Reconstruction of Sparse-View Tomography via Preconditioned Radon Sensing Matrix

Prasad Theeda[1], Praveen Kumar P. U.[1, 3], C. S. Sastry[1], and P. V. Jampana[2]

[1] Department of Mathematics, Indian Institute of Technology Hyderabad, India.

[2] Department of Chemical Engineering, Indian Institute of Technology Hyderabad, India , {ma13p1004, csastry, pjampana}@iith.ac.in

[3] Department of Computer Science, CHRIST University, Bengaluru, India, , praveen577302@gmail.com

Abstract

Computed Tomography (CT) is one of the significant research areas in the field of medical image analysis. As X-rays used in CT image reconstruction are harmful to the human body, it is necessary to reduce the X-ray dosage while also maintaining good quality of CT images. Since medical images have a natural sparsity, one can directly employ compressive sensing (CS) techniques to reconstruct the CT images. In CS, sensing matrices having low coherence (a measure providing correlation among columns) provide better image reconstruction. However, the sensing matrix constructed through the incomplete angular set of Radon projections typically possesses large coherence. In this paper, we attempt to reduce the coherence of the sensing matrix via a square and invertible preconditioner possessing a small condition number, which is obtained through a convex optimization technique. The stated properties of our preconditioner imply that it can be used effectively even in noisy cases. We demonstrate empirically that the preconditioned sensing matrix yields better signal recovery than the original sensing matrix.

Keywords: Computed tomography, Compressive Sensing, Preconditioning, TVAL3

1 Introduction

X-ray Computed Tomography (CT) is widely used in hospitals and clinics for diagnosis and intervention. CT is a technique for reconstructing the cross-section of an object from measurements that are essentially the line integrals of it. The general image reconstruction in CT is a mathematical process that generates an image from X-ray projection data acquired at different angles around the

object. As X-rays are harmful to human bodies, minimization of risk posed by radiation receives the attention of many researchers [1, 2]. The basic objective of CT in medical use is to obtain high quality images from projection data with as little of radiation dosage as possible.

The Reconstruction methodology of CT images is divided into two categories, namely, analytical and iterative methods. The filtered back-projection (FBP) method [3] is a commonly used analytical method for image reconstruction in CT due to its computational efficiency and numerical stability [4]. Despite its popularity, the FBP suffers from systematic geometric distortion and streak artifacts when the measured projection data are not sufficient. When dealing with insufficient data, the iterative class of methods has better performance [5]. In recent years, the reconstruction methods [4] [6] [7] based on optimization have become popular due to their ability in reducing number of X-ray projection samples while maintaining good reconstruction fidelity. The basic premise behind these methods lies in efficiently exploiting natural sparsity present in CT images using ideas from the emerging field of Compressive Sensing.

Compressive Sensing (CS) theory [8, 9] relies on two fundamental principles, *Sparsity* and *Coherence*. Sparse signals $x \in \mathbb{R}^M$ can be recovered from small number m (where $m < M$) of measurements y satisfying $y = Ax$ with $k < m$ where k is the number of non zero elements in x. The results pertaining to the recovery guarantee depend on the coherence and the Restricted Isometry Property (RIP) of the sensing matrix A which we discuss in Section 3. CS theory states that matrices having smaller coherence give better recovery, which signifies the need for reducing the coherence of the sensing matrix.

Over the last few years, researchers [10, 11, 12, 13, 14, 15] proposed various methods to reduce the coherence of a given sensing matrix, which has potential to imply fewer measurements in reconstruction. Elad [10] proposed an iterative method which minimizes the $t-$averaged mutual coherence of a sensing matrix. The authors in [11] considered an application to MRI, which uses an incoherent criterion based on Point Spread Function (SPF). Duarte et. al. [12] considered the problem of making any subset of columns in sensing matrix as approximately orthogonal as possible by making its Gram matrix lie close to identity matrix. Xu et.al. [13] proposed an alternating minimization method to find a sensing matrix which is as close to an equiangular Tight Frame (ETF) as possible. The minimization of mutual coherence was investigated in a formulation of sequential non-smooth convex programming, which was solved by using subgradient projection algorithm in [15]. Most of the preconditioning methods that minimize coherence provide nonsquare matrices, which may not address the compressed data acquisition problem in a true sense. The square and invertible preconditioners, however, imply that the original under-determined as well as preconditioned systems are equivalent. To our knowledge, square and invertible preconditioners with good condition number, applicable to CT, are not available in the literature.

The recovery of images in CT from incomplete measurements is an ill-posed problem [16]. The reconstruction from a limited set of views (referred to as

Limited angle or sparse-view tomography) was addressed using wavelets by B. Sahener et. al [17]. Another wavelet based statistical inversion method was proposed in [18]. These methods efficiently incorporate a priori information about missing portions of data into reconstruction process. Of late, addressing this ill-posed problem in sparsity framework became an active direction of research, which does not in general require any a priori information about missing views. The work presented in [4] highlighted the importance of sparsity based methods in CT. The authors of [19] (and references therein) discussed the CT reconstruction in circular cone-beam geometry by constrained TV minimization. Ritschl et. al. [20] proposed an improved TV method within the Adaptive Steepest Descent-Projection Onto Convex Sets (ASD-POCS) for CT image reconstruction. In [7], a novel algorithm, called linearized Split-Bregman method, was proposed to efficiently solve the reconstruction of CT images. By approximating an intermediate matrix that arises in the Split-Bregman method, the authors of [21] addressed incomplete data problem in tomography. This being an approximate method does not involve any kind of preconditioning.

To the best of our knowledge, the methods available in the literature mostly are aimed at improving the performance of the solvers while efficiently exploiting the inherent sparsity in CT images. The present work, however, aims at improving the recovery property of the Radon sensing matrix via preconditioning. In particular, we determine a square and invertible preconditioner, through a convex optimization problem, that improves the coherence of the Radon sensing matrix while simultaneously maintaining small condition number for the precontidioner. Experimentally, we demonstrate that the improved coherence reduces the reconstruction error in CT. Consequently, for a given error tolerance, our preconditioning approach uses a reduced number of measurements even in the noisy setting.

The remainder of the paper is organized into several sections. In Section 2, we discuss the basics of CT, discretization of the Radon transform and reconstruction of images from reduced set of measurements. We record the relevant concepts of CS based techniques in Section 3. In Section 3.2, we discuss the impact of preconditioning on spare recovery. We propose a new preconditioning method in Section 4. We present experimental results and concluding remarks in Sections 5 and 6 respectively.

2   Basics of Computed Tomography

The X-ray source, together with primary collimators, provides a fine beam of radiation that passes through the object, the intensity of the beam is then measured by a detector. In general, for any θ and τ the projection data are measured as

$$(Rf)(\theta, \tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(p, q)\delta(p\cos\theta + q\sin\theta - \tau)dpdq, \qquad (1)$$

called the *Radon Transform*. In (1), f is the density function and δ-function selects the ray point set. The limits of p and q are given by the object size. The
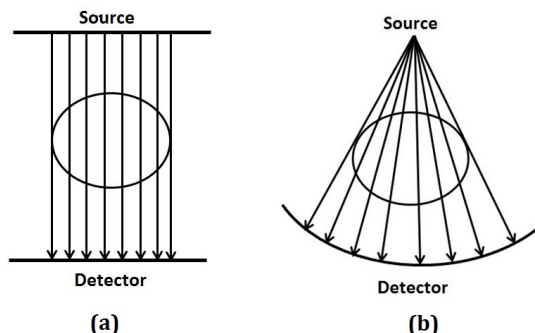
Fig. 1: Data acquisition in (a) Parallel-Beam and (b) Fan-Beam geometries.

task of reconstruction of the original image $f(p,q)$ from its projection represen-tation $R(\theta, \tau)$ is the problem of finding the inverse Radon transform. The data acquisition in CT is done in several ways such as parallel, fan and cone beam geometries. The parallel-beam and fan-beam modes of data acquisition, shown in Figure 1, are connected by a transformation, called *rebinning* [22].

2.1 Discretization of Radon Transform

Consider a ray corresponding to some view θ_i and radial parameter τ_j. In discrete setting, the Radon measurement may be rewritten as

$$(Rf)(\theta_i, \tau_j) = \sum_l a_{i,j}(l)p_l, \tag{2}$$

where $a_{i,j}(l) = \begin{cases} w_{ij} & \text{if } (\theta_i, \tau_j) \text{ ray hits } l^{th} \text{ pixel} \\ 0 & \text{else.} \end{cases}$

In formulating (2), one may consider bilinear interpolation of pixels that fall in the path of the ray. The value w_{ij} is the weight obtained through interpolation and p_l stands for the intensity value of lth pixel. The equation (2) may be rewritten as

$$(Rf)(\theta_i, \tau_j) = [a_{i,j}(1) \ldots a_{i,j}(M)][p_1 \ldots p_M]^T, \tag{3}$$

where $M = n^2$, the total number of pixels in the image of size $n \times n$. Using all θ_i and τ_j, one obtains a matrix system

$$\tilde{y} = \tilde{A}x, \tag{4}$$

where $\tilde{y}$ contains Radon measurements (i.e. $(Rf)(\theta_i, \tau_j)$) in vector form. Ac-cordingly, $\tilde{A}$ is a weighted matrix (whose elements are $a_{i,j}(l)$) and x is the vector whose elements are p_l. The row size of $\tilde{A}$ is dictated by the number of radial and angular sampling parameters and the column size of $\tilde{A}$ is the dimension of

the image in vector form. In generating the sensing matrix $\tilde{A}$, one may consider the following standard radial and angular sampling [22]:

Radial Sampling:

$$\tau_j = \delta\tau \left(j - \frac{n}{2} \right), \tag{5}$$

Angular Sampling:

$$\theta_i = \frac{i + 0.5}{n} \pi, \tag{6}$$

where $i, j \in \{0, 1, \cdots, n - 1\}$, $\delta\tau$ is the increment in radial sampling.

2.2 Reconstruction from reduced measurements

Let $\tilde{A}$ be the matrix corresponding to full set of measurements and $\tilde{y}$ its corresponding projection data. In sparse-view CT, one restricts the number of angles, and accordingly the row size of the sensing matrix $\tilde{A}$ gets restricted. Mathematically, for some row-restricted identity matrix R, the restricted projection set may be taken as $R\tilde{y}$ (denoted as y) with the corresponding sensing matrix $R\tilde{A}$ (denoted as A). Generally, in sparse-view CT, one deals with a small set of projection views, and hence the system in $y = Ax$ becomes under-determined, admitting thereby infinitely many solutions. The inherent sparsity present in CT images makes CS a natural choice [4] for recovering the underlying image.

3 Compressive Sensing

Compressed Sensing (CS) technique [8, 9] recovers $x \in \mathbb{R}^M$ from a few of its linear measurements $y \in \mathbb{R}^m$ through an efficient recovery process via the concept of sparsity. From the measurement vector y and the sensing mechanism, one obtains a system $y = Ax$, where A is an $m \times M$ ($m < M$) matrix. The measure for sparsity is provided by $\|\cdot\|_0$ norm, where $\|x\|_0 := |\{j \in \{1, 2, \ldots, M\} : x_j \neq 0\}|$. The sparsest solution can be obtained from the following minimization problem:

$$P_0(A, y) : \min_x \|x\|_0 \text{ subject to } \quad Ax = y.$$

This minimization problem is known to be NP-hard [8]. Two classes of methods, namely greedy methods [23] and convex relaxation of $P_0(A, y)$, are available for recovering the k-sparse x (that is, $\|x\|_0 \leq k$). The convex relaxation of $P_0(A, y)$ can be taken as

$$P_1(A, y) : \min_x \|x\|_1 \text{ subject to } \quad Ax = y.$$

Despite the solution to an under-determined linear system being non-unique, CS theory provides sufficient conditions under which unique recovery of the sparse signal is possible. The coherence, $\mu(A)$, of a matrix A is the largest absolute normalized inner product between different columns from A, that is,

$$\mu(A) = \max_{1 \leq i,j \leq M, \, i \neq j} \frac{|\mathbf{a}_i^T \mathbf{a}_j|}{\|\mathbf{a}_i\|_2 \|\mathbf{a}_j\|_2},$$

where, a_k stands for the k-th column in A. An $m \times M$ matrix A is said to satisfy the Restricted Isometry Property (RIP) [8, 9] of order k with constant δ_k $(0 < \delta_k < 1)$ if for all vectors $x \in \mathbb{R}^M$ with $\|x\|_0 \le k$, we have

$$(1 - \delta_k) \|x\|_2^2 \le \|Ax\|_2^2 \le (1 + \delta_k) \|x\|_2^2. \tag{7}$$

Roughly speaking, RIP measures the "overall conditioning" of the set of $m \times k$ submatrices of A and establishes sufficient condition for exact recovery. A smaller value for δ_k implies better sparse recovery properties. The following results ([8, 9]) establish the connection between δ_k and μ, and the equivalence relation between P_0 and P_1 problems respectively.

Theorem 3.1. *Suppose an $m \times M$ matrix A has the $(2k, \delta)$ restricted isometry property with $\delta < \sqrt{2} - 1$, then P_0 and P_1 have the same $k-$sparse solution if P_0 has a $k-$sparse solution.*

Proposition 3.1. *Suppose that $a_1, \ldots, a_M$ are the unit norm columns of the matrix A with coherence μ. Then A satisfies RIP of order k with constant $\delta_k = (k - 1)\mu$.*

The above result indicates that matrices with smaller coherence enable us to have smaller δ_k.

Coherence plays an important in establishing the guaranteed recovery of sparse signals via Orthogonal Matching Pursuit (OMP) or Basis Pursuit (BP), as summarized by the following result [8].

Theorem 3.2. *An arbitrary $k-$sparse signal x can be uniquely recovered as a solution to the problem $P_0(A, y)$ using OMP and BP, provided*

$$k < \frac{1}{2} \left(1 + \frac{1}{\mu(A)} \right). \tag{8}$$

$\square$

From (8), it is clear that smaller value for coherence results in better bounds on sparsity. Consequently, coherence reduction methods attain significance.

3.1 TV norm minimization and solvers

Total variation (TV) regularization is a generalization of the $\|\cdot\|_1$ regularization in compressive sensing problems. The advantage [24] of TV minimization is that it can recover not only sparse signals or images, but also dense staircase signals or piecewise constant images. Hence, it is a natural choice in problems such as the reconstruction in CT.

Among existing versions of TV regularization methods, *Total Variation Augmented Lagrangin and ALternating direction ALgorithm (TVAL3)* [24] is popular and outperforms other state-of-the-art solvers in the field.

For reconstructing sparse x from $y = Ax$, in total variation methods, one considers the following gradient based minimization:

$$\mathcal{P}_{tv} : \min_x \quad \|\nabla x\|_1 \text{ subject to } \quad Ax = y. \tag{9}$$

where $\|\nabla x\|_1 = \sum_{i,j} \sqrt{|x_{i+1,j} - x_{i,j}|^2 + |x_{i,j+1} - x_{i,j}|^2}$. This model is difficult to solve because of non-differentiability and non-linearity of TV-norm. The Alternating Direction Method (ADM) [25] minimizes the augmented Lagrangian function through alternating minimization schemes. The equivalent variant of the problem (9) is as follows:

$$\mathcal{P}_{tv1} : \begin{array}{c} \min_{d,x} \quad \|d\|_1 \\ \text{subject to} \quad Ax = y, \ d = \nabla x. \end{array} \tag{10}$$

Its corresponding augmented Lagrangian function and minimization problems are as follows:

$$\mathcal{P}_{al} : \min_{d,x} \quad \mathcal{L}_A(d, x) \tag{11}$$

where $\mathcal{L}_A(d, x)$ is augmented Lagrangian function

$$\mathcal{L}_A(d, x) = \|d\|_1 - \beta_1^T (\nabla x - d) + \frac{\beta_2}{2} \|\nabla x - d\|_2^2 - \beta_3^T (Ax - y) + \frac{\beta_4}{2} \|Ax - y\|_2^2.$$

The alternating direction method is used here to solve the problem (11). For a fixed x, the minimizer d_i for all i can be obtained via following formula:

$$d_i = \max \left\{ \|\nabla_i x - \frac{(\beta_1)_i}{\beta_2}\| - \frac{1}{\beta_2}, 0 \right\} \frac{\nabla_i x - \frac{(\beta_1)_i}{\beta_2}}{\|\nabla_i x - \frac{(\beta_1)_i}{\beta_2}\|}.$$

For a fixed d_i, with respect to x, one minimizes the quadratic in 11 approximately by taking steepest descent step. After each steepest descent step, d_i is updated and the process is repeated until 11 results in a solution which is approximate with respect to the prescribed tolerance [26]. Let $(\hat{d}, \hat{x})$ be an approximate solution to (11). The multipliers are then updated with the following two formulas for all i:

$$(\beta_1)_i = (\beta_1)_i - \beta_2 (\nabla_i \hat{x} - \hat{d}_i),$$

$$\beta_3 = \beta_3 - \beta_4 (A\hat{x} - y).$$

The performance of TVAL3 depends on the parameters β_2, β_4. In general, β_2 is chosen in accordance with the noise level in the measurement vector y and the sparsity level of the underlying signal x. An elaborate description of TVAL3 along with the selection of β_2, β_4 can be found in [24, 26] and the references therein.

3.2 Impact of preconditioning on sparse recovery

To begin with, we analyze the impact of preconditioning on sparse recovery as relevant to CT. For a nonsingular $m \times m$ matrix P, the system of equations $y = Ax$ and $Py = PAx$ have same set of solutions. Consequently, the problems $P_0(A, y)$, $P_0(PA, Py)$ and $P_1(A, y)$, $P_1(PA, Py)$ are equivalent respectively. In view of this, Theorem 3.2 may be restated as follows:

Theorem 3.3. *Let P be any nonsingular matrix. An arbitrary $k-$sparse signal x can be uniquely recovered as a solution to the problem $P_0(PA, Py)$ using OMP and BP, provided*

$$k < \frac{1}{2}\left(1 + \frac{1}{\mu(PA)}\right).$$ (12)

□

When $\mu(PA) < \mu(A)$, (12) shows improved bound on the sparsity parameter for the new system $PAx = Py$. The improvement in the bound is expected to contribute to reduction in the number of measurements to be used in reconstruction. The objective of present study is to determine a square and invertible matrix P possessing small condition number, and to analyze its bearing on reconstruction in CT.

4   On designing preconditioners

The discretization of Radon transform along with integral approximation on a rectangular grid makes Radon inversion ill-conditioned [27]. Further, the preconditioner with large condition number amplifies the power of noise, if present, in Radon measurements. Due to these reasons, designing of preconditioner with low condition number attains importance. The problem of determining the preconditioner P, with small condition number such that $\mu(PA) < \mu(A)$, is posed as the following optimization problem:

$$\mathcal{P}_\mu : \begin{array}{c} \min\limits_{P \neq 0} \\ \end{array} \quad \max\limits_{i \neq j} \frac{|\langle P\mathbf{a}_i, P\mathbf{a}_j\rangle|}{\|P\mathbf{a}_i\|_2 \|P\mathbf{a}_j\|_2},$$
$$\text{subject to} \quad \kappa(P) \leq \gamma.$$ (13)

This problem, however, is non-convex. We now reformulate it as follows.

The coherence of the preconditioned matrix PA is the maximum absolute value of off-diagonal elements of the Gram matrix $A^T P^T PA$. Therefore, we approximate the matrix $A^T P^T PA$ to the identity matrix I in terms of the Frobenius norm with a constraint on the condition number, which is stated as follows:

$$\mathcal{P} : \begin{array}{c} \min\limits_{X} \\ \end{array} \quad \|A^T XA - I\|_F$$
$$\text{subject to} \quad \kappa(X) \leq \gamma,$$ (14)

for some suitable choice of γ. In (14), X is $P^T P$, a positive definite matrix and $\kappa(X)$ is the spectral condition number of X, which is defined [28] as follows:

$$\kappa(X) = \begin{cases} \frac{\lambda_{\max}(X)}{\lambda_{\min}(X)}, & \text{if } \lambda_{\min}(X) > 0, \\ \infty, & \text{if } \lambda_{\min}(X) = 0 \text{ and } \lambda_{\max}(X) > 0, \\ 0, & \text{if } X = 0, \end{cases}$$ (15)

where $\lambda_{\max}(X)$ and $\lambda_{\min}(X)$ are the maximum and minimum eigenvalues of X respectively. The constrained optimization problem may be considered in the following unconstrained form:

$$\mathcal{P} : \min_X \quad \|A^T X A - I\|_F^2 + \eta \kappa(X), \tag{16}$$

for some suitable choice of η. We replace the nonconvex function $\kappa(X)$ with $\nu\|X\|_F^2 - \lambda \log \det(X)$ and obtain the following unconstrained convex optimization problem:

$$\mathcal{P}_c : \min_X \quad \|A^T X A - I\|_F^2 - \lambda \log \det(X) + \nu\|X\|_F^2, \tag{17}$$

where $\det(\cdot)$ denotes the determinant and $\|\cdot\|_F$ denotes the Frobenius norm. $\log \det(X)$ and $\|X\|_F$ respectively account for the sum of logarithm of eigen values of X and the maximum eigen value of X. The presence of "negative $\log \det(X)$" in (17) enforces the smallest eigen value, in particular, to be away from zero. Consequently, for a suitable choice of λ, minimization of $\|X\|_F^2 - \lambda \log \det(X)$ reduces the gap between largest and smallest eigen values of X, and hence leads to an approximate minimization of the nonconvex function $\kappa(X)$. The cost functions involving the negative log-determinant penalties and upper-bounding $\kappa(X)$ in terms of such penalties were studied in [29, 30].

The objective function in (17), $f(X) = \|A^T X A - I\|_F^2 - \lambda \log \det(X) + \nu\|X\|_F^2$ is convex and differentiable. The derivative [31] of each term in $f(X)$ with respect to X is

$$\nabla_X \|A^T X A - I\|_F^2 = 2A(A^T X A - I)A^T \tag{18}$$

$$\nabla_X \log \det X = X^{-T} \tag{19}$$

$$\nabla_X \|X\|_F^2 = 2X. \tag{20}$$

Therefore, the derivative of $f(X)$ is $\nabla_X f(X) = 2A(A^T X A - I)A^T - \lambda X^{-T} + 2\nu X$. Setting $\nabla_X f(X)$ to 0, we obtain the following equation, which provides optimal X.

$$2A(A^T X A)A^T - 2AA^T - \lambda X^{-T} + 2\nu X = 0. \tag{21}$$

The equation (21) can be solved by using gradient methods. In particular, we obtain X by solving it using conjugate gradient method [32]. Finally, we determine the preconditioner P by splitting X as $P^T P$ via Cholesky decomposition. The paprameters λ and ν are crucial in generating the preconditioner with good condition number, which are in general tuned experimentally. The pseudo code of solver for (17) is shown in algorithm (1). After determining P, we solve $Py = PAx$ for sparse x, which stands for CT image, using TVAL3 as detailed in Sec 3.1.

Algorithm 1 Pseudo code for determining preconditioner from (17)

- Part A: On solving (21):
  1: Given initial points X_0, $r_0 = \nabla_X f(X_0)$, $p_0 = -r_0$, $0 < \epsilon \ll 1$ and $K \in \mathbb{Z}^+$. Set k = 0.
  2: Choose α by using backtracking line search [33]
  3: Update X :
  $$X_{k+1} = X_k + \alpha p_k$$
  stop if $\|X_{k+1} - X_k\|_2 < \epsilon$ or $k + 1 > K$. Otherwise go to next step.
  4: Update r_k and p_k

  $$\gamma_k = \frac{\|r_k\|_F}{\|p_k\|_F}$$
  $$r_{k+1} = \nabla_X f(X_{k+1})$$
  $$p_{k+1} = -r_{k+1} + \gamma_k p_k$$

  5: Set $k := k + 1$, return to step 2.

- Part B: On obtaining P:
  Find P such that $X = P^T P$ via Cholesky decomposition

Remark 4.1. *By setting λ and ν in (17) to 0, one obtains the closed form for the preconditioner as $(AA^T)^{-1}$, provided A has full rank. In particular, if $A = UDV^T$, the Singular Value Decomposition of A, then $(AA^T)^{-1} = (UDD^T U^T)^{-1}$. Consequently, the preconditioner in this case takes the form $(DD^T)^{-\frac{1}{2}} U^T$, which is denoted as P_{int}. For some suitable choice of η, the authors of [34] regularized the preconditioner as $(DD^T + \eta I)^{-\frac{1}{2}} U^T$, which is denoted as P_{reg}. As P_{int} and P_{reg} are square matrices such as P proposed in the current work, a comparison of performances obtained through P_{int}, P_{reg} and P seems justified.*

Remark 4.2. *By determining a rectangular preconditioner, $\tilde{P}$, of size $m \times M$, one may also formulate a system $\tilde{P}\tilde{y} = \tilde{P}\tilde{A}x$ where $\tilde{y}$ and $\tilde{A}$ are respectively the full set of Radon measurements and the associated sensing matrix as defined in section 2.1. Such a preconditioner $\tilde{P}$ merely acts as a dimentionality reduction operator by converting an M-vector to an m-vector. In a true sense it does not use any reduced data set as the computation of y from $\tilde{P}\tilde{y}$ in this approach, in general, requires the full set of Radon measurements. In the present work, however, the system $\tilde{y} = \tilde{A}x$ is initially recast as $y = Ax$, where $y = R\tilde{y}$ and $A = R\tilde{A}$, with R standing for the row-restricted identity matrix. This step is then followed by preconditioning. Consequently, the proposed method addresses compressed data acquisition problem in true sense.*

5 Experimental Results

Initially we generated the sensing matrix (A) of size $|\theta|n \times M$ for different subsets of angles θ, with $|\theta| = 20, 30, 40, 50, 60$, ($15.6\%, 23.4\%, 31.2\%, 39\%, 46.8\%$ of total data respectively), by drawing θ uniformly from the regular angular grid generated as in (6) and by taking the radial samples as per (5). Now, for each sensing matrix A (that is, for each subset of angles θ), we obtained a preconditioner P such that the coherence of the preconditioned matrix, PA, is smaller than that of A and the condition number of P is small, by solving the convex optimization problem (17). It can be noted that the construction of A (and hence P) depends only on the choice of radial and angular samples, and n (and not on the image data). Hence A and P may be designed via an off-line process.

5.1 Finding Preconditioner

As finding P involves an iterative process (detailed in Algorithm 1), we considered the initial guess for P as P_{reg} (which is defined in Remark 4.1) and obtained P by carrying out the steps in Algorithm 1. The reasons for choosing P_{reg} as the initial guess are (i) P almost coincides with P_{reg} when $\nu = \lambda = 0$, and (ii) P_{int} has a very bad condition number for all subsets of angles. With a view towards comparing the performance of our preconditioner P against P_{reg}, we report in Tables 1, 2 the condition numbers $\kappa(P_{reg})$, $\kappa(P)$ and associated coherences $\mu(P_{reg}A)$, $\mu(PA)$. From Tables 1, 2, it is clear that, against P_{reg}, P has small condition number along with small value for $\mu(PA)$. Experimentally we found the values $\eta = 10^{-5}$ (in generating P_{reg}), $\nu = 1$ and $\lambda = 50$ (in generating P) to be better choices among others. From Table 2, it can be stated that new preconditioner provides marginally suboptimal values for coherence. This is an expected behavior due to the fact that the improvement in condition number comes with an additional constraint as in Eq. 14, which marginally increases coherence. Overall, the results reported in tables show the promise that preconditioning offers towards the reconstruction in CT from reduced measurements. We carried out entire simulation work, reported in this paper, in MATALB (2014a) environment on a machine having 32 GB RAM and Intel Xeon processor with speed of 2.20 GHz employing a 64-bit Windows 10 operating system.

5.2 Image reconstruction

For experimental purpose we used two different test images, namely Shepp-Logan and Brain MRI images, each of size 128×128. Using TVAL3 solver, we reconstructed the underlying image x from (Py, PA), where y is the vector of Radon measurements generated for different sets of parameters as discussed at the beginning of this section. In order to evaluate the accuracy of reconstruction results quantitatively, we computed the Peak Signal to Noise Ratio (PSNR) and the structural similarity index (SSIM) between the original image and the

reconstructed image. The PSNR value [35] is computed as

$$PSNR = 10 * \log_{10} \frac{(\max(x^o))^2}{MSE},$$
(22)

where MSE is defined as

$$MSE = \frac{\sum_{i=1}^{m} \sum_{j=1}^{n} (x_{i,j}^o - x_{i,j})^2}{m \times n},$$
(23)

with $x_{i,j}^o$ and $x_{i,j}$ being the $(i,j)^{th}$ pixel values of the original and reconstructed images respectively. Higher PSNR values indicate better reconstruction. The structural similarity (SSIM) index [36] is another effective way of measuring the similarity between the original and reconstructed images. Suppose ρ and t are local image patches taken from the same location of two images being compared. The local SSIM index measures three similarities of the image patches: the similarity of luminance $l(\rho,t)$, the similarity of contrast $c(\rho,t)$, and the similarity of structures $s(\rho,t)$. The local SSIM is defined as

$$S(\rho,t) = [l(\rho,t)]^{k_1} \cdot [c(\rho,t)]^{k_2} \cdot [s(\rho,t)]^{k_3},$$

$$S(\rho,t) = \left(\frac{2\mu_\rho\mu_t + c_1}{\mu_\rho^2 + \mu_t^2 + c_1}\right)^{k_1} \left(\frac{2\sigma_\rho\sigma_t + c_2}{\sigma_\rho^2 + \sigma_t^2 + c_2}\right)^{k_2} \left(\frac{2\sigma_{\rho t} + c_3}{\sigma_\rho\sigma_t + c_3}\right)^{k_3},$$
(24)

where μ_ρ and μ_t are local means, σ_ρ and σ_t are local standard deviations, $\sigma_{\rho t}$ is cross-correlation after removing the means, c_1, c_2 and c_3 are the regularization constants for the luminance, contrast and structural terms and $k_1 > 0$, $k_2 > 0$ and $k_2 > 0$ are parameters used to adjust the relative importance of the three components. The SSIM score of the entire image is then computed by pooling the SSIM map, i.e. by simply averaging the SSIM map. Higher SSIM value indicates better quality in image reconstruction.

5.2.1 Reconstruction from noiseless measurements

To begin with, we report the results in the case where there is no noise in the measurements, that is, $Py = PAx$. The PSNR and SSIM values, reported in Table 3 both for initial sensing matrix A and preconditioned sensing matrix PA for different sets of Radon measurements, indicate improvement in reconstruction quality that preconditioning brings in. The images reconstructed through the original sensing matrix A and preconditioned sensing matrices $P_{reg}A$, PA are shown in Figure 3 respectively for $|\theta| = 20$, 40 and 60. From the results in Table 3 and Figure 3, we conclude that PA results in better performance than its counterparts $P_{reg}A$ and A. The corresponding results and the improvement in reconstruction for MRI test image (Figure 2(b)) are shown in Table 4 and Figure 4.

Tab. 1: The condition numbers of preconditioners as well as preconditioned sensing matrices for diferent subsets of angles. This table shows the improvement provided by the new preconditioner P.

$|\theta|$	Size of matrix	$\kappa(P_{reg})$	$\kappa(P)$	$\kappa(A)$	$\kappa(PA)$
20	2560×16384	91.1745	35.2728	91.1759	2.5848
30	3840×16384	119.1704	41.4987	119.1725	2.8717
40	5120×16384	2.3091e+04	5.0474e+03	5.1585e+15	1.1815e+13
50	6400×16384	2.5812e+04	5.3047e+03	9.9531e+14	1.2731e+13
60	7680×16384	2.8274e+04	5.5568e+03	1.1822e+15	1.2888e+13

Tab. 2: Coherences of the sensing matrices A, $P_{reg}A$ and PA of size $|\theta|n \times M$ for different subsets of angles θ with $|\theta| = 20$, 30, 40, 50, 60. The terms in the brackets of second column refer to compression provided by the Radon sensing matrix. This table indicates that the new preconditioner (P) provides marginally suboptimal value for coherence, compared to the values provided by P_{reg}. This expected behaviour is due to the additional constraint in (14), which improves the condition number significantly.

$|\theta|$	Size of Sensing matrix	$\mu(A)$	$\mu(P_{reg}A)$	$\mu(PA)$
20	2560×16384 (15.6%)	0.9645	0.9143	0.9260
30	3840×16384 (23.4%)	0.8817	0.7586	0.7879
40	5120×16384 (31.2%)	0.7986	0.7102	0.7168
50	6400×16384 (39%)	0.7544	0.5818	0.5929
60	7680×16384 (46.8%)	0.7769	0.6201	0.6310

5.2.2 Image reconstruction from noisy measurements

The noisy Radon measurements, denoted y_ϵ, were generated by considering the sensing matrix A and the original phantom image x such that $y_\epsilon = Ax + \epsilon$ where ϵ was taken as $\frac{\eta}{\|Ax\|_2}$ with η representing 2-norm of a vector of white Gaussian noise. In other words, as opposed to (y, A), the pair (y_ϵ, A) was used in the execution of the algorithm. The PSNR and SSIM values for Shepp-logan Phantom and Brain MRI images, reported in Tables 5, 6 for the sensing matrices A, $P_{reg}A$ and PA, indicate that the proposed preconditioning improves the reconstruction performance. The corresponding results and the improvement in reconstruction for Shepp-logan Phantom and Brain MRI test images (Figure 2) are shown in Tables 5, 6 and Figures 5, 6.

To summarize, despite the fall in coherence obtained via the preconditioning appearing not so significant, its impact on the reconstruction is significant in reducing reconstruction error. Consequently, for a given error tolerance, preconditioning of CT sensing matrix amounts to reducing the number of measurements. It may be reiterated that our approach remains focused on improving the sparse recovery properties of Radon sensing matrix as opposed to improving the performance of solvers. It may be possible to improve the performance of preconditioning based reconstruction further by adopting other solvers [7].
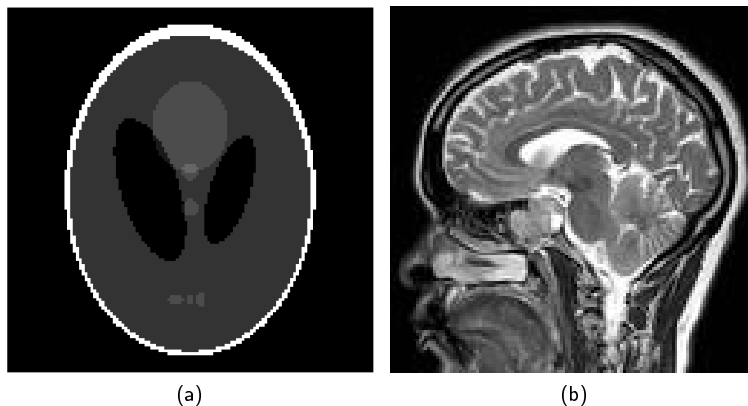
<center>(a)   (b)</center>

Fig. 2: Ground truth Shepp-Logan phantom and Brain images

Tab. 3: Reconstruction quality through PSNR and SSIM measures in noiseless case for the Phantom test image: The sensing matrices A, $P_{reg}A$ and PA are of size $|\theta|n \times M$. The PSNR and SSIM values are shown for $n = 128$ and $|\theta| = 20, 30, 40, 50$ and 60 angles. Here θ represents subset of angles used in reconstruction. The values reported in this table indicate the improvement obtained by the proposed preconditioning.

Size of sensing matrix	PSNR (dB) with			SSIM with		
	A	$P_{reg}A$	PA	A	$P_{reg}A$	PA
2560×16384 (15.6%)	26.6962	27.6218	28.9158	0.9629	0.9324	0.9767
3840×16384 (23.4%)	27.5491	29.9888	30.0952	0.9670	0.9761	0.9839
5120×16384 (31.2%)	30.8413	33.3227	34.2677	0.9865	0.9908	0.9940
6400×16384 (39%)	32.2684	43.0107	52.6897	0.9909	0.9972	0.9998
7680×16384 (46.8%)	32.2573	43.0505	62.4464	0.9907	0.9978	0.9999

Tab. 4: Reconstruction quality through PSNR and SSIM measures in noiseless case for the MRI test image: In line with Table 3, the values reported in this table indicate the improvement obtained by the proposed preconditioning.

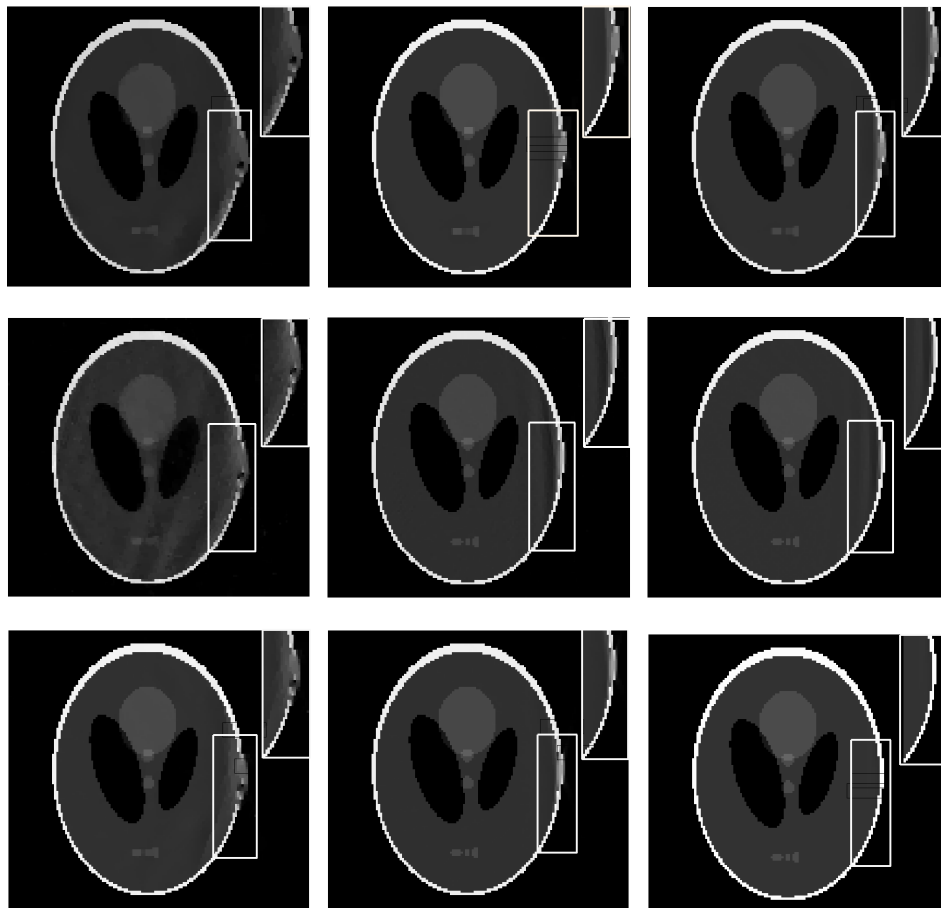| Size of sensing matrix	PSNR (dB) with			SSIM with		
	A	$P_{reg}A$	PA	A	$P_{reg}A$	PA
2560×16384 (15.6%)	17.0998	17.1289	17.5741	0.4825	0.4946	0.5253
3840×16384 (23.4%)	18.1166	18.7103	19.3386	0.5741	0.5936	0.6383
5120×16384 (31.2%)	19.9713	20.6244	21.7107	0.6793	0.7010	0.7588
6400×16384 (39%)	20.3146	21.9110	23.0272	0.7028	0.7633	0.8150
7680×16384 (46.8%)	20.9552	23.5582	24.6289	0.7395	0.8241	0.8639

Fig. 3: Reconstruction in noiseless case for the Phantom test image: The images on the first row correspond to the reconstruction obtained through A for 20 angles (left), 40 angles (middle) and 60 angles (right). The first row corresponding to the initial sensing matrix A and the second and third rows respectively stand for those obtained via $P_{reg}A$ and PA. This figure shows the improvement in reconstruction via proposed preconditioning. The small portions of images highlight the better reconstruction provided by the proposed preconditioner.
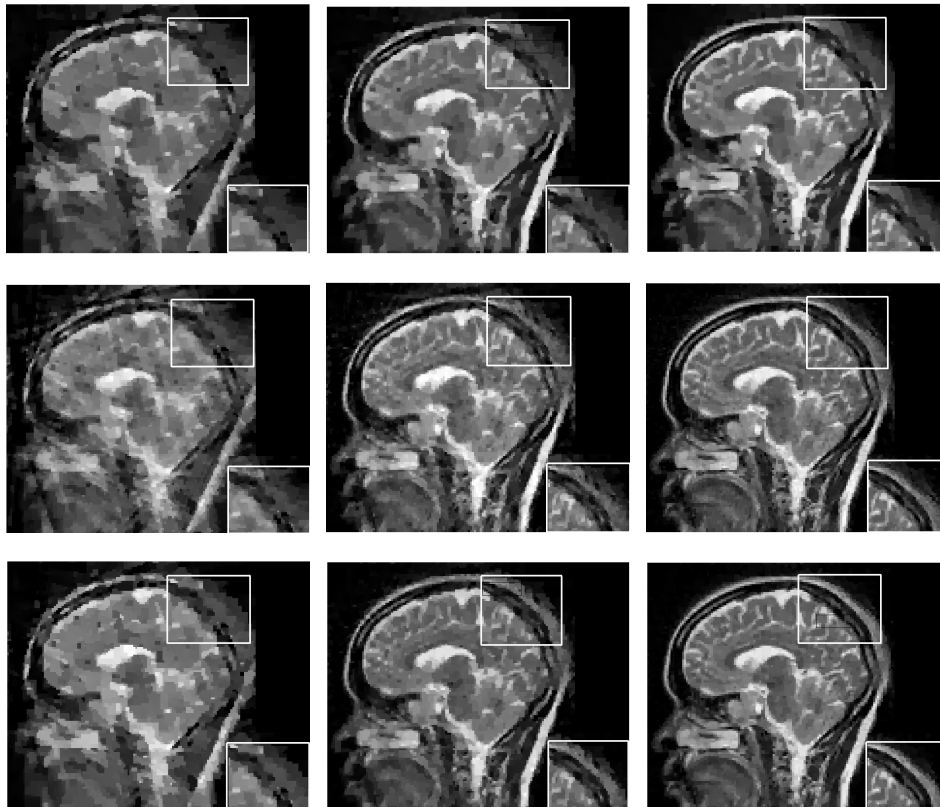
Fig. 4: Reconstruction in noiseless case for MRI test image: In line with Figure 3, this
figure shows the efficacy of proposed preconditioner on MRI image.

Tab. 5: Reconstruction quality through PSNR and SSIM measures in noisy case for
the Phantom test image: In line with Table 3, the values reported in this table
indicate the improvement in reconstruction obtained via P in noisy setting.

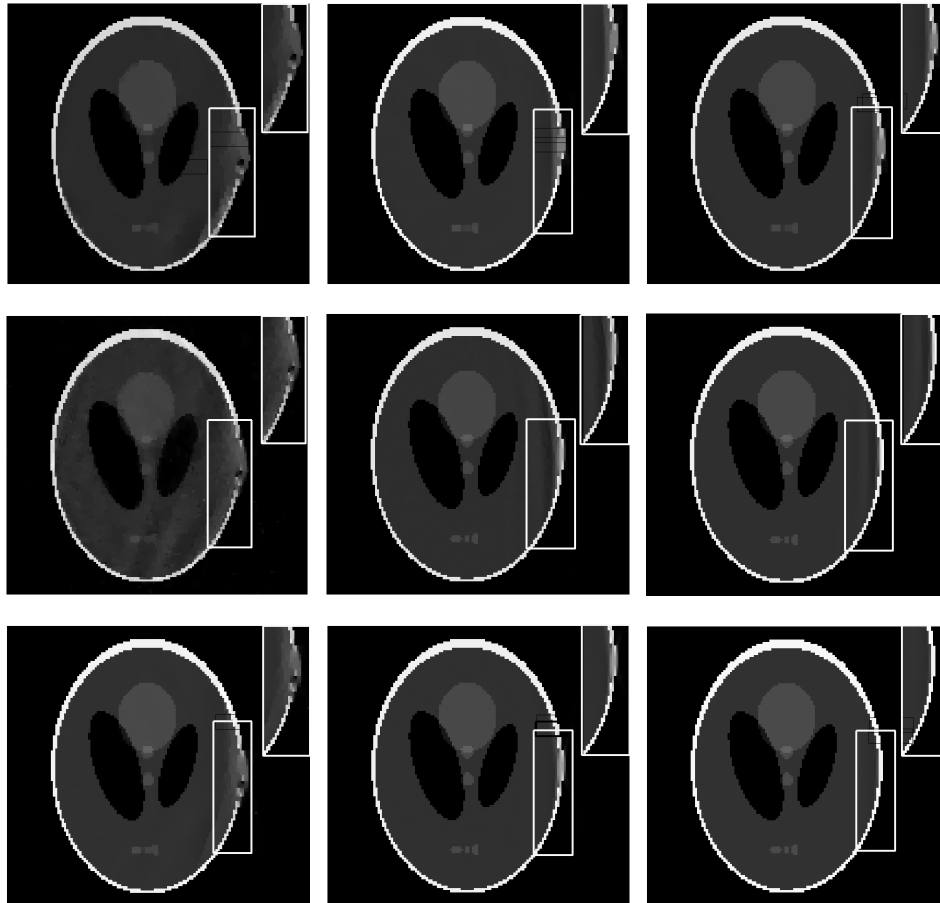| Size of sensing matrix	PSNR (dB) with			SSIM with		
	A	$P_{reg}A$	PA	A	$P_{reg}A$	PA
2560×16384 (15.6%)	26.5204	27.2283	28.8814	0.9592	0.9689	0.9745
3840×16384 (23.4%)	27.4991	29.6401	29.9838	0.9637	0.9722	0.9811
5120×16384 (31.2%)	30.8746	33.0376	34.0853	0.9855	0.9876	0.9923
6400×16384 (39%)	32.1912	41.1746	53.0983	0.9886	0.9953	0.9992
7680×16384 (46.8%)	32.1203	42.6366	50.3459	0.9897	0.9959	0.9989

Fig. 5: Reconstruction in noisy case for the Phantom test image: In line with Figure 3, the simulation results are shown in noisy setting with A, $P_{reg}A$ and PA for different sets of angles.

Tab. 6: Reconstruction quality through PSNR and SSIM measures in noisy case for the MRI test image

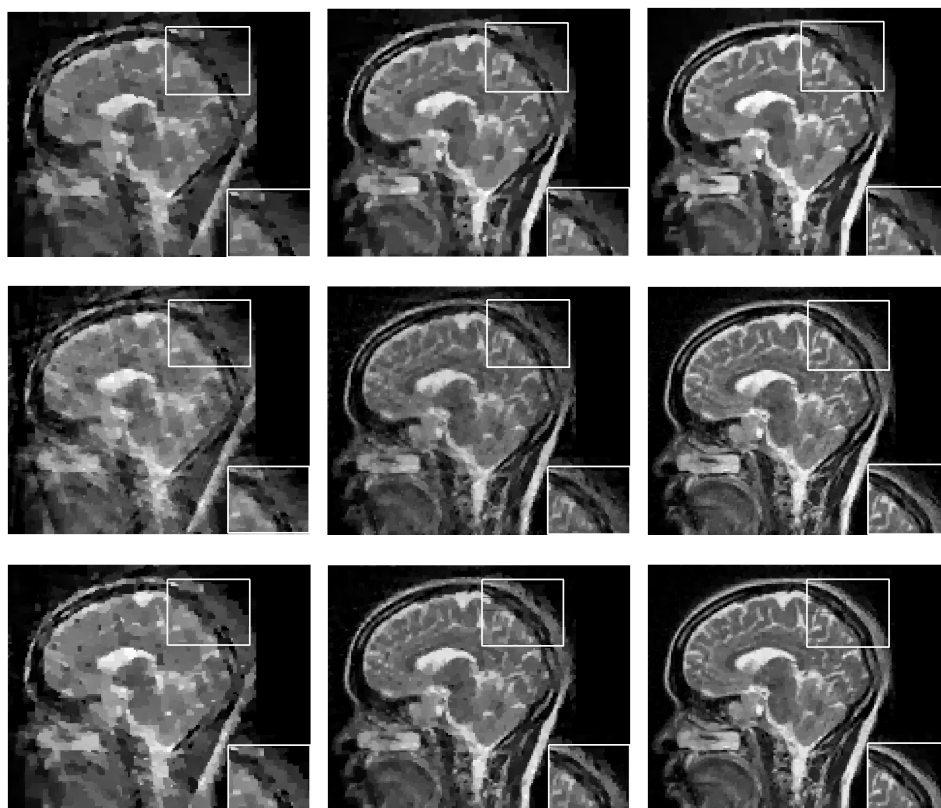| Size of sensing matrix	PSNR (dB) with			SSIM with		
	A	$P_{reg}A$	PA	A	$P_{reg}A$	PA
2560×16384 (15.6%)	17.1130	17.2274	17.5686	0.4806	0.4958	0.5255
3840×16384 (23.4%)	18.0831	18.9329	19.3256	0.5711	0.5650	0.6370
5120×16384 (31.2%)	19.9527	20.6254	21.6984	0.6786	0.7009	0.7570
6400×16384 (39%)	20.3004	21.9110	23.0139	0.7017	0.7633	0.8139
7680×16384 (46.8%)	20.9543	23.5582	24.6467	0.7392	0.8241	0.8641

Fig. 6: Reconstruction in noisy case for MRI test image: In line with Figure 5, the simulation results are shown in noisy setting with A, $P_{reg}A$ and PA for different sets of angles.

6 Conclusion

In this paper, we proposed a new method for finding a square and invertible preconditioner that reduces the coherence of Randon sensing matrix while improving its conditioning. The simulation results, discussed through the measures such as PSNR and SSIM, indicate that the proposed preconditioning method improves reconstruction quality for different data sets of reduced sizes.

References

[1] D. P. Frush, L. F. Donnelly, N. S. Rosen, Computed tomography and radiation risks: What pediatric health care providers should know, Pediatrics 112 (2003) 951–957.

[2] C. S. Sastry, P. C. Das, Wavelet based multilevel backprojection algorithm for parallel and fan beam scanning geometries, International Journal of Wavelets, Multiresolution and Information Processing 4 (3) (2006) 523 – 545.

[3] J. Jan, Medical Image Processing, Reconstruction and Restoration: Concepts and Methods, CRC Press, 2005.

[4] X. C. Pan, E. Y. Sidky, M. Vannier, Why do commercial ct scanners still employ traditional, filtered back-projection for image reconstruction?, Inverse Problems 25 (12) (2009) 1230009.

[5] M. Beister, D. Kolditz, W. A. Kalender, Iterative reconstruction methods in x-ray ct, Europian Journal of Medical Physics 28 (2012) 94–108.

[6] J. S. Jorgensen, P. C. Hansen, S. Schmidt, Sparse image reconstruction in computed tomography, Kgs. Lyngby: Technical University of Denmark, PHD-2013 (293).

[7] C. Chen, G. Xu, A new linearized split bregman iterative algorithm for image reconstruction in sparse view x-ray computed tomography, Computers and Mathematics with Applications 71 (8) (2016) 1537 – 1559.

[8] M. Elad, Sparse and Redundant Representations: from theory to applications in signal processing, Springer, 2010.

[9] S. Foucart, H. Rauhut, A Mathematical Introduction to Compressive Sensing, Birkhuser, 2013.

[10] M. Elad, Optimized projections for compressed sensing, IEEE Trans. on Signal Processing 55 (12) (2007) 5695–5702.

[11] M. Lustig, D. Donoho, J. Pauly, Sparse mri: The application of compressed sensing for rapid mr imaging, Magnetic resonance in medicine 58 (2007) 1182–1195.

[12] J. M. Duarte-Carvajalino, G. Sapiro, Learning to sense sparse signals: Simultaneous sensing matrix and sparsifying dictionary optimization, IEEE Trans. Image Proc. 18 (7) (2009) 1395–1408.

[13] J. Xu, Y. Pi, , Z. Cao, Optimized projection matrix for compressive sensing, EURASIP Journal on advances in signal processing (2010) 1–8.

[14] A. Vahid, S. Ferdowsi, S. Sanei, A gradient-based alternating minimization approach for optimization of the measurement matrix in compressive sensing, Signal Processing 92 (2012) 999–1009.

[15] W. Lu, T. Hinamoto, Design of projection matrix for compressive sensing by non-smooth optimization, IEEE International Symposium on Circuits and Systems (ISCAS) (2014) 1279–1282.

[16] F. Natterer, The Mathematics of Computerized Tomography, Society for Industrial and Applied Mathematics, 2001.

[17] B. Sahiner, A. E. Yagle, Limited angle tomography using wavelets, Nuclear Science Symposium and Medical Imaging Conference, (1993) 1912–1916.

[18] M. Rantala, S. Vanska, S. Jarvenpaa, M. Kalke, M. Lassas, J. Moberg, S. Siltanen, Wavelet-based reconstruction for limited-angle x-ray tomography, IEEE Tran. on Medical Imaging 25 (2) (2006) 210–217.

[19] E. Sidky, X. Pan, Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization, Phys. Med. Biol. 53 (17) (2008) 4777.

[20] L. Ritschl, F. Bergner, C. Fleischmann, M. Kachelrie, Improved total variation-based ct image reconstruction applied to clinical data, Physics in Medicine and Biology 56 (6) (2011) 1545 − 1561.

[21] T. Prasad, P. U. P. Kumar, C. S. Sastry, P. V. Jampana, Reconstruction of sparse-view tomography via banded matrices, In: Mukherjee S. et al. (eds) Computer Vision, Graphics, and Image Processing. ICVGIP 2016. Lecture Notes in Computer Science, Springer, Cham. 10481 (2017) 204–215.

[22] A. C. Kak, M. Slaney, Principles of Computerized Tomographic Imaging, Society for Industrial and Applied Mathematics, 2001.

[23] J. A. Tropp, Greed is good: Algorithmic results for sparse approximation, IEEE Trans. Inform. Theory 50 (10) (2004) 2231–2242.

[24] C. Li, W. Yin, H. Jiang, Y. Zhang, Improved total variation-based ct image reconstruction applied to clinical data, Physics in Medicine and Biology 56 (6) (2011) 1545 − 1561.

[25] D. W. Peaceman, H. H. Rachford, The numerical solution of parabolic and elliptic differential equations, Journal of the Society for Industrial and Applied Mathematics 3 (1955) 28–41.

[26] C. Li, W. Yin, H. Jiang, Y. Zhang, An efficient augmented lagrangian method with applications to total variation minimization, Computational Optimization and Applications 56 (3) (2013) 507–530.

[27] B. T. Kelley, V. K. Madisetti, The fast discrete radon transform-i: Theory, IEEE Trans.on Image Proc. 2 (3) (1993) 382–400.

[28] Z. Lu, T. K. Pong, Minimizing condition number via convex programming, SIAM J. Matrix Anal. Appl. 32 (2011) 1193–1211.

[29] C. Wang, D. Sun, K.-C. toh, Solving log-determinant optimization problems by a newton-cg primal proximal point algorithm, SIAM J. Optim. 20 (6) (2010) 2994–3013.

[30] S. Ravishankar, Y. Bresler, Learning sparsifying transforms, IEEE Trans. on Signal Proc. 61 (5) (2013) 1072–1086.

[31] D. S. Bernstein, Matrix Mathematics, Princeton University Press, 2005.

[32] R. Pytlak, Conjugate Gradient Algorithms in Nonconvex Optimization, Springer, 2009.

[33] A. Beck, Introduction to nonlinear Optimization Theory, Algorithms and Applications with MATLAB, SIAM, Philadelphia, 2014.

[34] A. Jin, B. Yazici, V. Ntziachristos, Light illumination and detection patterns for fluorescence diffuse optical tomography based on compressive sensing, IEEE Trans. on Image Proc. 23 (6) (2014) 2609–2624.

[35] J. Moreno, B. Jaime, S. Saucedo, Towards no-reference of peak signal to noise ratio, International Journal of Advanced Computer Science and Applications (IJACSA) 4 (1) (2013) 123–130.

[36] Z. Wang, A. C. Bovik, H. Sheikh, E. P. Simoncelli, Image quality assessment: From error visibility to structural similarity, IEEE Trans. on Image Processing 13 (4) (2004) 600–612.