted p-isometry property and its application for nonconvex compressive sensing," *Adv. Comput. Math.*, vol. 37, no. 3, pp. 441 – 452, Oct. 2012.

[6] D. Ge, X. Jiang, and Y. Ye, "A note on complexity of Lp minimization," *Math. Program.*, vol. 129, no. 2, pp. 285 – 299, Oct. 2011.

[7] I.Y. Chun and T. Talavage, "Efficient compressed sensing statistical x-ray/CT reconstruction from fewer measurements," in *Proc. Intl. Mtg. on Fully 3D Image Recon. in Rad. and Nuc. Med*, Lake Tahoe, CA, Jun. 2013, pp. 30 – 33.

[8] ——, "Fast non-convex statistical compressed sensing MRI reconstruction based on approximated $Lp(0 < p < 1)$-quasi-norm with fewer measurements than using L1-norm," in *Proc. 21st Intl. Soc. Mag. Res. Med.*, Salt Lake City, UT, Apr. 2013.

[9] A. M. Grigoryan and S. S. Agaian, *Multidimensional Discrete Unitary Transforms: Representation, Partitioning, and Algorithms.* CRC Press, 2003.

[10] E. J. Candes, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted $l1$ minimization," *J. Fourier Anal. Appl.*, vol. 14, no. 5-6, pp. 877 – 905, Dec. 2008.

[11] T. Goldstein and S. Osher, "The split Bregman method for L1-regularized problems," *SIAM Journal on Imag. Sciences*, vol. 2, no. 2, pp. 323 – 343, Apr. 2009.

[12] W. Yin, S. Osher, D. Goldfarb, and J. Darbon, "Bregman iterative algorithms for l_1-minimization with applications to compressed sensing," *SIAM Journal on Imag. Sciences*, vol. 1, no. 2, pp. 143 – 168, Mar. 2008.

[13] B. Recur, P. Desbarats, and J. P. Domenger, "Radon and Mojette projections equivalence for tomographic reconstruction using linear systems," in *Proc. 16th WSCG*, Plzen, Czech Republic, Feb. 2008, pp. 191 – 198.

[14] A. Grigoryan, "Image reconstruction from finite number of projections: method of transferring geometry," *IEEE Trans. Image Process.*, vol. 22, no. 12, pp. 4738 – 4751, Dec. 2013.