Article

A New Method of Measurement Matrix Optimization for Compressed Sensing Based on Alternating Minimization

Renjie Yi [1,*] **, Chen Cui** [1] **, Biao Wu** [2] **and Yang Gong** [1]

1   Institute of Electronic Countermeasure, National University of Defense Technology, Hefei 230000, China; kycuichen@163.com (C.C.); gongyang17@nudt.edu.cn (Y.G.)
2   Huayin Ordnance Test Center, Weinan 714000, China; wbuiao@163.com
*   Correspondence: yirenjie17@nudt.edu.cn; Tel.: +86-182-2665-6895

Abstract: In this paper, a new method of measurement matrix optimization for compressed sensing based on alternating minimization is introduced. The optimal measurement matrix is formulated in terms of minimizing the Frobenius norm of the difference between the Gram matrix of sensing matrix and the target one. The method considers the simultaneous minimization of the mutual coherence indexes including maximum mutual coherence μ_{max}, t-averaged mutual coherence μ_{ave} and global mutual coherence μ_{all}, and solves the problem that minimizing a single index usually results in the deterioration of the others. Firstly, the threshold of the shrinkage function is raised to be higher than the Welch bound and the relaxed Equiangular Tight Frame obtained by applying the new function to the Gram matrix is taken as the initial target Gram matrix, which reduces μ_{ave} and solves the problem that μ_{max} would be larger caused by the lower threshold in the known shrinkage function. Then a new target Gram matrix is obtained by sequentially applying rank reduction and eigenvalue averaging to the initial one, leading to lower. The analytical solutions of measurement matrix are derived by SVD and an alternating scheme is adopted in the method. Simulation results show that the proposed method simultaneously reduces the above three indexes and outperforms the known algorithms in terms of reconstruction performance.

Keywords: compressed sensing; measurement matrix; Equiangular Tight Frame; mutual coherence

1. Introduction

Compressed sensing (CS) [1] can sample the sparse or compressible signals at a sub-Nyquist rate, which brings great convenience for data storage, transmission, and processing. By adopting the reconstruction algorithms, the signal can be exactly reconstructed from the sampled data. As a marvel way of signal processing, CS is applied in different fields such as image encryption [2], wideband spectrum sensing [3], wireless sensor network data processing [4], etc.

The original signal $x \in R^{N \times 1}$ is assumed to have a sparse representation in a known domain $\Psi \in R_N^{N \times L}(N \le L)$ as $x = \Psi s$ where Ψ is the dictionary matrix and s is a K-sparse signal. The incomplete measurement $y \in R^{M \times 1}$ is obtained through the linear model

$$y = \Phi x = Ds \tag{1}$$

where $\Phi \in R^{M \times N}(M < N)$ is called the measurement matrix and $D = \Phi\Psi$ is the sensing matrix.

Some specific properties of Φ have great impacts on the reconstruction performance. In [5,6], spark and restricted isometry property (RIP) are respectively proposed as the sufficient conditions on Φ to recovery guarantee. However, computing the spark of a matrix has combinatorial complexity and certifying RIP for a matrix requires combinatorial search, that is to say, these tasks are NP-hard and difficult to accomplish. To a large extent, the coherence between Φ and Ψ reflects the performance of meeting the above conditions.

Actually, the coherence is equivalent to the mutual coherence of D. Since the mutual coherence can be easily manipulated to provide recovery guarantees, it is commonly used to measure the performance of Φ. The frequently-used mutual coherence indexes include the maximum mutual coherence μ_{max} [7], t-averaged mutual coherence μ_{ave} [8] and global mutual coherence μ_{all} [9], which respectively represent the maximum, the average and the sum of squares of the correlation between any distinct pair of columns in D.

The first attempt to consider the optimal design of Φ is given in [8]. The simulation results carried out in [8] show that the optimized Φ leads to smaller μ_{ave} and a substantially better CS reconstruction performance is obtained. Then the optimization of Φ becomes an important issue in CS. The recent works try to find the optimal Φ which has excellent performance in reducing the mutual correlation by minimizing the Frobenius norm of the difference between the Gram matrix $G = D^{'}D$ and the target Gram matrix G_t. The main work focuses on designing G_t and finding the best Φ.

In [8], G_t is obtained by shrinking the off-diagonal entries in G. The shrinkage technique reduces μ_{ave} but it is time-consuming. Furthermore, μ_{max} is still large, which ruins the worst-case guarantees of the reconstruction algorithms. A suitable point between the current solution and the one obtained using a new shrinkage function is chosen to design the G_t in [10]. It is of very strong competitiveness in μ_{ave} and μ_{max}. However, the optimal point is hard to determine and the unsuitable point may seriously degrade the algorithm's performance. In [9], G_t is obtained by averaging the eigenvalues of G. Simulation results show that μ_{all} is reduced effectively but reducing μ_{ave} and μ_{max} is hard to be guaranteed, which means μ_{ave} and μ_{max} may maintain large values. Duarte-Carajalino and Sapiro [11] set G_t as an identity matrix. Since D is overcomplete and G cannot be an identity matrix, simply minimizing the difference between G and G_t does not imply low μ_{max} [12]. In [13–17], G_t is chosen from a set of relaxed Equiangular Tight Frames (ETF) [18]. The set can be formulated as $S = \left\{ G_t \in R^{L \times L} : G_t = G_t{'}, diag G_t = 1, \max_{i \neq j} |G_t(i,j)| \leq \mu_{welch} \right\}$ where μ_{welch} denotes the Welch bound [19] and $G_t(i,j)$ denotes the (i,j)th entry of G_t. However, the maximum absolute value of off-diagonal entries in G is almost always greater than μ_{welch}. In this case, the optimization usually implies a solution D with low μ_{ave} but high μ_{max}. In summary, the target Gram matrices mentioned above only focus on a certain mutual coherence index, and fail to take into account μ_{ave}, μ_{max} and μ_{all} simultaneously. When a certain index is targeted, the other indexes may not decrease significantly or even increase. Therefore, Φ is not 'good' enough and the reconstruction performance is well below par.

After designing G_t, the next step is to find the 'best' Φ by approaching G to G_t. In [8–11,17], G is obtained by applying SVD to G_t primarily and then the square root $\hat{D}$ of G is built as $\hat{D}^{'}\hat{D} = G$. At last, Φ is obtained by $\Phi = \hat{D}\Psi^{\dagger}$ where $\dagger$ denotes the Moore Penrose pseudoinverse. This kind of method is intuitive, but the generalized pseudoinverse poses problems of calculation accuracy and robustness [15]. In [14,15], gradient algorithm and quasi-Newtonian algorithm are respectively utilized to attain Φ. Firstly, the cost function $F(\Phi)$ with Φ as the variable is constructed. Then the search direction is determined by the derivative of $F(\Phi)$. Finally, Φ is obtained with a fixed step size. However, choosing a suitable step size which has a great influence on the accuracy of the solution requires a lot of comparison work. Moreover, the gradient algorithm and quasi-Newtonian algorithm cannot converge until a certain number of iterations is accomplished, resulting in high computational cost. In [11,16], the method for designing Φ shares the same concept as K-SVD [20], that is to update a matrix row by row. Eigenvalue decomposition is required to find the square root of the maximum eigenvalue for each row, which results in a significant increase in the calculation. For solving this problem, Hong et al. [16] utilize the power method instead of eigenvalue decomposition. However, the eigenvalue obtained by power method is the one with the largest absolute value. When the eigenvalue is negative, eigenvalue decomposition is still necessary.

The primary contributions of this paper are threefold:

- The new target Gram matrix G_t targets μ_{ave}, μ_{max}, and μ_{all} of D simultaneously is designed. Firstly, a new shrinkage function whose threshold exceeds μ_{welch} is utilized to determine the initial target Gram matrix. Then G_t is obtained by sequentially applying rank reduction and eigenvalue averaging to the initial matrix.
- Analytical solutions of the measurement matrix $\mathbf{\Phi}$ to minimize the difference between $G = \mathbf{\Psi}'\mathbf{\Phi}'\mathbf{\Phi}\mathbf{\Psi}$ and G_t are derived by SVD.
- Based on alternating minimization, an iterative method is proposed to optimize the measurement matrix. The simulation results confirm the effectiveness of the proposed method in decreasing the mutual coherence indexes and improving reconstruction performance.

The remainder of this paper is organized as follows. Some basic definitions related to mutual coherence indexes and frames are described in the next section. The main results are presented in Section 3, where the solutions to the G_t design are characterized and a class of the solutions to the optimal $\mathbf{\Phi}$ is derived in detail. The procedure of our method and the discussion can be also found in Section 3. In Section 4, simulations are carried out to confirm the effectiveness of the proposed method. In the end, the conclusion is drawn.

2. Mutual Coherence Indexes and ETFs

2.1. Mutual Coherence Indexes

Rewrite $D = [d_1, d_2 \cdots d_L] \in R^{M \times L}$ where $d_i \in R^{M \times 1}$ and $\|d_i\|_2 = 1$. Denote $g_{ij} = d'_i d_j$ the entry at the position of row i and column j in G, where $i, j = 1, 2 \cdots L$. Here, we quote the definitions of the mutual coherence indexes as that presented by Donoho [7], Elad [8], and Zhao [9].

Definition 1. *For a matrix D, the maximum mutual coherence μ_{max} is defined as the largest absolute and normalized inner product between all columns in D that can be described as*

$$\mu_{max} = \max_{i \neq j} \left| d'_i d_j \right| = \max_{i \neq j} |g_{ij}| \tag{2}$$

Definition 2. *For a matrix D, the t-averaged mutual coherence μ_{ave} is defined as the average of all absolute and normalized inner products between different columns in D that are above t and can be described as*

$$\mu_{ave} = \frac{\sum\limits_{i \neq j} (|g_{ij} \geq t|) |g_{ij}|}{\sum\limits_{i \neq j} (|g_{ij} \geq t|)} \tag{3}$$

Definition 3. *For a matrix D, the global mutual coherence μ_{all} is defined as the sum of squares of normalized inner products between all columns in D that can be described as*

$$\mu_{all} = \sum_{i \neq j} g_{ij}^2 \tag{4}$$

As shown in [5], the original signal can be exactly reconstructed as long as $K < (1 + 1/\mu_{max})/2$. The conclusion is true from a worst-case standpoint which means that μ_{max} does not do justice to the actual behavior of sparse representations. Therefore, Elad considers that an "average" measure of mutual coherence, namely μ_{ave}, is more likely to describe its true behavior. Different from the previous two indexes, μ_{all} reflects the overall property of D.

In fact, the purpose of reducing the mutual coherence indexes of D is to attain G that meets the following requirements: (1) The maximum absolute value of off-diagonal entries in G is sufficiently small; (2) The number of off-diagonal entries with large absolute value

is minimized; (3) The average of off-diagonal entries with large absolute value is as small as possible. However, when a certain mutual coherence index is targeted solely, we cannot guarantee that the obtained G will fully meet the requirements. Therefore, the decrease of a certain index does not always mean better $\mathbf{\Phi}$ and improved reconstruction performance. When the three indexes are reduced simultaneously, the requirements are better satisfied and better performance is obtained.

2.2. ETFs

It is shown in [19] that μ_{max} of $D \in R_M^{M \times L}$ is lower bounded by

$$\mu_{max} \geq \mu_{welch} = \sqrt{\frac{L-M}{M(L-1)}}$$

The bound is achievable for ETF. Here, we recall the definition of ETF [18].

Definition 4. *Let F be a $M \times L$ matrix whose columns are $f_1, f_2 \cdots f_L$. The matrix F is called an equiangular tight frame if it satisfies three conditions*

(1) *Each column has a unit norm: $\|f_i\|_2 = 1$ for $i = 1, 2 \cdots L$.*
(2) *The columns are equiangular. For some nonnegative θ, we have $f'_i f_j = \theta$ when $i, j = 1, 2 \cdots L$ and $i \neq j$.*
(3) *The columns form a tight frame. That is, $FF' = (L/M)I_M$ where I_M is an $M \times M$ identity matrix.*

Sustik et al. [18] show that a real $M \times L$ ($1 < M < L - 1$) ETF exists on if $L \leq \min\{M(M+1)/2, (L-M)(L-M+1)/2\}$ holds. Furthermore, $\sqrt{M(L-1)/(L-M)}$ and $\sqrt{(L-M)(L-1)/M}$ must be odd integers when $L \neq 2M$, M, and $2M - 1$ must be an odd number and the sum of two squares respectively when $L = 2M$. Fickus et al. [21] surveys some known construction of ETFs and tabulates existence for sufficiently small dimensions. The above studies show that M and L must meet some exacting requirements when an ETF is available for D. However, it is really difficult to meet the requirements in practice, which means the maximum absolute value of the off-diagonal entries in G is usually significantly larger than μ_{welch}.

3. The Proposed Method

The off-diagonal entries in G actually are the inner products between different columns in D. Reducing those entries is likely to lead to lower mutual coherence indexes and better performance. The most straightforward approach is to replace large off-diagonal values with small ones. However, it is impossible to solve $\mathbf{\Phi}$ from a certain G because of the inequality of rank between $\mathbf{\Psi}' \mathbf{\Phi}' \mathbf{\Phi} \mathbf{\Psi}$ and G when the approach is adopted. Therefore, a feasible approach is to minimize the difference between G and G_t that can be formulated as

$$\min \|G_t - G\|_F^2 \tag{5}$$

where $G = \mathbf{\Psi}' \mathbf{\Phi}' \mathbf{\Phi} \mathbf{\Psi}$. This problem can be solved by alternating minimization strategy [14,16], which iteratively minimizes (5) to find the desired $\mathbf{\Phi}$. The idea is to update G_t and $\mathbf{\Phi}$ alternatively and repeat this proceeding until a stop criterion is reached. In this section, we design G_t firstly and then derive the analytical solutions of $\mathbf{\Phi}$. Finally, an iterative method is proposed to optimize the measurement matrix based on alternating minimization.

3.1. The Design of G_t

It can be seen from (5) that G_t plays an important role in measurement matrix optimization. In recent works, G_t is frequently set as the relaxed ETF matrix, which is obtained by applying the following shrinkage function

$$G_t(i,j) = \begin{cases} g_{ij}, & |g_{ij}| \leq \varsigma \\ sign(g_{ij})\varsigma, & otherwise \end{cases} \tag{6}$$

where $\varsigma = \mu_{welch}$ for $i, j = 1, 2 \cdots L$ and $i \neq j$. Such a scheme in designing G_t guarantees that the off-diagonal entries with large value of G will be intensively constrained, which means lower μ_{ave} and μ_{max}.

Recall from Section 2.2 that the Welch bound is not achievable for G in most cases. As shown in [14,16], different ς yields different results and μ_{welch} is not the optimal value. Li [22] et al. found that a smaller μ_{max} is available when ς is slightly larger than μ_{welch}. Inspired by [22], we propose an improved shrinkage function which divides the entries in G into three segments through two thresholds. One of the thresholds is μ_{welch} and the other is larger than μ_{welch}. The function is as follows

$$G_t(i,j) = \begin{cases} sign(g_{ij})Thr, & |g_{ij}| \geq Thr \\ sign(g_{ij})\mu_{welch}, & \mu_{welch} \leq |g_{ij}| < Thr \\ g_{ij}, & |g_{ij}| < \mu_{welch} \end{cases} \tag{7}$$

where $Thr = \mu_{welch} + c$ and $0 < c < \mu_{welch}$. As can be seen from Equation (7), the maximum absolute value of off-diagonal entries in G_t is raised from μ_{welch} to Thr. According to the previous analysis, the new function is likely to lead to a further reduction in μ_{max} while maintaining the advantage of Equation (6) with respect to μ_{ave}.

After shrinkage, G_t becomes full rank generally [8], that is $Rank(G_t) = L$. However, the rank of G is identically equal to M. Thus, we consider mending this by forcing a rank M. A new target Gram matrix, denoted as G_{t_M}, is obtained by solving

$$\min \|G_t - G_{t\_M}\|_F^2, \text{ s.t. } Rank(G_{t\_M}) = M \tag{8}$$

The solutions to this problem are given by Theorem 1 below.

Theorem 1. *Let $G_t \in R_L^{L \times L}$ be the matrix obtained by applying the shrinkage operation shown as Equation (7) to G and $G_t = P\Lambda P'$ be the eigendecomposition of G_t. P is orthonormal with dimension L and $\Lambda = diag(\lambda_1, \lambda_2 \cdots \lambda_L)$ with $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_L|$. The solutions of the minimization problem defined by (8) are characterized by*

$$G_{t\_M} = PA\Lambda P' \tag{9}$$

where $A = \begin{bmatrix} I_M & 0 \\ 0 & 0 \end{bmatrix} \in R^{L \times L}$.

Proof. Denote $G_{t\_M} = X'X$ where $X \in R_M^{M \times L}$. Let $X = U_X \begin{bmatrix} \Sigma_X & 0 \end{bmatrix} V'_X$ be an SVD of X, where $U_X \in R^{M \times M}$ and $V_X \in R^{L \times L}$ are unitary. Then G_{t_M} can be rewritten as

$$G_{t\_M} = V_X \begin{bmatrix} \Sigma_X^2 & 0 \\ 0 & 0 \end{bmatrix} V'_X \tag{10}$$

Denote $f = \|G_t - G_{t\_M}\|_F^2$. By substituting G_{t_M} with $G_{t\_M} = X'X$, can be rewritten as a function of matrix X. Let $\partial f / \partial X$ be the derivative of f with respect to X. The optimal X should satisfy $\partial f / \partial X = 0$. Equivalently, we have

$$XX'X = XG_t \tag{11}$$

It then follows from $X = U_X \begin{bmatrix} \Sigma_X & 0 \end{bmatrix} V'_X$ that

$$\begin{bmatrix} \Sigma_X^2 & 0 \\ 0 & 0 \end{bmatrix} = AV'_X G_t V_X \tag{12}$$

Substituting Equation (12) into Equation (10), we obtain

$$G_{t\_M} = V_X AV'_X G_t \tag{13}$$

It turns out from the unitary invariance with Equation (13) that

$$f = \left\| \left(I_L - V_X AV'_X \right) G_t \right\|_F^2 \tag{14}$$

With a few manipulations, we conclude that the solution of (8) is equivalent to solving

$$\max tr\left(G_t' V_X AV'_X G_t \right) \tag{15}$$

where $tr()$ denotes the matrix trace operation. Noting that $A = A'A$ and $G_t = P\Lambda P'$, the problem in (15) is equivalent to

$$\max \left\| AV'_X P\Lambda \right\|_F^2 \tag{16}$$

Denote $B = V'_X P$ and rewrite B as $B = [b_1, b_2 \cdots b_L]$ where $b_i \in R^{L \times 1}$ for $i = 1, 2 \cdots L$. Rewrite A as $A = [e_1, e_2 \cdots e_M, 0, 0 \cdots 0]'$, where $e_j \in R^{L \times 1}$ denotes a unit vector with the ith entry is equal to 1 for $j = 1, 2 \cdots M$. Then, it is easy to obtain that

$$AV'_X P\Lambda = \begin{bmatrix} \lambda_1 e'_1 b_1 & \lambda_2 e'_1 b_2 & \cdots & \lambda_L e'_1 b_L \\ \lambda_1 e'_2 b_1 & \lambda_2 e'_2 b_2 & \cdots & \lambda_L e'_2 b_L \\ \vdots & & & \\ \lambda_1 e'_M b_1 & \lambda_2 e'_M b_2 & \cdots & \lambda_L e'_M b_L \\ 0 & 0 & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

Let $b_{ji} = e'_j b_i$ be the jth entry of b_i. It can be shown with some manipulations that the problem in (16) is equivalent to

$$\max \sum_{i=1}^{L} \sum_{j=1}^{M} \lambda_i^2 b_{ji}^2 \tag{17}$$

With $|\lambda_1| \ge |\lambda_2| \ge \cdots \ge |\lambda_L|$ and $\sum_{i=1}^{L} \sum_{j=1}^{M} b_{ji}^2 = M$, it is straightforward that $\sum_{i=1}^{L} \sum_{j=1}^{M} \lambda_i^2 b_{ji}^2$ reaches the maximum value $\lambda_1^2 + \lambda_2^2 + \cdots + \lambda_M^2$ only if $\sum_{j=1}^{M} b_{ji}^2 = 1$ for $i = 1, 2 \cdots M$. In this case, $\sum_{j=1}^{M} b_{ji}^2 = 0$ and $\sum_{j=1}^{M} b_{ij}^2 = 0$ hold for $i = M + 1, M + 2 \cdots L$. Rewrite B as

$B = \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix}$ where $B_1 \in R^{M \times M}$. Accordingly, all of the entries in both B_2 and B_3 are zero. Noting that B is unitary, it can be shown that B_1 and B_4 are unitary. Hence, we have

$$
\begin{aligned}
G_{t\_M} &= V_X A V'_X P \Lambda P' \\
&= P B' A B \Lambda P' \\
&= P A \Lambda P'
\end{aligned}
\tag{18}
$$

As can be seen, the rank of G_{t_M} is equal to M. The proof is then completed. □

After the operation of rank reduction, the rank of G_{t_M} is equal to that of G. Additionally, G_{t_M} is most similar to G_t in terms of Frobenius norm. Inspired by [9], we reduce the sum of squares of all off-diagonal values of G_{t_M}, namely $\hat{\mu}_{all}$, by eigenvalue averaging. When minimizing the difference between G and G_{t_M}, a smaller $\hat{\mu}_{all}$ is more likely to lead to a smaller μ_{all}. $\hat{\mu}_{all}$ can be formulated as

$$
\hat{\mu}_{all} = \sum_{i,j=1}^{L} \hat{g}_{ij}^2 - \sum_{i=1}^{L} \hat{g}_{ii}^2
\tag{19}
$$

where $\hat{g}_{ij}$ denotes the (i,j)th entry of G_{t_M} and $\hat{g}_{ii} = 1$ holds for $i = 1, 2 \cdots L$.

Noting that $\sum_{i,j=1}^{L} g_{ij}^2 = \|\hat{G}_t\|_F^2$ and $G_{t\_M} = P \hat{\Lambda} P'$ where $\hat{\Lambda} = diag(\lambda_1, \lambda_2 \cdots \lambda_M, 0 \cdots 0)$, $\hat{\mu}_{all}$ can be rewritten as

$$
\hat{\mu}_{all} = \sum_{i=1}^{M} \lambda_i^2 - \sum_{i=1}^{L} \hat{g}_{ii}^2
\tag{20}
$$

Assuming that $\sum_{i=1}^{M} \lambda_i$ is invariable, it then follows from the Cauchy BuniakowskySchwarz Inequality that $\sum_{i=1}^{M} \lambda_i^2$ takes the minimum value only if $\lambda_i = \frac{1}{M} \sum_{i=1}^{M} \lambda_i$ for $i = 1, 2 \cdots M$. Let $\hat{\lambda} = \frac{1}{M} \sum_{i=1}^{M} \lambda_i$ and $\hat{\Lambda} = diag\left(\underbrace{\hat{\lambda}, \hat{\lambda} \cdots \hat{\lambda}}_{M}, 0 \cdots 0 \right)$, a new target Gram matrix denoted as G_{t_opt} with M equal all non-zero eigenvalues is given by

$$
G_{t\_opt} = P \hat{\Lambda} P'
\tag{21}
$$

Recall that G_{t_M} is most similar to G_t in terms of Frobenius norm, it means that G_{t_M} is of good competitiveness in μ_{ave} and μ_{max}. Furthermore, as a variant of G_{t_M}, G_{t_opt} reduces the sum of squares of all off-diagonal values of G_{t_M}, leading to a better performance in minimizing μ_{all}. Therefore, G_{t_opt} is more likely to be an ideal solution of target Gram matrix which leads to better μ_{ave}, μ_{max}, and μ_{all} simultaneously.

3.2. The Analytical Solutions of Φ

After obtaining the target Gram matrix G_{t_opt}, the next step of the optimization is to find the best Φ. To handle the problem, we try to find the optimal solution by minimizing the difference between $\Psi' \Phi' \Phi \Psi$ and G_{t_opt} as

$$
\min \left\| G_{t\_opt} - \Psi' \Phi' \Phi \Psi \right\|_F^2
\tag{22}
$$

Let $D = U_D \begin{bmatrix} \Sigma_D & 0 \end{bmatrix} V'_D$ be an SVD of D, where $U_D \in R^{M \times M}$ and $V_D \in R^{L \times L}$ are unitary. Similarly, Let $\Psi = U_\Psi \begin{bmatrix} \Sigma_\Psi & 0 \end{bmatrix} V'_\Psi$ be an SVD of Ψ, where $U_\Psi \in R^{N \times N}$ and $V_\Psi \in R^{L \times L}$ are unitary. The solutions to this problem are given by Theorem 2 below.

Theorem 2. *Let G_{t_opt} be the matrix shown as Equation (21) and $\Lambda_M \in R^{M \times M}$ be the Mth principal submatrix of $\hat{\Lambda}$. Then the solutions of the minimization problem defined by (22) are characterized by*

$$\Phi_{opt} = U_Z \begin{bmatrix} (\Lambda_M)^{\frac{1}{2}} & 0 \end{bmatrix} P' V_{\Psi} \begin{bmatrix} \Sigma_{\Psi}^{-1} & 0 \end{bmatrix}' U'_{\Psi} \tag{23}$$

where $U_Z \in R^{M \times M}$ is an arbitrary unitary matrix.

Proof. Assume that the off-diagonal values of Σ_D and Σ_{Ψ} are non-zero, Φ can be written as

$$\Phi = U_D \begin{bmatrix} \Sigma_D & 0 \end{bmatrix} V'_D V_{\Psi} \begin{bmatrix} \Sigma_{\Psi}^{-1} & 0 \end{bmatrix}' U'_{\Psi} \tag{24}$$

By substituting Φ in (22) with Equation (24), it can be shown with some manipulations that the solutions of the problem in (22) are equivalent to the solutions of

$$\min \left\| V'_D P \hat{\Lambda} P' V_D - \begin{bmatrix} \Sigma_D^2 & 0 \\ 0 & 0 \end{bmatrix} \right\|_F^2 \tag{25}$$

Let $Z = V'_D P \hat{\Lambda} P' V_D$ and z_i be the ith diagonal entry of Z. Denote $\Lambda_Z = diag(z_1, z_2 \cdots z_M)$. With further manipulations, we simplify (25) to

$$\min \left\| G_{t\_opt} \right\|_F^2 + \left\| \Lambda_Z - \Sigma_D^2 \right\|_F^2 - \left\| \Lambda_Z \right\|_F^2 \tag{26}$$

Obviously, the minima are achievable only if $\Lambda_Z = \Sigma_D^2$ holds and $\| \Lambda_Z \|_F^2$ takes the maximum value. Let $U = V'_D P$ and u_{ij} be the (i, j)th entry of U where $i, j = 1, 2 \cdots L$. Noting that the top M diagonal entries of $\hat{\Lambda}$ are all equal to $\hat{\lambda}$, we have

$$z_i = \hat{\lambda} \left(u_{i1}^2 + u_{i2}^2 + \cdots + u_{iM}^2 \right) \tag{27}$$

It is worth noting that $\| \Lambda_Z \|_F^2 = \sum_{i=1}^{M} z_i^2$ and $\sum_{i=1}^{L} z_i = M\hat{\lambda}$ where $0 \leq z_i \leq \hat{\lambda}$. Hence, it is clear that the maxima of $\sum_{i=1}^{M} z_i^2$ are reached when z_i takes the maximum value $\hat{\lambda}$ for $i = 1, 2 \cdots M$. Rewrite U as $U = \begin{bmatrix} U_1 & U_2 \\ U_3 & U_4 \end{bmatrix}$ where $U_1 \in R^{M \times M}$. Since $z_i = \hat{\lambda}$ holds for $i = 1, 2 \cdots M$, we have $u_{i1}^2 + u_{i2}^2 + \cdots + u_{iM}^2 = 1$ and $U_2 = 0$, $U_3 = 0$ accordingly. As U_Z is a unitary matrix, it is easy to verify that U_1 and U_4 are both unitary matrices. Then it follows that

$$V'_D = \begin{bmatrix} U_1 & 0 \\ 0 & U_4 \end{bmatrix} P' \tag{28}$$

Substituting Equation (28) into Equation (24), the optimal solution is obtained by

$$\Phi_{opt} = U_Z \begin{bmatrix} (\Lambda_M)^{\frac{1}{2}} & 0 \end{bmatrix} P' V_{\Psi} \begin{bmatrix} \Sigma_{\Psi}^{-1} & 0 \end{bmatrix}' U'_{\Psi} \tag{29}$$

The proof is then completed. $\square$

3.3. Comments

According to Sections 3.1 and 3.2, the procedure for measurement matrix optimization has been summarized in Algorithm 1.

Algorithm 1. The proposed optimization method.

Input: Dictionary matrix $\boldsymbol{\Psi} \in R_N^{N \times L}$ which has an SVD form of $\boldsymbol{\Psi} = \boldsymbol{U_\Psi} \begin{bmatrix} \boldsymbol{\Sigma_\Psi} & 0 \end{bmatrix} \boldsymbol{V'_\Psi}$, number of iterations Iter, constant c, Welch bound μ_{welch}.

Output: Measurement matrix $\boldsymbol{\Phi}_{opt}$.

Initialization: Initialize $\boldsymbol{\Phi}_0 \in R^{M \times N}$ to a random matrix, initial $\boldsymbol{U}_Z \in R^{M \times M}$ to a unitary matrix.

For $l = 1$ to *Iter* do

1. Compute the sensing matrix $\boldsymbol{D} = \boldsymbol{\Phi}_l \boldsymbol{\Psi}$ and normalize the columns in $\boldsymbol{D}$.
2. Compute Gram matrix $\boldsymbol{G} = \boldsymbol{D'D}$.
3. Shrink $\boldsymbol{G}$ and obtain $\boldsymbol{G}_t$ by

$$\boldsymbol{G}_t(i,j) = \begin{cases} sign\left(g_{ij}\right)(\mu_{welch} + c), & \left|g_{ij}\right| \geq \mu_{welch} + c \\ sign\left(g_{ij}\right)\mu_{welch}, & \mu_{welch} \leq \left|g_{ij}\right| < \mu_{welch} + c \\ g_{ij}, & \left|g_{ij}\right| < \mu_{welch} \end{cases}$$

4. Apply eigenvalue decomposition to $\boldsymbol{G}_t$ and obtain $\boldsymbol{G}_t = \boldsymbol{P\Lambda P'}$.
5. Compute the average of the top M diagonal entries in $\boldsymbol{\Lambda}$, denoted as $\hat{\lambda}$.
6. Construct $\boldsymbol{\Lambda}_M \in R^{M \times M}$ as $\boldsymbol{\Lambda}_M = diag(\hat{\lambda}, \hat{\lambda} \cdots \hat{\lambda})$.
7. Update $\boldsymbol{\Phi}$ by $\boldsymbol{\Phi}_l = \boldsymbol{U}_Z \begin{bmatrix} (\boldsymbol{\Lambda}_M)^{\frac{1}{2}} & 0 \end{bmatrix} \boldsymbol{P'V_\Psi} \begin{bmatrix} \boldsymbol{\Sigma}_\Psi^{-1} & 0 \end{bmatrix}' \boldsymbol{U'_\Psi}$.

end

return $\boldsymbol{\Phi}_{Iter}$

Noting that $\boldsymbol{G}_t$ plays an important role in measurement matrix optimization, Algorithm 1 takes μ_{ave}, μ_{max} and μ_{all} into consideration simultaneously when designing the $\boldsymbol{G}_t$. By minimizing (5), the Gram matrix is most similar to $\boldsymbol{G}_t$ in terms of Frobenius norm, leading to maintain the advantage of $\boldsymbol{G}_t$ in reducing the mutual coherence indexes. Therefore, Algorithm 1 is effective in reducing μ_{ave}, μ_{max}, and μ_{all}.

In the shrinkage function, a different threshold yields different results. Inspired by [22], we propose a shrinkage function shown as (7) which has a new threshold $\mu_{welch} + c$. We have not derived the optimal value of c in theory, but setting c to a proper value can also lead to a moderate result.

After averaging the eigenvalues of $\boldsymbol{G}_{t\_M}$, the first term on the right part of (20) is minimized. However, the diagonal entries of $\boldsymbol{G}_{t\_M}$ change accordingly. Hence, we can't assure that $\hat{\mu}_{all}$ reaches the minima, that is to say, $\boldsymbol{G}_{t\_opt}$ may not be the optimal solution in terms of μ_{all}. Fortunately, we find that the change of $\sum_{i=1}^{M} \lambda_i^2$ is much greater than that of $\sum_{i=1}^{L} \hat{g}_{ii}^2$ in (20), which means our approach is effective in reducing μ_{all}.

The proposed algorithm is an iterative one. The main complexity of Algorithm 1 for each iteration is located at steps 1, 2, 4, and 7. For those steps, the flops required are $O(MNL)$, $O(ML^2)$, $O(L^3)$, and $O(L^3)$ respectively. Hence, the complexity of Algorithm 1 is approximate to be $O(IterL^3)$. Since the complexity for similar algorithms in [8,16,17] which apply eigenvalue decomposition or SVD is no less than $O(IterL^3)$, the proposed algorithm has not increased the complexity significantly.

4. Simulation Results and Discussion

In this section, we conduct simulations to predetermine a suitable c firstly. Then, we examine the mutual coherence indexes and reconstruction performance of the proposed method and compare them with the well-established similar algorithms given in [8,16,17] by presenting the empirical results. Last, we verify the effectiveness of our method with various measurement matrices and dictionary matrices. The iteration number Iter is set to 100 and t is set to μ_{welch}. For a given dictionary matrix $\boldsymbol{\Psi} \in R^{80 \times 120}$, $\boldsymbol{x} \in R^{120 \times 1}$ has a sparse representation as $\boldsymbol{x} = \boldsymbol{\Psi}\boldsymbol{s}$ where $\boldsymbol{s}$ is K-sparse and each non-zero entry is randomly positioned with a Gaussian distribution of i.i.d. zero-mean and unit variance. Orthogonal Matching Pursuit (OMP) [23] algorithm is employed in signal reconstruction. Denote

$\varepsilon = \|x_e - x\|_2 / \|x\|_2$ the reconstruction error where x_e is the reconstructed signal. The reconstruction is identified as a success, called exact reconstruction, provided $\varepsilon \leq 10^{-6}$. Denote P_{suc} the percentage of successful reconstruction. In Sections 4.1–4.3, Φ_0 and Ψ are both Gaussian random matrices.

4.1. The Choice of c

Since the analytical solution of c is extremely difficult, here, we conduct a serious of simulations to find a suitable c. Figure 1 illustrates the change tendency of mutual coherence indexes and P_{suc} with argument c. We fix the row number of M to 28, the sparsity to 8, and varies c from 0 to 0.16. The experiment is performed for 1000 random sparse ensembles and the results are recorded.

When $c = 0$, the shrinkage function shown as (7) is the same as (6). As can be seen from the graphs, when c increases, μ_{ave} increases, μ_{max} and μ_{all} decrease first and then increase, P_{suc} increases firstly and then decreases. μ_{max} and μ_{all} reach the minima when $c = 0.02$ and $c = 0.03$ respectively. It is worth noting that appropriate increase of c leads to decrease of μ_{max} and μ_{all} but increase of μ_{ave}. When $c = 0.01$, better μ_{max} and μ_{all} are obtained and the loss in μ_{ave} is tolerable. Moreover, P_{suc} reaches the maxima. Therefore, 0.01 may be a moderate value for c and c is set to 0.01 in the simulations in Sections 4.2–4.4.

4.2. Comparing the Mutual Coherence Indexes

This section presents a series of simulations to compare our method with algorithms given in [8,16,17] on the three mutual coherence indexes of D obtained by $D = \Phi_{opt}\Psi$ where Φ_{opt} is the optimized measurement matrix. For convenience, each method is denoted as Propose, Elad, Hong, and Entezari. The down-scaling factor for Elad is set to 0.95. The inner iteration number for Hong is set to 2, which means K-SVD is applied twice in every updating of Φ. The point is set to 0.5 to update the G_t in Entezari.

Figure 2 illustrates the change tendency of mutual coherence indexes with iteration number for $M = 28$. As can be seen from the figure, the indexes corresponding to different algorithms all change monotonously with the iteration number. When μ_{max} and μ_{ave} converge, the number of iterations required by our method is almost equal to that of Hong and significantly less than that of Elad. When μ_{all} converges, the number of iterations required by our method is equivalent to that of Entezari and significantly less than that of Hong and Elad.

Figure 3 presents the histogram of the absolute off-diagonal values of $\left(\Phi_{opt}\Psi\right)'\Phi_{opt}\Psi$ for $M = 28$. It is seen from the figure that Elad and Entezari have long tails, showing that the number of off-diagonal values that exceed 0.34 is relatively large. The tail of Hong is shorter than that of Elad and Entezari, and reaches the maximum of 0.34. Compared with Hong, our method has a shorter tail which reaches the maximum of 0.32 and has more off-diagonal values below the μ_{welch} (0.1662).

In conclusion, while effectively reducing μ_{max} and μ_{all}, our method can maintain a small μ_{ave} at the same time. Additionally, the number of iterations required for the convergence of each index of our method is significantly less than that of Elad. Therefore, from the view of mutual coherence indexes, the measurement matrix obtained by our method has better properties than the other three methods. This coincides with the theoretical results obtained in the Section 3.

4.3. Comparing the Reconstruction Performance

Case 1. *Comparison of the P_{suc} in the noiseless case.*

In this case, we conduct two separate CS experiments, first by fixing $K = 8$ and varying M from 12 to 44 and second by fixing $M = 28$ and varying K from 4 to 20. Each experiment is performed for 1000 random sparse ensembles and the number of exact reconstruction is recorded.

Figures 4 and 5 reveal that the P_{suc} of our method is the highest, which indicates its superiority over the other three methods.
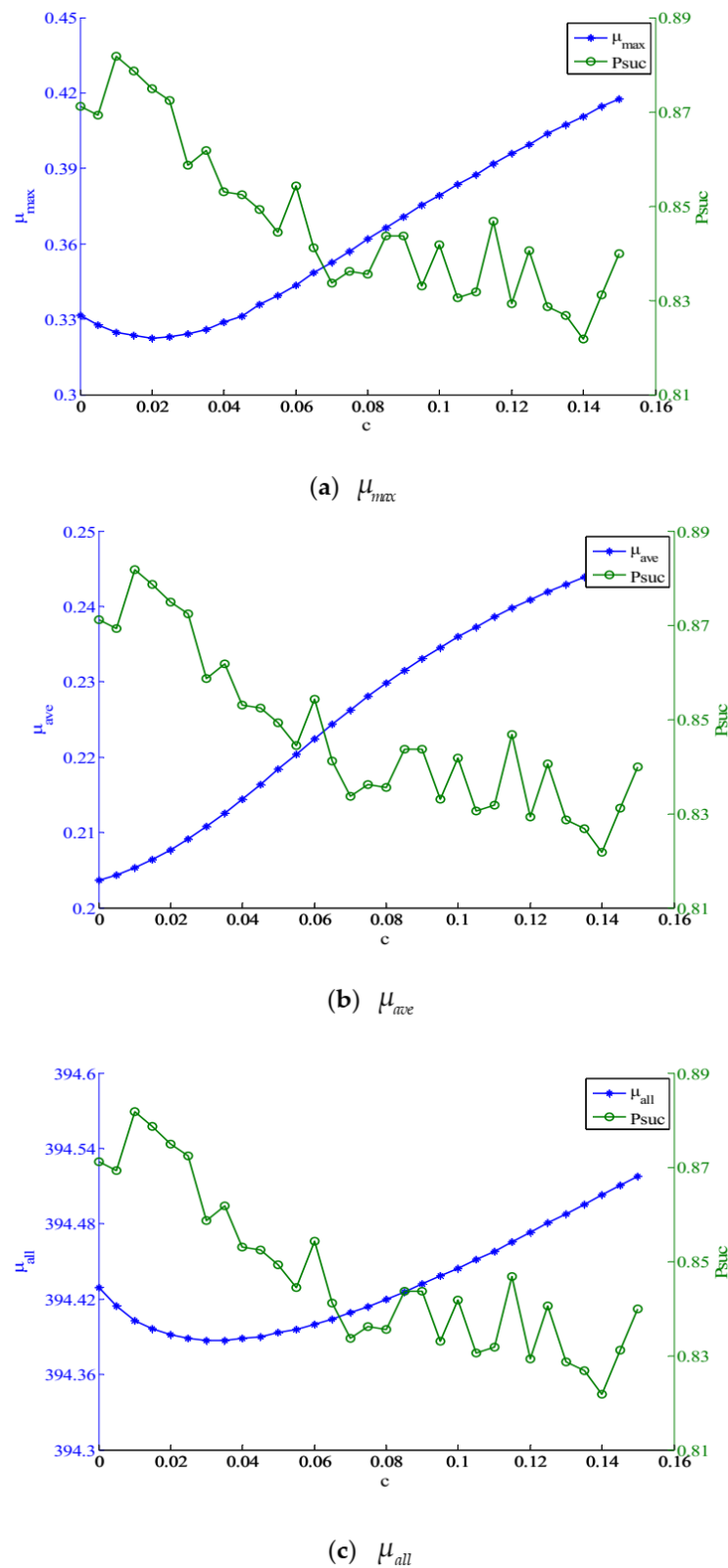

(a) μ_{max}

(b) μ_{ave}

(c) μ_{all}

Figure 1. (a) μ_{max} and P_{suc} versus c; (b) μ_{ave} and P_{suc} versus c; (c) μ_{all} and P_{suc} versus c, both with $M = 28$ and $K = 8$.

(**a**) μ_{max}

(**b**) μ_{ave}
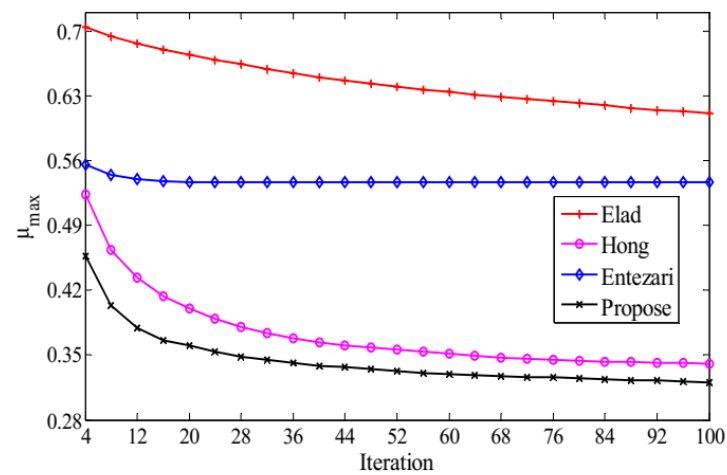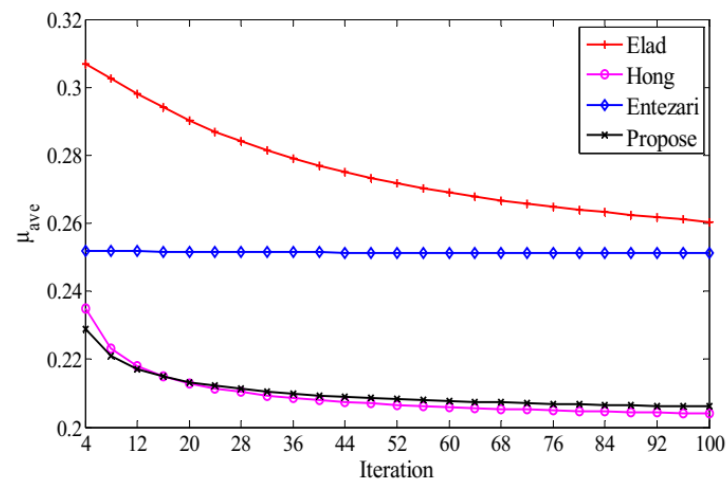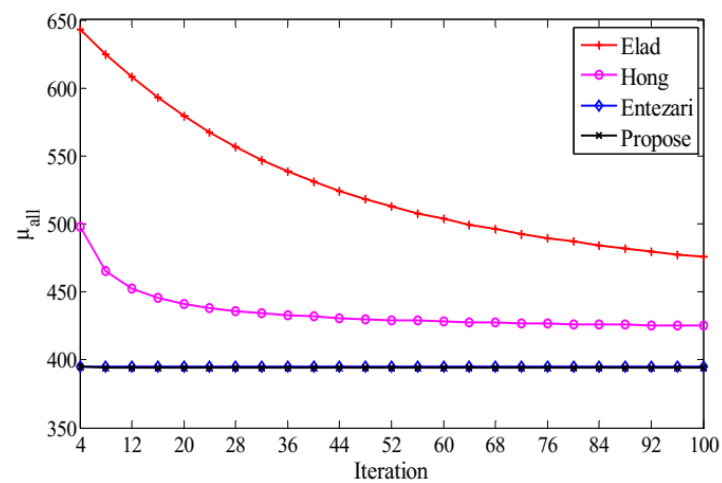
(**c**) μ_{all}

Figure 2. The convergence results: (**a**) evolution of μ_{max}, (**b**) μ_{ave} and (**c**) the evolution of μ_{all}, all versus iteration number, where $M = 28$.
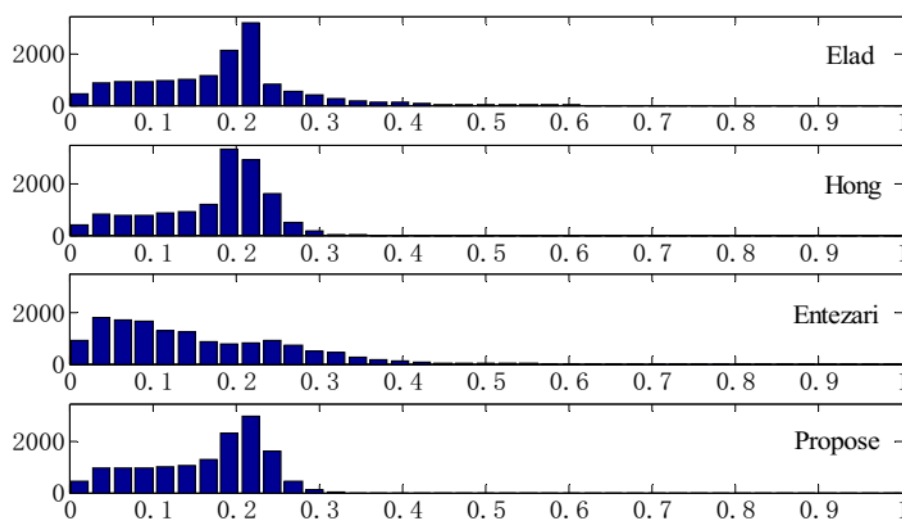
Figure 3. Histogram of the absolute off-diagonal values of $\left(\mathbf{\Phi}_{opt}\mathbf{\Psi}\right)'\mathbf{\Phi}_{opt}\mathbf{\Psi}$ for $M=28$.

Table 1. μ_{max} by Elad, Hong, Entezari, and proposed method versus measurement dimension M.

M	Elad	Hong	Entezari	Propose	μ_{welch}
12	0.8971	0.8884	0.8465	0.8072	0.2750
16	0.8157	0.6488	0.7517	0.5211	0.2337
20	0.7610	0.4625	0.6710	0.4218	0.2050
24	0.6972	0.3920	0.5874	0.3618	0.1833
28	0.6186	0.3403	0.5351	0.3223	0.1662
32	0.5460	0.3069	0.4911	0.2919	0.1520
36	0.4879	0.2804	0.4472	0.2659	0.1400
40	0.4454	0.2591	0.4177	0.2487	0.1296
44	0.3802	0.2404	0.3927	0.2302	0.1205

Table 2. μ_{ave} by Elad, Hong, Entezari, and proposed method versus measurement dimension M.

M	Elad	Hong	Entezari	Propose	μ_{welch}
12	0.4168	0.4528	0.4141	0.3608	0.2750
16	0.3507	0.2964	0.3532	0.2979	0.2337
20	0.3016	0.2566	0.3111	0.2581	0.2050
24	0.2431	0.2272	0.2772	0.2288	0.1833
28	0.2212	0.2042	0.2512	0.2060	0.1662
32	0.2056	0.1854	0.2302	0.1879	0.1520
36	0.1937	0.1700	0.2118	0.1727	0.1400
40	0.1841	0.1567	0.1961	0.1600	0.1296
44	0.1760	0.1454	0.1824	0.1491	0.1205

Table 3. μ_{all} by Elad, Hong, Entezari, and proposed method versus measurement dimension M.

M	Elad	Hong	Entezari	Propose
12	1220.20	2638.47	1081.18	1080.19
16	898.80	802.77	780.92	780.16
20	721.41	630.91	605.72	600.13
24	571.34	512.24	480.67	480.12
28	474.76	425.25	394.86	394.39
32	404.35	358.49	330.55	330.11
36	353.96	307.11	280.55	280.10
40	315.03	265.40	240.55	240.10
44	285.51	231.43	207.79	207.38
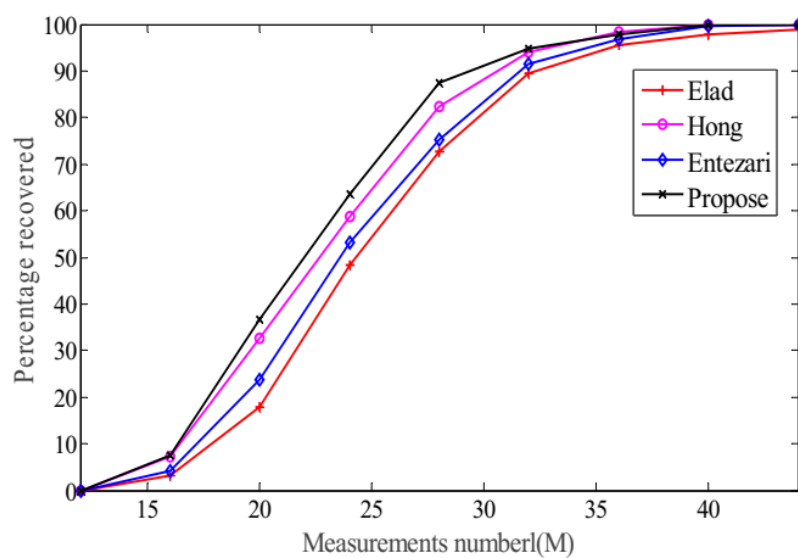
Figure 4. The change tendency of P_{suc} with M while $K = 8$ in the noiseless case.
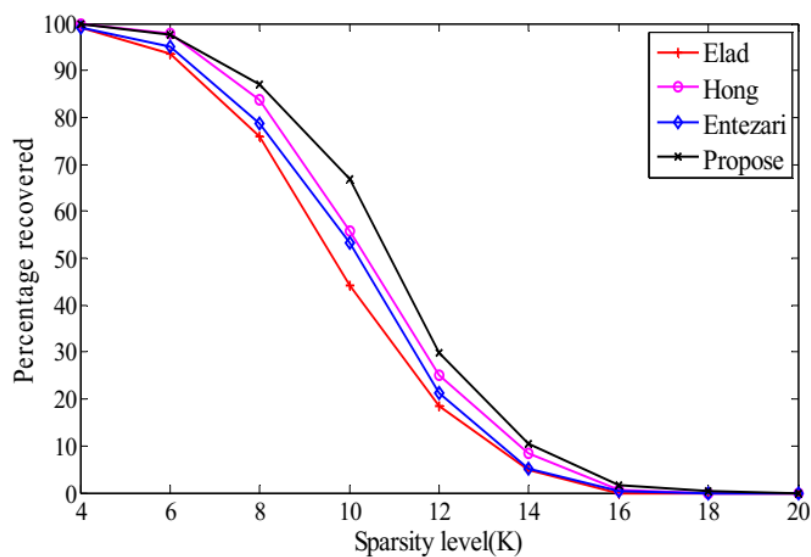
Figure 5. The change tendency of P_{suc} with K while $M = 28$ in the noiseless case.

Case 2. *Comparison of the ε in the noisy case.*

To show the robustness of the proposed method in noisy cases we consider the noisy model $y = \mathbf{\Phi}x + v$ where v is the vector of additive Gaussian noise with zero means. We conduct the experiment by fixing $M = 28$, $K = 8$, and varying SNR from 10 to 50 dB. The experiment is performed for 1000 random sparse ensembles and the average reconstruction error is recorded. From Figure 6, we can see that the reconstruction errors decrease with the increase of SNR, and the error of the proposed method is smaller than that of the others.
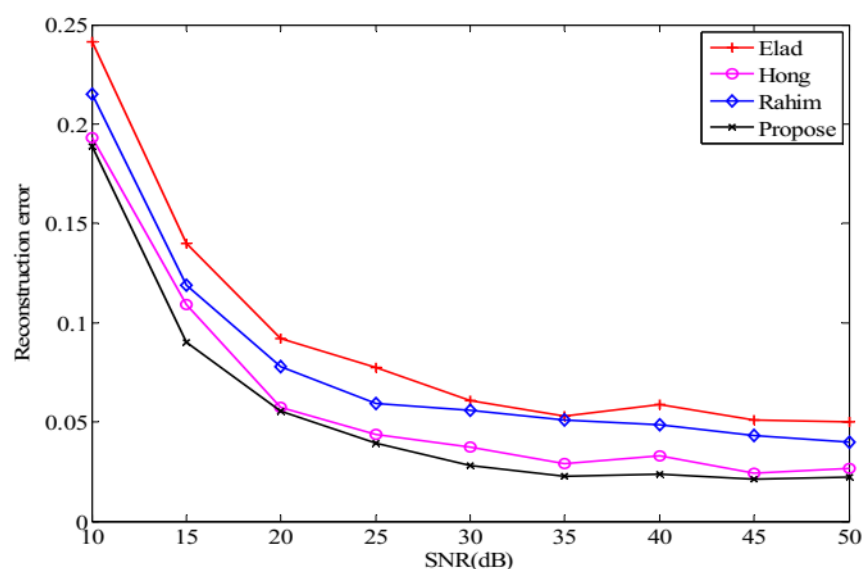
Figure 6. The change tendency of ε with SNR while $M = 40$ and $K = 8$.

Table 3 presents that μ_{all} of the Entezari is slightly larger than that of our method. It is interesting to note that the number of off-diagonal entries with smaller absolute values in the Entezari is significantly larger than that of our method from Figure 3. Moreover, it can be seen from Table 2 that μ_{ave} of Hong is slightly lower than that of our method. However, the simulation results show that our method outperforms the others in terms of reconstruction performance. It is also worthy noting that our method reduces μ_{ave}, μ_{max}, and μ_{all} simultaneously, leading to better reconstruction performance in CS. This implies that a single mutual coherence index cannot accurately reflect the actual performance of the methods, and verifies the necessity of using multiple indexes simultaneously in measurement matrix optimization.

4.4. Different Kinds of $\mathbf{\Phi}$ and $\mathbf{\Psi}$ Optimized by the Proposed Methods

To analyze the performance of our method with various measurement matrices and dictionary matrices, a serious of simulations are carried out in this section. We choose the measurement matrix as a Gaussian random matrix and a Bernoulli random matrix, and choose the dictionary matrix as a Gaussian random matrix and the DCT matrix, respectively. We compare the mutual coherence indexes and the reconstruction performance before and after optimization. When $\mathbf{\Psi}$ is the Gaussian random matrix, $\mathbf{\Phi}$ belongs to $R^{M \times 80}$ and $\mathbf{\Psi}$ belongs to $R^{80 \times 120}$. When $\mathbf{\Psi}$ is the DCT matrix, $\mathbf{\Phi}$ belongs to $R^{M \times 120}$ and $\mathbf{\Psi}$ belongs to $R^{120 \times 120}$. Each experiment is performed for 1000 random sparse ensembles.

The mutual coherence indexes of different measurement matrices $\mathbf{\Phi}$ with different dictionary matrices $\mathbf{\Psi}$ are shown in Figure 7. As seen from the simulations, all the optimized measurement matrices produce smaller μ_{max}, μ_{ave}, and μ_{all} than the random ones.
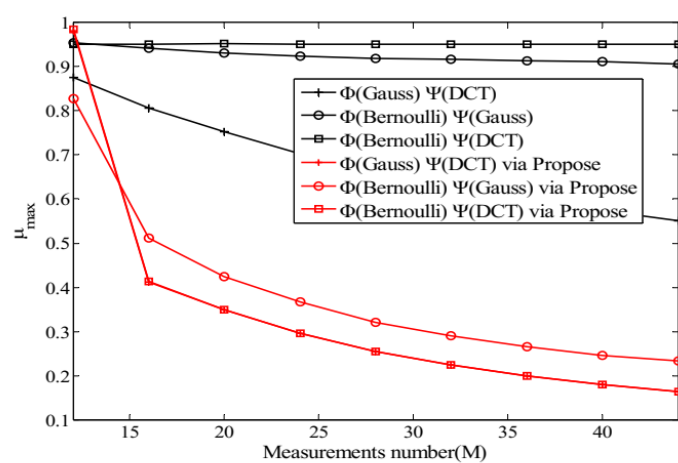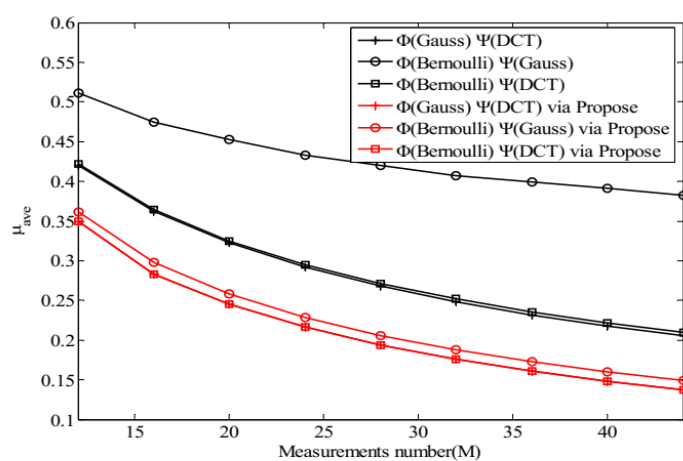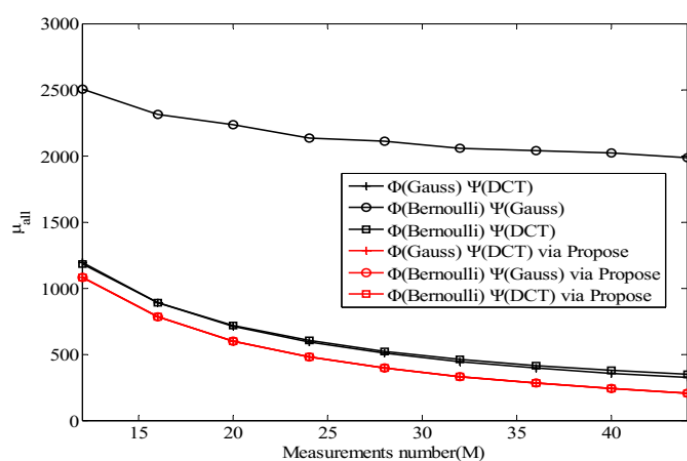
(**a**) μ_{max}

(**b**) μ_{ave}

(**c**) μ_{all}

Figure 7. The evolution of (**a**) μ_{max}, (**b**) μ_{ave} and (**c**) μ_{all}, all versus measurements number with different *D*.

Figures 8 and 9 present the reconstruction performance of OMP with the optimized measurement matrices and the random ones. It is seen from the graphs in these figures that all the optimized matrices outperform the random ones in terms of the percentage of exact reconstruction.
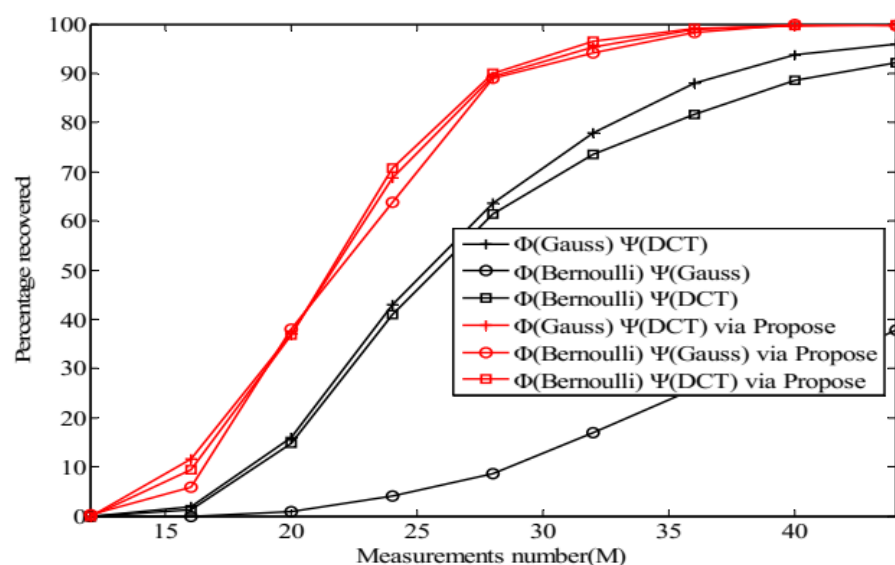
Figure 8. The change tendency of P_{suc} with M while $K = 8$ in the noiseless case.
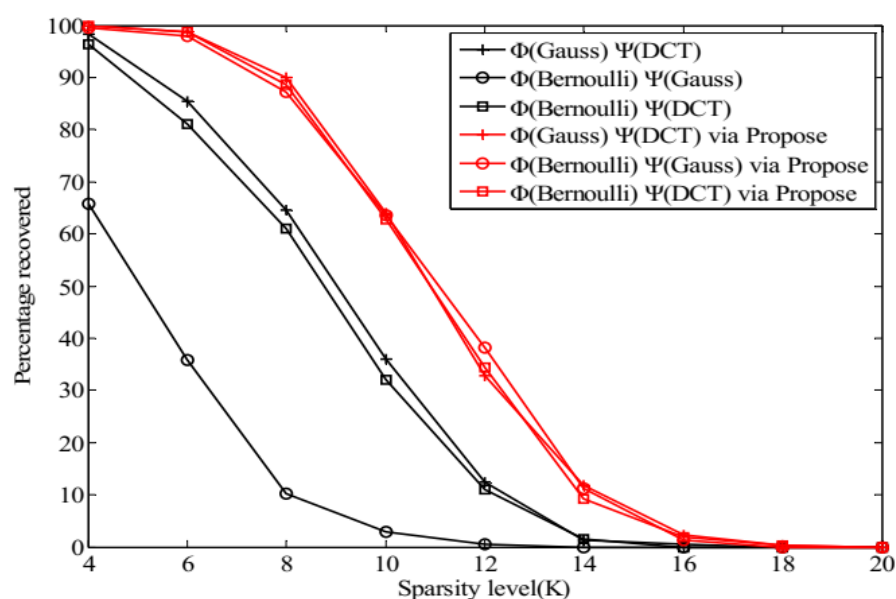
Figure 9. The change tendency of P_{suc} with K while $M = 28$ in the noiseless case.

5. Conclusions

This paper focused on the optimization of measurement matrix for compressed sensing. To decrease μ_{max}, μ_{ave}, and μ_{all} simultaneously, we designed a new target Gram matrix which was obtained by applying a new shrinkage function to the Gram matrix and updated by performing rank reduction and eigenvalue averaging. Then, we characterized the analytical solutions of the measurement matrix by SVD. Based on alternating minimization, we proposed an iterative method to optimize the measurement matrix. The simulation results show that the proposed method reduces μ_{max}, μ_{ave}, and μ_{all} simultaneously and outperforms the existing algorithms in terms of reconstruction performance. In addition,

the proposed method is computationally less expensive than some existing algorithms in the literature.

As detailed, we gave the optimal value of c under a fixed matrix scale through simulation. When the scale changes, the value of c in Section 4.1 may no longer be applicable. Therefore, it is meaningful to find the theoretical 'optimal value' of c. Furthermore, noting that lower mutual coherence indexes mean potentially higher reconstruction performance, further efforts are needed to decrease the indexes simultaneously.

References

1. Donoho, D.L. Compressed sensing. *IEEE Trans. Inf. Theory* **2006**, *52*, 1289–1306. [CrossRef]
2. Xie, Y.; Yu, J.; Guo, S.; Ding, Q.; Wang, E. Image Encryption Scheme with Compressed Sensing Based on New Three-Dimensional Chaotic System. *Entropy* **2019**, *21*, 819. [CrossRef]
3. Fang, Y.; Li, L.; Li, Y.; Peng, H.; Yang, Y. Low Energy Consumption Compressed Spectrum Sensing Based on Channel Energy Reconstruction in Cognitive Radio Network. *Sensors* **2020**, *20*, 1264. [CrossRef]
4. Martinez, J.A.; Ruiz, P.M.; Skarmeta, A.F. Evaluation of the Use of Compressed Sensing in Data Harvesting for Vehicular Sensor Networks. *Sensors* **2020**, *20*, 1434. [CrossRef]
5. Donoho, D.L.; Elad, M. Optimally sparse representation in general (nonorthogonal) dictionaries via 1 minimization. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 2197–2202. [CrossRef]
6. Candes, E.J.; Tao, T. Decoding by linear programming. *IEEE Trans. Inf. Theory* **2005**, *51*, 4203–4215. [CrossRef]
7. Stark, D.P.B. Uncertainty Principles and Signal Recovery. *SIAM J. Appl. Math.* **1989**, *49*, 906–931.
8. Elad, M. Optimized Projections for Compressed Sensing. *IEEE Trans. Signal Process.* **2007**, *55*, 5695–5702. [CrossRef]
9. Shaohai, H.U. An Optimization Method for Measurement Matrix Based on Eigenvalue Decomposition. *Signal Process.* **2012**. [CrossRef]
10. Yan, W.; Wang, Q.; Shen, Y. Shrinkage-Based Alternating Projection Algorithm for Efficient Measurement Matrix Construction in Compressive Sensing. *IEEE Trans. Instrum. Meas.* **2014**, *63*, 1073–1084. [CrossRef]
11. Duartecarvajalino, J.M.; Sapiro, G. Learning to Sense Sparse Signals: Simultaneous Sensing Matrix and Sparsifying Dictionary Optimization. *IEEE Trans. Image Process.* **2009**, *18*, 1395–1408. [CrossRef]
12. Lu, C.; Li, H.; Lin, Z. Optimized Projections for Compressed Sensing via Direct Mutual Coherence Minimization. *Signal Process.* **2018**, *151*, 45–55. [CrossRef]
13. Xu, J.; Pi, Y.; Cao, Z. Optimized projection matrix for compressive sensing. *EURASIP J. Adv. Signal Process.* **2010**, *2010*, 560349. [CrossRef]
14. Abolghasemi, V.; Ferdowsi, S.; Sanei, S. A gradient-based alternating minimization approach for optimization of the measurement matrix in compressive sensing. *Signal Process.* **2012**, *92*, 999–1009. [CrossRef]
15. Zheng, H.; Li, Z.; Huang, Y. An Optimization Method for CS Projection Matrix Based on Quasi-Newton Method. *Acta Electron. Sin.* **2014**, *42*, 1977–1982.
16. Hong, T.; Bai, H.; Li, S.; Zhu, Z. An efficient algorithm for designing projection matrix in compressive sensing based on alternating optimization. *Signal Process.* **2016**, *125*, 9–20. [CrossRef]
17. Entezari, R.; Rashidi, A. Measurement matrix optimization based on incoherent unit norm tight frame. *AEU Int. J. Electron. Commun.* **2017**, *82*, 321–326. [CrossRef]
18. Sustik, M.A.; Tropp, J.A.; Dhillon, I.S.; Heath, R.W. On the existence of equiangular tight frames. *Linear Algebra Its Appl.* **2007**, *426*, 619–635. [CrossRef]
19. Welch, L. Lower bounds on the maximum cross correlation of signals (Corresp.). *IEEE Trans. Inf. Theory* **1974**, *20*, 397–399. [CrossRef]
20. Aharon, M.; Elad, M.; Bruckstein, A.M. K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation. *IEEE Trans. Signal Process.* **2006**, *54*, 4311–4322. [CrossRef]

21. Fickus, M.; Mixon, D.G. Tables of the existence of equiangular tight frames. *arXiv* **2015**, arXiv:1504.00253.
22. Li, G.; Zhu, Z.; Yang, D.; Chang, L.; Bai, H. On Projection Matrix Optimization for Compressive Sensing Systems. *IEEE Trans. Signal Process.* **2013**, *61*, 2887–2898. [CrossRef]
23. Tropp, J.A.; Gilbert, A.C. Signal Recovery from Random Measurements Via Orthogonal Matching Pursuit. *IEEE Trans. Inf. Theory* **2007**, *53*, 4655–4666. [CrossRef]