Optimized Sensing Matrix Design Based on Parseval Tight Frame and Matrix Decomposition

Aiping Yang*, Jinxia Zhang, and Zhengxin Hou
School of Electronic Information Engineering, Tianjin University, Tianjin 300072, China
Email: yangaiping@tju.edu.cn; 15822977075@163.com; zhengxinhou@163.com

Abstract—Recent efforts have shown that the reconstruction performance could be improved with optimized sensing matrix according to a given dictionary for a compressed sensing (CS) system. The existed optimizing conditions are mainly used to address the worst-case performance of CS recovery. Considering the quality of a sensing matrix with respect to the mean squared error (MSE) performance of the Oracle estimator, Chen *et al.* proposed the sensing matrix based on Parseval tight frame, which exhibits superior performance in relation to other existed designs. However, the equivalent sensing matrix under this design framework couldn't achieve the optimal mutual coherence. In light of the matrix decomposition theory, the bigger the smallest singular value, the stronger non-correlation of the columns of the matrix have. We further optimize the sensing matrix combining with the matrix decomposition theory, so as to achieve the optimal statistical reconstruction and the optimal mutual coherence performance at the same time. Through the approximate QR decomposition and the mean singular value decomposition (SVD), we adjust the singular values of the sensing matrix, so as to reduce the correlation of the matrix. A great number of experiments show that the proposed optimized sensing matrix realizes the minimum of the reconstructed error compared to other designs in the literature with different sparse recovery algorithms.

Index Terms—compressed sensing, sensing matrix, MSE, Parseval tight frame, QR decomposition, SVD decomposition.

I. INTRODUCTION

Compressed sensing (CS) theory [1] has received much attention in recent years which states that under sparse conditions, the signal can be sampled via non-adaptive linear projections and meanwhile still keeping the original structure. Then the original high-dimensional signals can be reconstructed exactly from these projections whose dimension is much lower by solving the sparse-constraint optimization problems. As the key step of the compressed sensing system, it is of great importance to construct reasonable and effective sensing matrix. From the restricted isometry property (RIP) [2] and the mutual coherence theory [3], many sensing methods have been proposed [4]-[6]. In 2006, Elad *et al.* pointed out that CS reconstruction accuracy could be improved with optimized sensing matrices according to a given dictionary comparing to the non-adaptive random matrices. Based on this point, Elad *et al*. [7] put forward the concept of t-averaged mutual coherence firstly and then proposed an iterative algorithm to optimize the sensing matrix by minimizing the t-averaged mutual coherence. Elad's method improved the reconstruction performance, but as iterative, it is time-consuming. Based on Elad's research, Duarte-Carvajalino and Sapiro [8] addressed the problem by making any subset columns of the equivalent sensing matrix as orthogonal as possible, or equivalently, making the Gram matrix as closely as possible to identity matrix. Then they introduced an algorithm to iteratively optimize both the sensing matrix and the over-complete dictionary simultaneously. The reconstructed performance of this method is not ideal as the equivalent sensing matrix is over-complete rather than orthogonal. Further, Xu *et al.* considered the equiangular tight frame (ETF) as their target design and proposed an iterative algorithm to make the sensing matrix approach that design [9]. The restriction on the dimensions of the equiangular tight frame greatly limits this method's application in compressed sensing.

All these above optimized designs constrain the conditions that the sensing matrix needs to satisfy by taking the worst-case performance of sparse recovery as the target. However, the actual reconstruction performance is often much better than the worst-case, so that this viewpoint can be too conservative. In addition, the above designs are all inherently based on the mutual coherence which is difficult to operate and they are all complex iterative algorithms.

Considering the statistics of the CS process, it's of great practical importance to take the good expect-case recovery performance as the design target. In [10], from the statistical significance, Chen *et al.* demonstrated that the good equivalent sensing matrix should be a Parseval tight frame by capitalizing on the mean squared error (MSE) of the Oracle estimator whose performance has been shown to act as a benchmark to the performance of various common sparse recovery algorithms. However, the equivalent sensing matrix which is frame-based doesn't achieve the optimal mutual coherence performance. Researches have found that, the linear correlation of the matrix is closed with its singular values [11]. Then through in-depth study of the sensing matrix design method based on Parseval tight frame, we further reduce the mutual coherence between the sensing matrix and the sparsifying matrix by adjusting the singular values of the sensing matrix through matrix

Manuscript received June 17, 2013; revised July 15, 2013.
This work was supported by the National Natural Science Foundation of China under Grant No. 61002027.
Corresponding author email: yangaiping@tju.edu.cn.
doi:10.12720/jcm.8.7.456-462

decomposition to achieve both the optimal reconstruction and the optimal non-coherence performance.

The remainder of this paper is organized as follows. We mainly describe the CS theory in Section II. Section III present the design method of statistical compressed sensing matrix by capitalizing on the MSE of the Oracle estimator whose performance has been shown to act as a benchmark to the performance of standard sparse recovery algorithms. In Section IV, we introduce the relationship between the singular values of a matrix and its correlation. Further we put forth the matrix decomposition methods for reducing the correlation and proved the related theories, including mean singular value based SVD decomposition and approximate QR decomposition. In Section V, the sensing matrix is optimized by combining the Parseval tight frame-based method with the matrix decomposition theory mentioned in section IV. By adjusting the singular values, we further reduce the correlation and propose the sensing matrix design method which can achieve the optimal statistical reconstruction performance and optimal non-coherence. Section VI are the experiments, the proposed method is compared with other state-of-the-art methods. Finally, concluding remarks are presented in Section VII.

II. Compressed Sensing Theory

The sampling model of compressed sensing is

$$\boldsymbol{y} = \boldsymbol{\Phi f} + \boldsymbol{n} \tag{1}$$

where $\boldsymbol{y} \in R^m$ is the measurement signal vector, $\boldsymbol{f} \in R^n$ is the original signal vector, $\boldsymbol{\Phi} \in R^{m\times n} (m \le n)$ is the sensing matrix, $\boldsymbol{n} \sim \mathrm{N}(0,\sigma^2 \boldsymbol{I}_m)$ is a zero-mean white Gaussian noise vector with variance σ^2.We assume that the original signal is sparse in some basis, i.e.,

$$\boldsymbol{f} = \boldsymbol{\Psi x} \tag{2}$$

where $\boldsymbol{\Psi} \in R^{n\times \hat{n}} (n \le \hat{n})$ is a matrix that represents the sparsifying basis, e.g., an orthonormal or over-complete dictionary [12], and $\boldsymbol{x}$ is a sparse representation vector of f , i.e., $\|\boldsymbol{x}\|_{l_0} << n$ .Then we can rewrite (1) as

$$y = \boldsymbol{\Phi\Psi} x + n = Ax + n \tag{3}$$

where $\boldsymbol{A} = \boldsymbol{\Phi\Psi} \in R^{m\times \hat{n}}$ is called the equivalent sensing matrix.

To recover the sparse signal representation $\boldsymbol{x}$ from the measurement $\boldsymbol{y}$, one can resort to the l_1 norm constrained optimization problems:

$$\hat{\boldsymbol{x}} = \arg\min \|\boldsymbol{x}\|_1 \quad s.t. \|\boldsymbol{Ax} - \boldsymbol{y}\|_2 \le \varepsilon \tag{4}$$

by solving (4), we can get $\hat{\boldsymbol{x}}$ and further the reconstructed original signal by $\hat{\boldsymbol{f}} = \boldsymbol{\Psi}\hat{\boldsymbol{x}}$ exactly. In (4), ε is an estimate of the noise level. The Eq. (4) is also known as the basis pursuit de-noise (BPDN) problem.

It has been established that the well-known RIP [1] provides a guarantee for exact or near exact recovery of a sparse signal representation $\boldsymbol{x}$ from the measurement vector $\boldsymbol{y}$ via the l_1 minimization in (4). Note that the RIP is a sufficient but not necessary condition for successful reconstruction and it may be too strict. It's also too difficult to use the RIP property to guide the design of sensing matrix in practice. Another way to evaluate a sensing matrix is the mutual coherence between the sensing matrix $\boldsymbol{\Phi}$ and the sparsifying matrix $\boldsymbol{\Psi}$ which is defined as below:

$$\mu(\boldsymbol{\Phi},\boldsymbol{\Psi}) = \sqrt{n} \max_{1\le k,j\le n} \left|\left\langle \boldsymbol{\varphi}_k, \boldsymbol{\psi}_j \right\rangle\right| \tag{5}$$

The smaller μ, the greater probability that $\boldsymbol{\Phi}$ satisfies RIP [2]. As mentioned above, Elad's method [7], Sapiro's method [8], and Xu *et al.*'s method [9] are all inherently mutual coherence based approaches.

III. The Sensing Matrix Design Method Based on Parseval Tight Frame

From the viewpoint of statistics, Chen deduced the conditions that the sensing matrix should satisfy by taking the good expected-case reconstruction performance as the target. Specifically, the goal of the sensing matrix design relates to the minimization of the MSE in estimating a sparse random vector x corrupted by a random Gaussian noise vector n from the measurement y , given by:

$$\mathrm{MSE}(\boldsymbol{\Phi}) = \mathrm{E}_{\boldsymbol{x},\boldsymbol{n}}(\|\mathrm{F}(\boldsymbol{\Phi\Psi}\, \boldsymbol{x} + \boldsymbol{n}) - \boldsymbol{x}\|_2^2) \tag{6}$$

where $F(\cdot)$ denotes a specified estimator, here corresponding to sparse recovery algorithms such as the basis pursuit de-noise (BPDN) and orthogonal matching pursuit (OMP), etc. $\mathrm{E}_{\boldsymbol{x},\boldsymbol{n}}(\cdot)$ denotes the expectation with respect to the joint distribution of the random signal vector $\boldsymbol{x}$ and the noise vector $\boldsymbol{n}$. More representatively, choosing the oracle MSE, which represents the best achievable performance for any unbiased estimator, as a benchmark to the performance of various sparse recovery algorithms. Accordingly, the oracle estimator MSE incurred in the estimation of a sparse deterministic vector $\boldsymbol{x}$ in the presence of a standard Gaussian noise vector $\boldsymbol{n}$, according to the model in (1), is given by:

$$\begin{aligned} &\mathrm{MSE}_{\boldsymbol{n}}^{\mathrm{oracle}}(\boldsymbol{A},\boldsymbol{x}) \\ &= \mathrm{E}_{\boldsymbol{n}}(\|\mathrm{F}^{\mathrm{oracle}}(\boldsymbol{Ax}+\boldsymbol{n}) - \boldsymbol{x}\|_2^2) \\ &= \sigma^2 \mathrm{Tr}((\boldsymbol{E}_J^{\mathrm{T}} \boldsymbol{A}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{E}_J)^{-1}) \end{aligned} \tag{7}$$

where $\mathrm{E}_{\boldsymbol{n}}(\cdot)$ denotes expectation with respect to the distribution of the random vector $\boldsymbol{n}$, $\mathrm{Tr}(\cdot)$ denotes the trace of a matrix and $\boldsymbol{E}_J$ denotes the matrix that results from the identity matrix by deleting the set of columns out of the support J .

Consequently, the average value of the oracle MSE is given by

$$\begin{aligned}\mathrm{MSE}^{\mathrm{oracle}}(\boldsymbol{\Phi}) &= \mathrm{E}_{\boldsymbol{x}}(\mathrm{MSE}_{\boldsymbol{n}}^{oracle}(\boldsymbol{\Phi}, \boldsymbol{x}) \\ &= \sigma^2 \mathrm{E}_J(\mathrm{Tr}((\boldsymbol{E}_{\boldsymbol{J}}^{\mathrm{T}} \boldsymbol{D}^{\mathrm{T}} \boldsymbol{\Phi}^{\mathrm{T}} \boldsymbol{\Phi} \boldsymbol{D} \boldsymbol{E}_{\boldsymbol{J}})^{-1}))\end{aligned} \tag{8}$$

where $\mathrm{E}_{\boldsymbol{x}}(\cdot)$ denote the expectation with respect to the distribution of the random vector $\boldsymbol{x}$. In order to obtain the optimal reconstruction performance, the MSE above should be as small as possible. Then posing the optimization problem:

$$\begin{aligned}&\min_{\boldsymbol{Q}} \mathrm{E}_J(\mathrm{Tr}((\boldsymbol{E}_J^{\mathrm{T}} \boldsymbol{Q} \boldsymbol{E}_J)^{-1})) \\ &\mathrm{s.t.} \boldsymbol{Q} \succ 0, \mathrm{Tr}(\boldsymbol{Q}) = m, \mathrm{rank}(\boldsymbol{Q}) \le m\end{aligned} \tag{9}$$

In (9), $\boldsymbol{Q} = \boldsymbol{A}^{\mathrm{T}} \boldsymbol{A}$ is the coherence matrix of the equivalent sensing matrix $\boldsymbol{A}$, $\boldsymbol{Q} \succ 0$ means that the matrix $\boldsymbol{Q}$ is positive semi-defined, $\mathrm{E}_J(\cdot)$ denotes expectation with respect to the random support J, $\mathrm{Tr}(\cdot)$ and $\mathrm{rank}(\cdot)$ stand for the trace and the rank of a matrix, respectively.

The optimization problem (9) is solved by considering the closest convex-relaxation problem by ignoring the rank constraint in (9) .i.e.,

$$\min_{\boldsymbol{Q}} \mathrm{E_J}(\mathrm{Tr}((\boldsymbol{E}_{\mathrm{J}}^{\mathrm{T}} \boldsymbol{Q} \boldsymbol{E}_{\mathrm{J}})^{-1})) \quad \mathrm{s.t.} \quad \boldsymbol{Q} \succ 0, \mathrm{Tr}(\boldsymbol{Q}) = m \tag{10}$$

It has been proven that the solution to (10) is a $\hat{n} \times \hat{n}$ matrix $(m/\hat{n})\boldsymbol{I}_{\hat{n}}$, $\boldsymbol{I}_{\hat{n}}$ denotes the $\hat{n} \times \hat{n}$ identity matrix. Obviously, $(m/\hat{n})\boldsymbol{I}_{\hat{n}}$ is not the feasible solution to the original optimization problem (9), because $\mathrm{rank}((m/\hat{n})\boldsymbol{I}_{\hat{n}}) = \hat{n} \ge m$. Therefore， The $m \times \hat{n}$ matrix $\boldsymbol{A}$ whose coherence matrix $\boldsymbol{Q} = \boldsymbol{A}^{\mathrm{T}} \boldsymbol{A}$ is closest to the matrix $(m/\hat{n})\boldsymbol{I}_{\hat{n}}$ can be used. i.e.,

$$\min_{\boldsymbol{A}} \left\| \boldsymbol{A}^T \boldsymbol{A} - \frac{m}{\hat{n}} \boldsymbol{I}_{\hat{n}} \right\|_{\mathrm{F}}^2 \quad \mathrm{s.t.} \quad \mathrm{Tr}(\boldsymbol{A}^T \boldsymbol{A}) = m \tag{11}$$

It can be proven that the solution to (11) is the $m \times \hat{n}$ Parseval tight frame [13].

Based on the above analysis, good equivalent sensing matrices ought to be close to a Parseval tight frame [13]. In addition, in order to achieve good sensing performance, the sensing energy cost should be as small as possible. Then getting the following optimization problem:

$$\min_{\hat{\boldsymbol{\Phi}}} \|\hat{\boldsymbol{\Phi}}\|_{\mathrm{F}}^2 \quad \mathrm{s.t.} \quad \hat{\boldsymbol{\Phi}} \boldsymbol{\Psi} \boldsymbol{\Psi}^{\mathrm{T}} \hat{\boldsymbol{\Phi}}^{\mathrm{T}} = \boldsymbol{I}_m \tag{12}$$

The solution to (12) is

$$\hat{\boldsymbol{\Phi}} = \boldsymbol{U}_{\hat{\boldsymbol{\Phi}}} \boldsymbol{\Lambda}_{\hat{\boldsymbol{\Phi}}} \boldsymbol{J}_{\boldsymbol{n}} \boldsymbol{U}_{\boldsymbol{\Psi}}^{\mathrm{T}} \tag{13}$$

where $\boldsymbol{U}_{\hat{\boldsymbol{\Phi}}}$ is an arbitrary orthonormal matrix and $\Lambda_{\hat{\boldsymbol{\Phi}}} = \left[\mathrm{Diag}\left(1/\lambda_m^{\boldsymbol{\Psi}}, 1/\lambda_{m-1}^{\boldsymbol{\Psi}}, \cdots, 1/\lambda_1^{\boldsymbol{\Psi}}\right) O_{m \times (n-m)} \right]$ is a matrix whose main diagonal entries are the singular values of $\boldsymbol{\Phi}$, among it $\lambda_1^{\boldsymbol{\Psi}}, \cdots, \lambda_m^{\boldsymbol{\Psi}}$ are the m singular values of $\boldsymbol{\Psi}$. Unlike the design methods in [7]-[9], the proposed sensing matrix has closed form instead of iterative and it can reach the minimum MSE for image recovery.

IV. Reduce the Correlation of a Matrix Based on the Matrix Decomposition

A. The Approximate QR Decomposition

From the matrix decomposition theory, the linear correlation of a matrix is closely related with its smallest singular value [11]. If the smallest singular value is bigger, then it can result in more non-correlation. The approximate QR decomposition can increase the smallest singular value and narrow the range of singular values of a matrix.

Theorem 1: After making standard QR decomposition to the matrix $\boldsymbol{\Phi}$, we get an upper triangular matrix $\boldsymbol{R}$ and a square matrix $\boldsymbol{Q}$, i.e, $\boldsymbol{\Phi} = \boldsymbol{Q}\boldsymbol{R}$. Then we get a new matrix $\hat{\boldsymbol{R}}$ by keeping the main diagonal elements of $\boldsymbol{R}$ unchanged and setting all the other elements to zero, further getting $\tilde{\boldsymbol{\Phi}} = \boldsymbol{Q}\hat{\boldsymbol{R}}$. So the smallest singular value of $\tilde{\boldsymbol{\Phi}}$ is bigger than that of $\boldsymbol{\Phi}$, and the biggest singular value of $\tilde{\boldsymbol{\Phi}}$ is smaller than that of $\boldsymbol{\Phi}$.

Proof: Firstly, the smallest singular value of $\tilde{\boldsymbol{\Phi}}$ is,

$$\begin{aligned}\sigma_{\min}(\tilde{\boldsymbol{\Phi}}) &= \sqrt{\lambda_{\min}(\tilde{\boldsymbol{\Phi}}\tilde{\boldsymbol{\Phi}}^{\mathrm{T}})} = \sqrt{\lambda_{\min}(\boldsymbol{R}\boldsymbol{R}^{\mathrm{T}})} \\ &= \sqrt{\frac{\hat{\boldsymbol{v}}^{\mathrm{T}} \boldsymbol{R}\boldsymbol{R}^{\mathrm{T}} \hat{\boldsymbol{v}}}{\hat{\boldsymbol{v}}^{\mathrm{T}} \hat{\boldsymbol{v}}}} \ge \sqrt{\min_{\boldsymbol{v}} \frac{\boldsymbol{v}^{\mathrm{T}} \boldsymbol{R}\boldsymbol{R}^{\mathrm{T}} \boldsymbol{v}}{\boldsymbol{v}^{\mathrm{T}} \boldsymbol{v}}} \\ &= \sqrt{\min_{\boldsymbol{v}} \frac{\boldsymbol{v}^{\mathrm{T}} \boldsymbol{\Phi}\boldsymbol{\Phi}^{\mathrm{T}} \boldsymbol{v}}{\boldsymbol{v}^{\mathrm{T}} \boldsymbol{v}}} = \sqrt{\lambda_{\min}(\boldsymbol{\Phi} \cdot \boldsymbol{\Phi}^{\mathrm{T}})} \\ &= \sigma_{\min}(\boldsymbol{\Phi})\end{aligned} \tag{14}$$

On the other hand, the biggest singular value of $\tilde{\Phi}$ is,

$$\begin{aligned}\sigma_{\max}(\tilde{\boldsymbol{\Phi}}) &= \sqrt{\lambda_{\max}(\tilde{\boldsymbol{\Phi}}\tilde{\boldsymbol{\Phi}}^{\mathrm{T}})} = \sqrt{\lambda_{\max}(\boldsymbol{R}\boldsymbol{R}^{\mathrm{T}})} \\ &= \sqrt{\frac{\hat{\boldsymbol{v}}^{\mathrm{T}} \boldsymbol{R}\boldsymbol{R}^{\mathrm{T}} \hat{\boldsymbol{v}}}{\hat{\boldsymbol{v}}^{\mathrm{T}} \hat{\boldsymbol{v}}}} \le \sqrt{\max_{\boldsymbol{v}} \frac{\boldsymbol{v}^{\mathrm{T}} \boldsymbol{R}\boldsymbol{R}^{\mathrm{T}} \boldsymbol{v}}{\boldsymbol{v}^{\mathrm{T}} \boldsymbol{v}}} \\ &= \sqrt{\max_{\boldsymbol{v}} \frac{\boldsymbol{v}^{\mathrm{T}} \boldsymbol{\Phi}\boldsymbol{\Phi}^{\mathrm{T}} \boldsymbol{v}}{\boldsymbol{v}^{\mathrm{T}} \boldsymbol{v}}} = \sqrt{\lambda_{\max}(\boldsymbol{\Phi} \cdot \boldsymbol{\Phi}^{\mathrm{T}})} \\ &= \sigma_{\max}(\boldsymbol{\Phi})\end{aligned} \tag{15}$$

where $v, \hat{v}$ are column vectors, they corresponding to the diagonal of matrix $\boldsymbol{R}$ after making the smallest and the biggest element to be 1, and the other elements all to 0, respectively.

B. The Mean SVD Decomposition Theory

Assume that $\boldsymbol{A}$ is a $m \times n$ singular matrix, the singular value decomposition of $\boldsymbol{A}$ is: $\boldsymbol{A} = \boldsymbol{U} \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} \boldsymbol{V}^{\mathrm{T}}$, the generalized inverse matrix of $\boldsymbol{A}$ is $\boldsymbol{A}^{+} = \boldsymbol{V} \begin{bmatrix} \Sigma^{-1} & 0 \\ 0 & 0 \end{bmatrix} \boldsymbol{U}^{\mathrm{T}}$, where $\boldsymbol{U}$ and $\boldsymbol{V}$ are $m \times m$, $n \times n$ unit orthogonal matrices, respectively. $\Sigma = diag(\sigma_1, \sigma_2, \cdots, \sigma_m)$, $\sigma_1 \geq \sigma_2 \cdots \geq \sigma_m$ are the singular values of $\boldsymbol{A}$.

Theorem 2: Assume that vectors $\boldsymbol{x}_1$, $\boldsymbol{x}_2$ satisfying $\boldsymbol{A}\boldsymbol{x}_1 = \boldsymbol{y}_1$, $\boldsymbol{A}\boldsymbol{x}_2 = \boldsymbol{y}_2$, respectively, after making $d_1 = \|\boldsymbol{x}_1 - \boldsymbol{x}_2\|_p / \|\boldsymbol{x}_1\|_p$, $d_2 = \|\boldsymbol{y}_1 - \boldsymbol{y}_2\|_p / \|\boldsymbol{y}_1\|_p$, then $d_1 / d_2 \leq k(\boldsymbol{A})$, where $k(\boldsymbol{A}) = \|\boldsymbol{A}\|_p * \|\boldsymbol{A}^+\|_p$.

Proof: Because $\boldsymbol{x}_1 - \boldsymbol{x}_2 = \boldsymbol{A}^+(\boldsymbol{y}_1 - \boldsymbol{y}_2)$

$$\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|_p = \|\boldsymbol{A}^+(\boldsymbol{y}_1 - \boldsymbol{y}_2)\|_p \leq \|\boldsymbol{A}^+\|_p \|\boldsymbol{y}_1 - \boldsymbol{y}_2\|_p$$

and so

$$\|\boldsymbol{y}_1\|_p \leq \|\boldsymbol{A}\|_p * \|\boldsymbol{x}_1\|_p ,$$

than is

$$\|\boldsymbol{x}_1\|_p \geq \frac{\|\boldsymbol{y}_1\|_p}{\|\boldsymbol{A}\|_p}$$

plugging it into the above inequality, we get the following result:

$$d_1 = \frac{\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|_p}{\|\boldsymbol{x}_1\|_p} \leq \frac{\|\boldsymbol{A}\|_p * \|\boldsymbol{A}^+\|_p * \|\boldsymbol{y}_1 - \boldsymbol{y}_2\|_p}{\|\boldsymbol{y}_1\|_p}$$

$$= k(\boldsymbol{A}) * \frac{\|\boldsymbol{y}_1 - \boldsymbol{y}_2\|_p}{\|\boldsymbol{y}_1\|_p} = k(\boldsymbol{A}) * d_2 .$$

Based on the above analysis, the smaller $k(\boldsymbol{A})$, the better reconstruction performance is. Through SVD, we can use the mean singular values so as to reduce the correlation of the matrix. In addition, the reduced $k(\boldsymbol{A})$ can result in good sensing performance.

Deduction: Assume that $\boldsymbol{A}$ is the original matrix with the biggest singular value p and the smallest singular value q , then the biggest singular value of $\boldsymbol{A}^+$ is *1/q* . After SVD of $\boldsymbol{A}$, replacing all the singular values with their mean value l and getting matrix $\boldsymbol{A}'$. Then,

$$\frac{k(\boldsymbol{A}')}{k(\boldsymbol{A})} = \frac{\|\boldsymbol{A}'\|_p * \|\boldsymbol{A}'^+\|_p}{\|\boldsymbol{A}\|_p * \|\boldsymbol{A}^+\|_p} \leq \frac{l * 1/l}{p * 1/q} \leq 1, \text{ i.e., } k(\boldsymbol{A}') \leq k(\boldsymbol{A}) .$$

V. Optimized Sensing Matrix Design Based on Parseval Tight Frame and Matrix Decomposition

Considering the oracle estimator MSE, the equivalent sensing matrix ought to be close to a Parseval tight frame. But the sensing matrix based on this theory doesn't have the optimal non-coherence. In this section, we optimize the performance of the sensing matrix by further reducing the mutual coherence between the sensing and the sparsifying matrix combining with the matrix decomposition theory described in part IV.

A. *An Optimized Sensing Matrix Design Based on Approximate QR Matrix Decomposition*

According to the approximate decomposition theory, the correlation of the matrix can be further reduced by adjusting its singular values. The following are the specific steps of the proposed optimized sensing matrix design algorithm based on approximate QR decomposition, supposing the sparsifying matrix is known.

step1: Constructing the sensing matrix $\boldsymbol{\Phi}_{m \times n}$ according to (13);

step2: After making standard QR decomposition to the matrix $\boldsymbol{\Phi}$, getting an upper triangular matrix $\boldsymbol{R}$ and a square matrix $\boldsymbol{Q}$, noted as $\boldsymbol{\Phi} = \boldsymbol{Q}\boldsymbol{R}$;

step3: Getting a new matrix $\hat{\boldsymbol{R}}$ by keeping the main diagonal elements of $\boldsymbol{R}$ unchanged and setting the other elements all to zero, further getting the new sensing matrix $\tilde{\boldsymbol{\Phi}} = \boldsymbol{Q}\hat{\boldsymbol{R}}$.

From *Theorem 1, a*pproximate QR decomposition can well reduce the original sensing matrix's condition number and narrow the range of singular values of the original matrix. This method further reduced the mutual coherence between the sensing matrix and the sparsifying dictionary. It has been evaluated that the new sensing matrix has better RIP constant which is more applicable to compressed sensing.

B. *An Optimized Sensing Matrix Design Based on Mean Singular Value Decomposition*

According to the SVD decomposition theory, we can optimize the sensing matrix by replacing all the singular values with their average value. The following are the specific steps of the proposed optimized sensing matrix design algorithm, supposing the sparsifying matrix is known.

Step1: Constructing the sensing matrix $\boldsymbol{\Phi}_{m \times n}$ according to (13);

Step2: Adopting SVD decomposition to $\boldsymbol{\Phi}_{m \times n}$, denoted as $\boldsymbol{\Phi} = \boldsymbol{U} \begin{bmatrix} \boldsymbol{\Sigma} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \boldsymbol{V}^{\mathrm{T}}$;

Step3: After step2, getting all the singular values of $\boldsymbol{\Phi}_{m \times n}$, $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_m$. Then calculating the mean of all these singular values $mean = (\sigma_1 + \sigma_2 + \cdots + \sigma_m) / m$;

Step4: Making $\bar{\boldsymbol{\Sigma}} = \mathrm{diag}(\sigma_1^{'}, \cdots, \sigma_m^{'})$, where $\sigma_1^{'} = \sigma_2^{'} = \cdots = \sigma_m^{'} = mean$;

Step5: Getting the new sensing matrix $\tilde{\boldsymbol{\Phi}} = \boldsymbol{U} \begin{bmatrix} \bar{\boldsymbol{\Sigma}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \boldsymbol{V}^{\mathrm{T}}$.

From the Theorem 2 and its deduction, after adopting SVD decomposition to the sensing matrix $\boldsymbol{\Phi}_{m \times n}$, noted as $\boldsymbol{\Phi} = \boldsymbol{U} \begin{bmatrix} \boldsymbol{\Sigma} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \boldsymbol{V}^{\mathrm{T}}$, we get the new sensing matrix $\tilde{\boldsymbol{\Phi}}_{m \times n}$

by modifying its singular values through mean value algorithm. Using $\tilde{\boldsymbol{\Phi}}_{m\times n}$ can achieve higher reconstruction accuracy than $\boldsymbol{\Phi}_{m\times n}$ in compressed sensing.

VI. Experimental Results and Analysis

In order to verify the effectiveness of the proposed methods, firstly we test for 1-D sparse signals with several typical existed sensing matrices and the proposed optimized matrices, respectively, and compare the relative MSE of reconstruction signals with different reconstruction algorithms; Secondly, we calculate the PSNR (dB) of reconstructed 2-D image based on block compressed sensing with different sampling rate. And we use the dictionary trained by the *K*-SVD algorithm [12] for sparse representation for image patches. At last, we give the histograms of the distribution of the off-diagonal entries of the coherence matrices with respect to the equivalent sensing matrices of the sensing matrix.

A. Comparing the Relative Error Rate of the Reconstructed Signals for Different Measurements

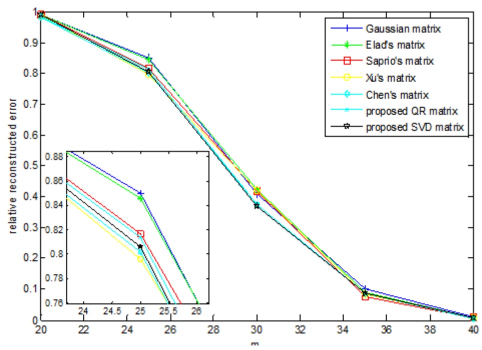


Figure 1. The relative reconstructed error rate vs. the number of measurements *m* using BPDN

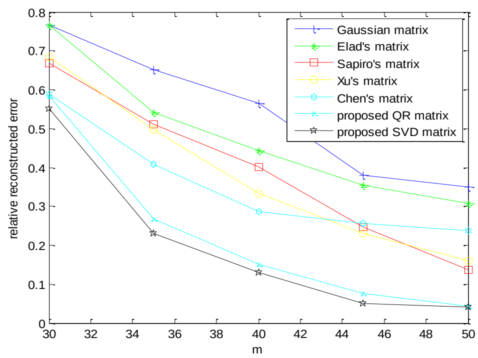


Figure 2. The relative reconstructed error rate vs. the number of measurements *m* using OMP

We generate a sparse signal with the length $n = 64$ and the sparsity $s = 10$ randomly by drawing its elements from i.i.d. zero mean and unit variance Gaussian distributions. The experimental result is calculated by averaging over 1000 trials. In each trial the relative error rate was evaluated as a function of m for different measurement matrices, in which m is the number of measurements. The measurements are corrupted by additive zero-mean Gaussian noise with variance $\sigma^2 = 10^{-4}$ .The specific results based on BPDN and orthogonal matching pursuit (OMP) [14] recovery algorithm are respectively shown in Fig. 1 and Fig. 2.

From Fig. 1 and Fig. 2, it's clear that the optimized sensing matrices in this paper are effective for 1-D sparse signals. Especially with OMP, the relative error rate with the optimized projections is much lower than others.

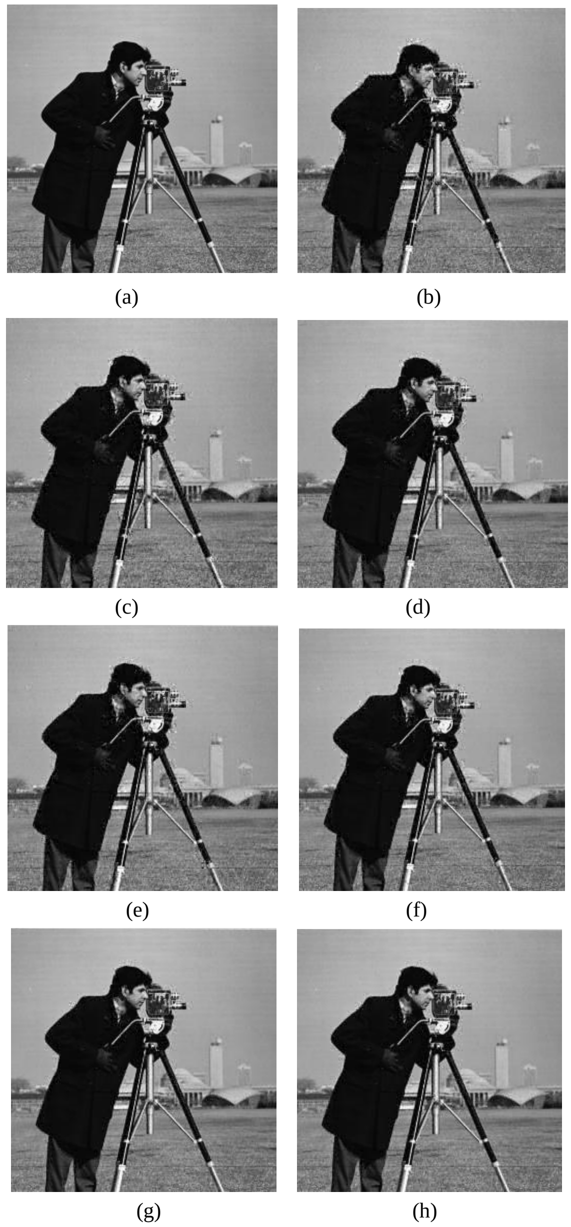


Figure 3. The reconstructed images based on different sensing matrices: (a) Original image, (b) Gaussian matrix, PSNR =28.40dB, (c) Elad's matrix, PSNR =29.13dB, (d) Sapario's matrix, PSNR =30.58dB, (e) Xu's matrix, PSNR =29.36dB, (f) Chen's matrix PSNR =30.85dB, (g)proposed QR based matrix, **PSNR =32.30dB**, (h) proposed SVD based matrix, **PSNR =32.33dB**.

B. Block Compressed Sensing and Reconstruction for the Image

In this part, we calculate the PSNR of the reconstructed image based on block compressed sensing [15] using different sensing matrices. Taking the 'Cameraman' image of size 256×256 for example. The image is partitioned into 1024 nonoverlapping patches of size 8×8, i.e., $n=64$. Instead of orthogonal basis, we use a dictionary of size 64×81 trained by the *K*-SVD for sparse representation of nonoverlapping patches. The number m of measurements for each patch is set to be equal to 40. And the measurements are corrupted by additive zero-mean Gaussian noise with the variance $\sigma^2=10^{-3}$. We use OMP to reconstruct each patch from its measurements owing to its fast execution. Fig. 3 gives the visual effect and the PSNR of reconstructed images with different sensing matrices.

From Fig. 3, we can easily find that the reconstruction performance is highly improved with matrices proposed in this paper, about 1.5dB higher than that of Chen's sensing matrix. In this test, the performance based on the SVD decomposition is close to the QR decomposition. In order to make further comparison, we give out the PSNR of reconstructed images based on various sensing matrices with different sampling rates in Table I. Fig. 4 shows us the PSNR curves versus subrate based on different sensing matrices.

From Table I and Fig. 4, for any sampling rate, the PSNR of reconstructed images can be improved significantly with the optimized sensing matrices proposed in this paper than other matrices.

TABLE I. THE PSNR (DB) OF RECONSTRUCTED IMAGE BASED ON DIFFERENT SENSING MATRICES WITH DIFFERENT SAMPLING RATES

Sensing matrix	Subrate				
	0.3	0.4	0.5	0.6	0.7
Gaussian matrix	22.04	24.41	27.03	28.87	30.14
Elad's matrix	22.94	24.68	26.73	29.08	30.41
Sapiro's matrix	21.28	24.05	29.17	30.42	30.55
Xu's matrix	22.55	24.59	29.09	30.53	30.64
Chen's matrix	23.44	24.53	30.26	30.85	30.72
QR matrix	23.60	26.53	31.05	32.18	32.72
SVD matrix	24.71	25.60	31.08	32.33	32.96

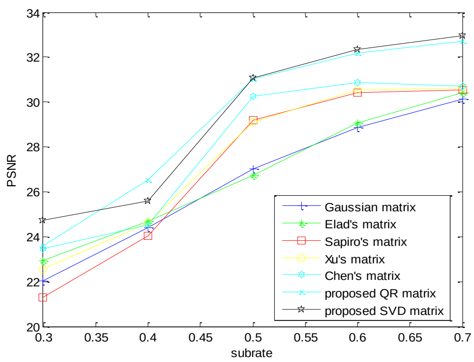
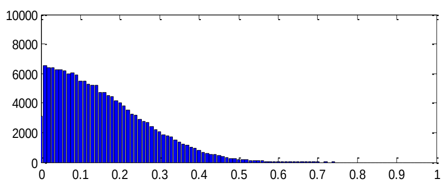
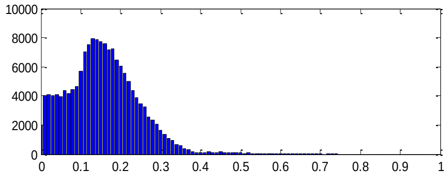


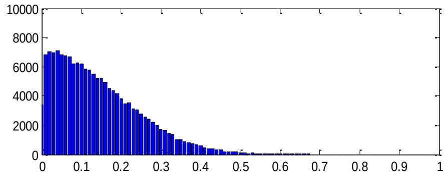
Figure 4. PSNR curves vs. subrate based on different sensing matrices



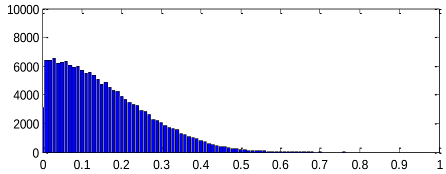
(a)



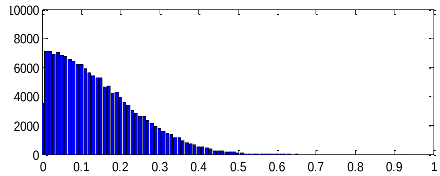
(b)



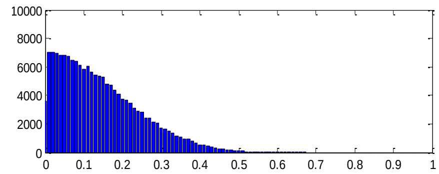
(c)



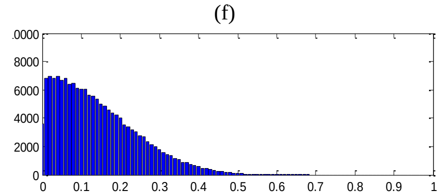
(d)



(e)



(f)



(g)

Figure 5. The histogram of the absolute off-diagonal entries of the Gram matrices of various measurement matrices: (a) Gaussian matrix, μ=0.7, (b) Elad matrix, μ=0.78, (c) Sapiro's matrix, μ=0.63, (d) Xu's matrix, μ=0.64, (e) Chen's matrix, μ=0.66, (f) proposed QR based matrix, **μ=0.61** (g) proposed SVD based matrix, **μ=0.60**.

C. Distribution of the Absolute Values of the off-Diagonal Entries of the Coherence Matrix

In this part, we compare the histograms of the absolute values of the off-diagonal entries of the Gram matrices of the proposed sensing matrices with other existed sensing matrices, namely, Gaussian random matrix, Elad's optimized matrix, Sapiro's learned matrix, Xu's matrix and Chen's frame based matrix. Fig. 5 gives the detailed results in which we use a random orthogonal dictionary $\boldsymbol{\Psi}\in \mathrm{R}^{200\times400}$ with entries drawn from i.i.d. zero mean and unit variance Gaussian distributions. In Fig. 5, the abscissa stands for the value of the elements and the ordinate stands for the number of elements fallen in each range. The measurement number m is equal to 30. From Fig. 5, the distributions of the off-diagonal entries of the two proposed approaches are obviously better than Gaussian matrix. In order to make comparison more clearly, the values of μ defined in Eq. (5) for various sensing matrices are given in Fig. 5. It's well known that the smaller the μ, the weaker the mutual coherence between the sensing matrix and the sparsifying dictionary is. We can find from Fig.5 that the proposed optimized sensing matrices have smaller μ than others.

VII. Conclusion

It's of great importance to design good and efficient sensing matrices for compressed sensing. The design method based on Parseval tight frame proposed by Chen is proven superior to the state-of-art designs. Based on this method, we propose optimized approaches to further reduce the mutual coherence of the coherence matrix combining with the matrix decomposition theory. The simulation results show that our proposed algorithm obviously improve the signal reconstruction performance. It offers a new idea to optimize the sensing matrix by considering the mutual coherence and the MSE performance at the same time.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (Grant No. 61002027). The authors would like to thank Xiaoyan Fu for her help and valuable suggestions to improve the presentation of the paper.

References

[1] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289-1306, 2006.

[2] E. J. Candès, "The restricted isometry property and its implications for compressed sensing," *C R Math Acad Sci Paris, Ser. I*, vol. 346, pp. 589-592, 2008.

[3] E. J. Candès and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, pp. 4203-4215, 2005.

[4] T. Do, L. Gan, N. Nguyen, and T. Tran, "Fast and efficient compressive sensing using structurally random matrices," *IEEE Trans. on Signal Processing*, vol. 60, no.1, pp. 139-154, 2012.

[5] W. Yin, S. Morgan, J. Yang, and Z .Yin, "Practical compressive sensing with toeplitz and circulant matrices," *Rice University CAAM Technical Report TR10-01*, Jan., 2010.

[6] R. Berinde and P. Indyk, "Sparse recovery using sparse random matrices," *MIT-CSAIL Technical Report*, 2008.

[7] M. Elad, "Optimized projections for compressed sensing," *IEEE Trans. on Signal Processing*, vol. 55, no. 12, pp. 5695-5702, 2007.

[8] J. M. Duarte-Carvajalino and G. Sapiro, "Learning to sense sparse signals: Simultaneous sensing matrix and sparsifying dictionary optimization," *IEEE Trans. Image Processing*, vol. 18, no. 7, pp. 1395-1408, 2009.

[9] J. Xu, Y. Pi, and Z. Cao, "Optimized projection matrix for compressive sensing," *EURASIP J. Adv. Signal Process*., vol. 43, pp. 1–8, 2010.

[10] W. Chen, M. R. D. Rodrigues, and I. J. Wassell, "Projection design for statistical compressive sensing: A tight frame based approach," *IEEE Trans. on Signal Processing*, vol. 61, no.8, pp. 2016–2029, 2013.

[11] G. Golub and C. Loan, "Matrix computations," *Johns Hopkins: Baltimore*, 1996.

[12] M. Aharon, M. Elad, and A. M. Bruckstein, "The K-SVD: An algorithm for designing of over-complete dictionaries for sparse representation," *IEEE Trans. on Signal Processing*, vol. 54, no. 11, pp. 4311-4322, 2006.

[13] J. A. Tropp, I. Dhillon, R. W. Heath, and T. Strohmer, "Designing structured tight frames via alternating projection," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 188-209, 2005.

[14] J. Tropp and A. C. Gilbert, "Signal recovery from partial information via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655-4666, 2007.

[15] L. Gan, "Block compressed sensing of natural images," in *Proc. International Conference on Digital Signal Processing*, Cardiff, UK, 2007, pp. 403-406.

Aiping Yang received her B.Ed. degree in mathematics education from Shandong Normal University, China, in 2000 and her M.S. degree in operations research and control theory from Tianjin University, China, in 2003 and her Ph.D. degree in signal and information processing from Tianjin University, China, in 2008. Now she is an associate professor in Tianjin University, China. Her current research interests include digital image and video processing, compressive sensing theory, and wavelet analysis.

Jinxia Zhang received her bachelor's degree in communication engineering from Shandong University, China, in 2011. Now she is a postgraduate student in school of electronic and information engineering, Tianjin University, China. She mainly researches compressive sensing theory, especially the CS matrix design methods.

Zhengxin Hou received the B.E. degree in electrical engineering from Peking University, China, in 1969 and the M.E. degree from Tianjin University, China, in 1982. Now he is a professor and supervisor of the Ph.D. students in Tianjin University, China. His current research interests include compressed sensing, digital signal processing, and wavelet analysis.

