Proceedings of the 32nd Chinese Control Conference
July 26-28, 2013, Xi'an, China

Projection Matrix Optimization Based on SVD for Compressive Sensing Systems

Qiuwei Li, Zhihui Zhu, Si Tang, Liping Chang and Gang Li
Zhejiang Provincial Key Laboratory for Signal Processing
College of Information Engineering
Zhejiang University of Technology, Hangzhou, Zhejiang, P.R. China
E-mail: liqiuweiss@gmail.com

Abstract: Sparse signals can be sensed with a reduced number of projections and then reconstructed if compressive sensing (CS) is employed. Traditionally, the projection matrix is chosen as a random matrix, but a projection sensing matrix that is optimally designed for a certain class of signals can further improve the reconstruction accuracy or further reduce the necessary number of measurement samples. This paper considers the problem of designing the projection matrix Φ for a compressive sensing system in which the dictionary Ψ is assumed to be given. The optimal projection matrix design is formulated in terms of finding those Φ such that the *Frobenius* norm of the difference between the Gram matrix of the equivalent dictionary $\Phi\Psi$ and the identity matrix is minimized. A novel algorithm based on SVD for optimal projection matrix searching is proposed to solve the corresponding minimization problem. Simulation results reveal that the signal recovery performance of sensing matrix obtained by proposed algorithm surpasses that of other standard sensing matrix designs.

Key Words: Compressed sensing, averaged coherence, optimization techniques

1 Introduction

Compressive sensing (CS) is an emerging framework, which states that sparse signals, that is, signals that have a concise linear representation on an appropriate dictionary, can be exactly recovered from a number of linear projections of dimension considerably lower than the number of samples required by the Shannon-Nyquist Theorem [1] - [3]. Over the last few years researchers have derived some new compressive sensing theory for a variety of structured sensing matrices [4]. In the same time, some progress has been made in making statements on the pursuit performance for optimized dictionaries [5] - [6].

In one the hand, given a signal vector $y \in \Re^N$ we consider measurement systems that acquire M linear measurements. We can represent this process mathematically as

$$z = \Phi y \tag{1}$$

where $\Phi \in \Re^{M\times N}$ and $z \in \Re^M$. The matrix Φ, called *measurement matrix* or *sensing matrix*, represents a *dimensionality reduction*, i.e., it maps $\Re^N$ into $\Re^M$, where M is typically much smaller than N. Note that in the standard CS framework, we assume that the measurements are *non-adaptive*, meaning that the rows of Φ are fixed in advance and do not depend on the previously acquired measurements. In this paper, we will optimize the *measurement matrix* or *sensing matrix* to obtain significant performance gains.

In the other hand, a given signal $y \in \Re^N$ can often be expressed as a linear combination of a small number of signals taken from a "resource" database, which is called the dictionary. Elements of the dictionary are typically unit norm functions called atoms. Let us denote the dictionary as Ψ, and the atoms as $\psi_k, k = 1, \cdots, L$, that is, $\Psi \triangleq \begin{bmatrix} \psi_1 & \psi_2 & \cdots & \psi_L \end{bmatrix} \in \Re^{N\times L}$, where L is the size of the dictionary. The dictionary is over-complete ($L > N$) when it spans the signal space and its atoms are linearly dependent. In that case, every signal can be represented as a linear combination of atoms in the dictionary :

$$y = \sum_{k=1}^{L} s_k \psi_k \triangleq \Psi s \tag{2}$$

where $s \triangleq \begin{bmatrix} s_1 & s_2 & \cdots & s_L \end{bmatrix}^{\mathcal{T}}$ is a coefficient vector that represents y in dictionary Ψ. A signal is said to be *compressible* if most of the coefficients of s are zero or they can be discarded without much loss of information of the signal. Denote $\bar{s}$ as the signal where only the K largest coefficients of s are kept and the rest are set to zero, and thus we have a new signal

$$\bar{y} \triangleq \Psi \bar{s}$$

If the value of the coefficients (sorted in decreasing order), decrease quickly, then y is well approximated by $\bar{y}$, when properly selecting both s and the dictionary Ψ.

By substituting y in (1) with (2), z can be rewritten as

$$z = \Phi\Psi s \triangleq As \tag{3}$$

where the matrix $A = \Phi\Psi = \begin{bmatrix} A_1 & A_2 & \cdots & A_L \end{bmatrix} \in \Re^{M\times L}$ is sometimes referred to as *equivalent dictionary* of the CS system. As $A \in \Re^{M\times L}$ with $M << L$, the *equivalent dictionary*A is over-complete. Thus for given measurement vector z and *equivalent dictionary*A, the coefficients vector s tends to be not unique. This is where the sparsity constraint comes into play.

The spark of a given matrix A, denoted as $spark(A)$, is defined as the smallest number of columns in A that are linearly dependent. This property of matrices for the study of the uniqueness of sparse solutions is of great importance. The $spark$ gives a simple criterion for uniqueness of sparse solutions. By definition, the vectors in the null-space of the matrix

$$As = \mathbf{0}$$

must satisfy

$$||s||_0 \geq spark(A) \tag{4}$$

where $||s||_0$ denotes the number of non-zero elements of vector s. Since these vectors combine lineally columns from A to generate the zero vector, and at least $spark(A)$ such columns are necessary by definition.

In another words, for a given z and A there exists at most one s such that $z = As$ if and only if

$$spark(A) > 2||s||_0 \quad (5)$$

As seen from the above, the larger the $spark$ of A, the bigger the signal space among which the CS systems can guarantee an exact recovery. For a given dictionary Ψ, the $spark$ of the equivalent dictionary A is determined by the sensing matrix Φ. It would be of great interest to design Φ such that $spark(A)$ is maximized.

One concludes that any K-sparse signal $y_0 = \Psi s_0$ can be uniquely determined from the measurement z given by

$$z = \Phi y_0$$

and that s_0 is the solution of the following constrained problem:

$$s_0 = arg \min_s ||s||_0 \quad s.t. \quad z = As \quad (6)$$

as long as (5) holds, that is, $spark(A) > 2K$.

Such a problem can be attacked by a number of algorithms which are classified into two groups. The first group includes greedy algorithms such as the matching pursuit (MP) and the orthogonal MP (OMP), which iteratively select locally optimal basis vectors.

In the second group, the algorithms are based on convex relaxation methods such as the basis pursuit (BP) or least absolute shrinkage and selection operator (LASSO), which solve the following problem:

$$s_0 = arg \min_s ||s||_1 \quad s.t. \quad z = As \quad (7)$$

The replacement of $|| \, ||_0$ by $|| \, ||_1$, the l_1 norm, converts the non-convex problem to a convex one.

Moreover, there are alternative properties of A that can be easily exploited to provide recovery guarantees. One of such properties is the *mutual coherence* of a matrix [5], which will be introduced in the next section.

The main objective and contribution of this paper are:

- **Objective**: in order to improve the performance of compressed sensing (signal reconstruction accuracy), we consider the problem of designing the projection matrix Φ for a compressive sensing system in which the dictionary Ψ is given. The optimal projection matrix design is formulated in terms of finding those Φ such that the *Frobenius* norm of the difference between the Gram matrix of the equivalent dictionary $\Phi\Psi$ and the identity matrix is minimized.
- **Contribution**: we investigate the problem of projection matrix design for sensing signals which are sparse in over-complete dictionaries and a novel algorithm based on SVD for optimal projection matrix searching is proposed in this paper. Experiments are given to show that the sensing matrix obtained using our proposed algorithm outperforms others in signal reconstruction accuracy.

The outline of this paper is arranged as follows. In Section 2, we define the criteria that is used for measuring the coherence of a matrix. In Section 3, some related work on the sensing matrix optimization problem is provided, and an iterative algorithm based on SVD for optimal projection matrix searching is derived to find an optimal sensing matrix that minimizes the coherence of the equivalent dictionary. Experiments are carried out in Section 4 to analyze the proposed algorithm. Simulations are also presented in Section 5 to show the effectiveness of our proposed method in improving signal reconstruction accuracy. Some concluding remarks are given in Section 6 to end this paper.

2 Preliminaries

The mutual coherence, denoted as $\mu(A)$, represents the worst-case coherence between any two columns (atoms) of A and is one of the most fundamental quantities associated with CS theory. As shown in [5], any K-sparse signal s_0 can be exactly recovered from the observation/measurement via

$$s_0 = arg \min_s ||s||_0 \quad s.t. \quad z = As$$

as long as

$$K < \frac{1}{2}[1 + \frac{1}{\mu(A)}] \quad (8)$$

The mutual coherence of this matrix is defined as

$$\mu(A) \triangleq \max_{1 \le i \ne j \le L} \frac{|A_i^{\mathcal{T}} A_j|}{\|A_i\|_2 \|A_j\|_2} \quad (9)$$

which measures the maximum linear dependency possibly achieved by any two columns of matrix A. Coherence is a blunt instrument since it only reflects the most extreme correlation in the matrix. Nevertheless, it is easy to calculate and it captures well the behavior of uniform matrices. Thus it can be used as a criteria of the CS system.

The (i, j)th element of the Gram matrix of A is defined as

$$g_{ij} \triangleq A_i^{\mathcal{T}} A_j$$

and

$$S_c \triangleq diag(g_{11}^{-1/2} \;\; \cdots \;\; g_{kk}^{-1/2} \;\; \cdots \;\; g_{LL}^{-1/2})$$

Thus the Gram matrix of $\bar{A} \triangleq AS_c$, denoted as $\bar{G} = \{\bar{g}_{ij}\}$, is normalized such that $\bar{g}_{kk} = 1, \; \forall \; k$. Obviously,

$$\mu(A) = \max_{i \ne j} |\bar{g}_{ij}|$$

It can be shown [9] that for a matrix $A \in \Re^{M \times N}$, $\mu(A)$ is bounded with

$$\underline{\mu} \le \mu(A) \le 1 \quad (10)$$

with the low bound given by

$$\underline{\mu} \triangleq \sqrt{\frac{L - M}{M(L-1)}} \quad (11)$$

Simulations have shown that the performance of an CS system is more related to the averaged mutual coherence, defined as [6]

$$\mu_t(A) \triangleq \frac{\sum_{\forall \; (i,j) \in \mathcal{S}_t} |\bar{g}_{ij}|}{N_t} \quad (12)$$

where $\mathcal{S}_t \triangleq \{(i,j) : |\bar{g}_{ij}| \geq t\}$ with $0 \leq t < 1$ a given number and N_t is the number of elements in the index set $\mathcal{S}_t$. It should be pointed out that $\mu_t(A)$ was used as an indicator of convergence for a proposed iteration procedure but not minimized directly in [6].

3 Sensing Matrix Optimization

In this section, we first provide some related work on the sensing matrix optimization problem, and then an iterative algorithm based on SVD for optimal projection matrix searching is derived to find an optimal sensing matrix that minimizes the coherence of the equivalent dictionary.

3.1 Related work

This subsection contains a brief survey of the important results in optimization of the projection matrix.

Elad proposed the first work of the optimal design of sensing matrix Φ in [6]. It is due to the fact that (8) is just a worst-case bound and can not reflect the average signal recovery performance that, instead of $\mu(A)$, an averaged mutual coherence, denoted as $\mu_t(A)$, was dealt with in [6]. Simulation results showed that the optimized sensing matrix with the procedure of reducing outperforms the one generated randomly in terms of signal recovery accuracy.

In [7] Duarte-Carvajalino and Sapiro produced a approach to learning the projection matrix for a given dictionary as

$$\min_{\Phi} ||\Pi_d^2 - \Pi_d^2 U_d^{\mathcal{T}} \Phi^{\mathcal{T}} \Phi U_d \Pi_d^2||_F^2 \tag{13}$$

where $|| \ ||_F$ denotes the *Frobenius* norm and $\Psi = U_d \begin{bmatrix} \Pi_d & \mathbf{0} \end{bmatrix} V_d^{\mathcal{T}}$ is a singular value decomposition (SVD) of the dictionary Ψ. The numerical procedure, though not globally optimal, was reported to be faster and for some situations, the obtained sensing matrix led to a more accurate signal recovery than the approach proposed in [6] . However, this numerical procedure used in (13) lost the original intension of making the Gram matrix as close to the identity matrix as possible due to several approximation procedures involved in.

Zelnik-Manor considered the following optimal sensing matrix design problem in [8] formulated as

$$\min_{\Phi \in \Re^{M \times N}} ||I_L - \Psi^{\mathcal{T}} \Phi^{\mathcal{T}} \Phi \Psi||_F^2 \tag{14}$$

where I_L denotes the identity matrix of dimension L. Compared with (13), (14) has a much clearer physical meaning. Noting that

$$||I_L - \Psi^{\mathcal{T}} \Phi^{\mathcal{T}} \Phi \Psi||_F^2 = \sum_{i \neq j} |g_{ij}|^2 + \sum_{k=1}^{L} |1 - g_{kk}|^2 \tag{15}$$

In the one hand, as we have seen

$$\mu_t(A) = \frac{\sum_{\forall \ (i,j) \in \mathcal{S}_t} |\bar{g}_{ij}|}{N_t}$$

Thus $\sum_{i \neq j} |g_{ij}|^2$ is related to the the averaged coherence, which is just the term we want to optimize. In the other hand, noting the second term $\sum_{k=1}^{L} |1 - g_{kk}|^2$ means the distance of equivalent *atoms* (or *frames*) $\{A_k\}$ to be one, where A_k is the kth column vector of the equivalent dictionary A. Obviously, this term included in (15) observes the purpose of normalizing the equivalent dictionary.

3.2 Problem formation and the proposed method

An Equiangular Tight Frame (ETF) of size $N \times M$ with $N \leq M$ is a matrix with normalized columns such that its Gram matrix $G = A^{\mathcal{T}} A$ satisfies

$$\forall k \neq j, \ |G_{i,j}| = \sqrt{\frac{L - M}{M(L-1)}} \tag{16}$$

As we have stated in (11), this is the smallest possible mutual coherence possible.

An ETF has a very nice averaged mutual coherence behavior and has been used in optimal dictionary design [10]. However, it is difficult to make the equivalent dictionary $A = \Phi\Psi$ an ETF with Φ only as the degrees of freedom, compared with those in a totally free A, are much reduced. Therefore, we extend the searching space to a more convex set Λ_ϵ

$$\Lambda_\epsilon \triangleq \{G_{ea} \in \Re^{L \times L} : \quad G_{ea} = G_{ea}^{\mathcal{T}}, G_{ea}(k,k) = 1, \forall k \\ \max_{i \neq j} |G_{ea}(i,j)| \leq \epsilon\} \tag{17}$$

in which $\epsilon > 0$ is a constant to control the searching space. When $\epsilon = \underline{\mu}$, the ideal ETF Grams of dimension L are confined in Λ_ϵ.

Based on the discussions above, we formulate the optimal sensing matrix design problems as below:

$$\min_{\Phi \in \Re^{M \times N}, G_{ea} \in \Lambda_{\underline{\mu}}} ||G_{ea} - \Psi^{\mathcal{T}} \Phi^{\mathcal{T}} \Phi \Psi||_F^2 \tag{18}$$

where the dictionary Ψ is assumed to be given and G_{ea} is the targeted Gram which belongs to the space $\Lambda_{\underline{\mu}}$.

Such a problem can be solved practically using *alternative minimization* based numerical procedure, which is outlined below:

Objective: To optimize Φ

Input: Parameters to be set:

- $\Psi \in \Re^{N \times L}$: the dictionary
- $\Phi \in \Re^{M \times N}$: the projection
- $iter$: number of iteration

Initialization: With Ψ given, an initial Φ_0, say randomly generated, and set $\Phi = \Phi_0$.

Loop: Set $k = 1$ and repeat $iter$ times.

- **Step I**: While $1 \leq k \leq iter$, compute $G = (\Phi\Psi)^{\mathcal{T}}(\Phi\Psi)$, then normalize it.
- **Step II**: Solve

$$G_{ea} = arg \min_{G_{ea} \in \Lambda_{\underline{\mu}}} ||G_{ea} - G||_F^2 \tag{19}$$

the solution of this problem is given in [11] as follows:

$$G_{ea}(i,j) = \begin{cases} G(i,j), & |G(i,j)| \leq \underline{\mu} \\ 1, & i = j \\ sign(G(i,j))\underline{\mu}, & otherwise \end{cases} \tag{20}$$

- **Step III**: With G_{ea} obtained above, find the optimal sensing matrix Φ:

$$\tilde{\Phi} = arg \min_{\Phi} ||G_{ea} - \Psi^{\mathcal{T}}\Phi^{\mathcal{T}}\Phi\Psi||_F \quad (21)$$

  and if $||G_{ea} - \Psi^{\mathcal{T}}\tilde{\Phi}^{\mathcal{T}}\tilde{\Phi}\Psi||_F < ||G_{ea} - \Psi^{\mathcal{T}}\Phi^{\mathcal{T}}\Phi\Psi||_F$ then $\Phi = \tilde{\Phi}$ and go to *Step I* with $k \rightarrow k+1$
- **Step IV**: End while

We propose a novel algorithm based on SVD in the following to solve the corresponding minimization problem (21) in Step II of the above algorithm.

Denote $V \triangleq (\Phi\Psi)^{\mathcal{T}} \triangleq \Psi^{\mathcal{T}}Y \in \Re^{L\times M}$. The Gram matrix can then be rewritten into

$$G \triangleq VV^{\mathcal{T}} = \sum_{k=1}^{M} v_k v_k^{\mathcal{T}} \quad (22)$$

Let

$$\Psi = U \left[\begin{array}{cc} \sum_{\Psi} & \mathbf{0} \end{array} \right] Q^{\mathcal{T}} \quad (23)$$

be an SVD of Ψ. It is easy to see that

$$\begin{aligned} v_k &= \Psi^{\mathcal{T}} y_k \triangleq \sum_{m=1}^{N} w_{mk} q_m \\ &= Q(:, 1:N) w_k, \ \forall\, k = 1, 2, \cdots, M \quad (24) \end{aligned}$$

Define

$$\Delta(\Phi) \triangleq G_{ea} - G \quad (25)$$

where G_{ea} is obtained through Step II of the above algorithm. For a given Φ, one wishes to update it with $\tilde{\Phi}$ such that

$$||\Delta(\tilde{\Phi})||_F \leq ||\Delta(\Phi)||_F$$

and this can be done using the procedure given below. The basic idea is to update the column vectors of

$$V = \left[\begin{array}{ccccc} v_1 & \cdots & v_k & \cdots & v_M \end{array} \right]$$

one by one, leading to a new matrix

$$\tilde{V} = \left[\begin{array}{ccccc} \tilde{v}_1 & \cdots & \tilde{v}_k & \cdots & \tilde{v}_M \end{array} \right]$$

Assume that the updating has been done for $k = 1, \cdots, m-1$, that is

$$\tilde{v}_j = v_j, \ \forall\, j \geq m$$

To simplify the explanation, let us define a sequence of matrices

$$\tilde{V}_{k-1} \triangleq \left[\begin{array}{cccccc} \tilde{v}_1 & \cdots & \tilde{v}_{k-1} & v_k & v_{k+1} & \cdots & v_M \end{array} \right]$$

and hence

$$\Delta_k \triangleq G_{ea} - \tilde{V}_k \tilde{V}_k^{\mathcal{T}}$$

It is assumed that

$$||\Delta(\Phi)||_F \geq ||\Delta_{k-1}||_F \geq ||\Delta_k||_F, \ \forall\, k \leq m-1 \quad (26)$$

Let $\{\lambda_{l,m}\}$ be the eigenvalue set of $\Delta_{m-1} + v_m v_m^{\mathcal{T}}$, which is actually equal to $G_{ea} - \tilde{V}_{m-1}\tilde{V}_{m-1}^{\mathcal{T}} + v_m v_m^{\mathcal{T}}$. Noting the symmetry of such a matrix, one has

$$\Delta_{m-1} + v_m v_m^{\mathcal{T}} = E_m \Lambda_m E_m^{\mathcal{T}}$$

where $\Lambda_m = diag(\lambda_{m1}, \ \cdots, \ \lambda_{mL})$ and E_m is the matrix formed with a set of (orth-normal) eigenvectors of $\Delta_{m-1} + v_m v_m^{\mathcal{T}}$.

Note

$$\Delta_m \triangleq G_{ea} - \tilde{V}_m \tilde{V}_m^{\mathcal{T}} = E_m \Lambda_m E_m^{\mathcal{T}} - \tilde{v}_m \tilde{v}_m^{\mathcal{T}}$$

with $\tilde{v}_m$ to be determined.

Obviously, if $\lambda_{mk} > 0$ is the maximal eigenvalue, one possible choice for $\tilde{v}_m$ is to take

$$\tilde{v}_m = \sqrt{\lambda_{mk}} E_m(:, k) \quad (27)$$

as for such a choice, $||\Delta_m||_F \leq ||\Delta_{m-1}||_F$ holds. Since all the vectors $\tilde{v}_k$ in $\tilde{V}$ should belong to the space spanned with $q_1, q_2, \cdots, q_N$, one has to use the best projection of $\tilde{v}_m$ on this space, which leads to

$$\tilde{v}_m = \sum_{k=1}^{N} w_{mk} q_k \quad (28)$$

where $w_{mk} = q_k^{\mathcal{T}} \tilde{v}_m$, $k = 1, \cdots, N$ as long as with the just updated $\tilde{V}_m$ given by (28), and then produce for $m+1$ until all the vectors are updated. With such updated $\tilde{V}$, then

$$\tilde{\Phi} = \tilde{V}^{\mathcal{T}} \Psi^{\mathcal{T}} (\Psi\Psi^{\mathcal{T}})^{-1} \quad (29)$$

4 Analysis of the Proposed Algorithm

To illustrate the behavior of the proposed algorithm and compare it with other sensing matrix designs including Gaussian matrix, Elad's algorithm [1] [6], DCS's algorithm [7] and ZRE's algorithm [8], we provide a demonstration in Fig. 1 and Fig. 2. We generate a random dictionary $N \times L$, and then compute the best projection matrix Φ with the four method. Both of the proposed algorithm and Elad's algorithm are initialized with a $M \times N$ random matrix Φ_0 and ran 1000 iterations.

Fig. 1 presents the distribution of the absolute value of the off-diagonal elements of the corresponding normalized Gram matrix to the five sensing matrix. As shown, there is a remarkable shift towards the origin of the histogram after optimized by our algorithm, with a shorter right tail which represents the higher values.

Fig. 2 illustrates the convergence of the averaged mutual coherence $\mu_t(A), t = 0.2$ for Elad's algorithm and proposed algorithm. As can be seen, our algorithm yields a smaller μ_t than that by Elad's algorithm at almost every iteration.

[1] This algorithm has two parameters γ and t. In our simulations, we set $t = 20\%, \gamma = 0.95$, unless there is additional instruction.

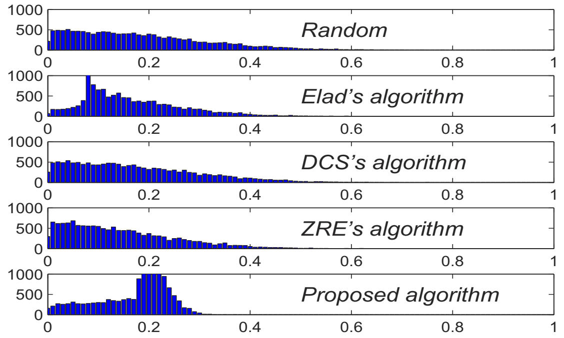

Fig. 1: Histogram of the absolute off-diagonal values of different Gram matrix ($N = 25$, $M = 80$ and $L = 120$).

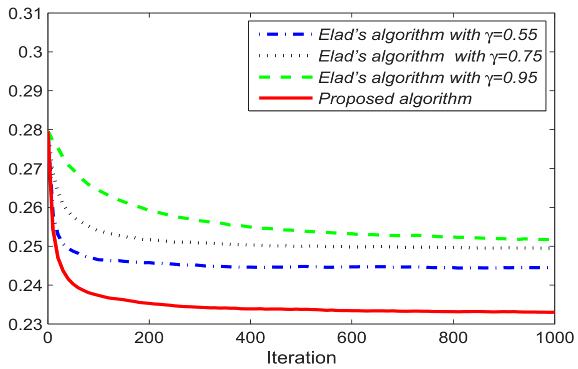

Fig. 2: Value of $\mu_t(A), t = 0.2$ as a function of the iteration for Elad's algorithm and proposed algorithm ($N = 25, M = 80$ and $L = 120$).

5 Computer Simulation Results

We now present some numerical experiments to evaluate the performance of the optimized projections via the signal recovery accuracy. We choose a dictionary $\Psi \in \Re^{N \times L}$ and synthesize 1000 test signals $\{y_j\}_{j=1}^{1000}$ by randomly generating K-sparse $L \times 1$ vectors $\{s_j\}_{j=1}^{1000}$, and computing $y_j = \Psi s_j$. Then we apply random sensing projection and designed projections to get measurements with $z = \Phi y_j$. OMP method is used to recover the sparse vectors $\hat{s}_j$ from the measurements by approximating the solution of

$$\hat{s}_j = arg \ \min \ ||s||_0 \quad s.t. \quad z_j = \Phi \Psi s$$

Then we reconstruct the signal $y_j = \Psi s_j$ and test the recovery error of the relevant CS system via

$$e_r = \frac{1}{1000} \sum_{k=1}^{1000} ||y_k - \hat{y}_k||_2^2 / ||y_k||_2^2, \ \ \hat{y}_k \ = \ \Psi \hat{s}$$

In the first experiment, the size of the CS system is $M = 25, \ N = 80 \ and \ L = 120$. The sparsity K varies in the range $[1, 7]$. The results are depicted in Fig. 3. As seen, our proposed algorithm can yield a better recovery accuracy than others for all sparsity levels.

The second experiment is similar to the first one, this time fixing sparsity $K = 4$, we vary M from 16 to 40. The results are shown in Fig. 4. As excepted, the results improves as M increases for all projections. Once again it is evident that our proposed sensing matrix outperforms the other sensing matrices. We should point out that both of the settings of the two experiments are the same as [6].

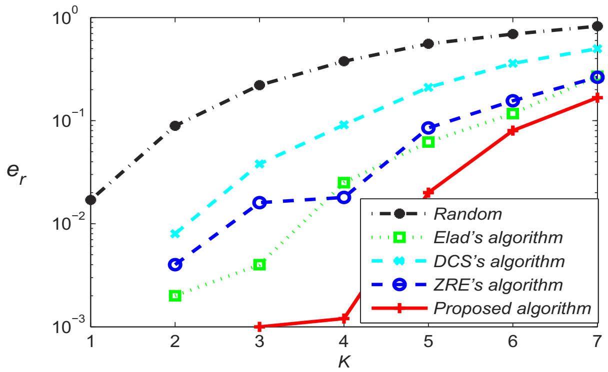

Fig. 3: Reconstruction error e_r as a function of the signal sparsity K for $M = 25, N = 80, L = 120$, with random projection and optimized projections. Note: a vanishing graph implies a zero error rate.

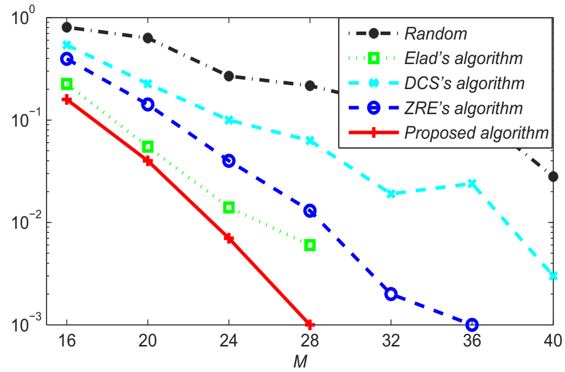

Fig. 4: Reconstruction error e_r as a function of the number of measurement M, $N = 80, L = 120, K = 4$, with random projection and optimized projections. Note: a vanishing graph implies a zero error rate.

6 Conclusions

In this paper, we investigate the problem of projection matrix design for sensing signals which are sparse in overcomplete dictionaries and a novel algorithm based on SVD for optimal projection matrix searching is proposed in this paper. Experiments are given to show that the sensing matrix obtained using our proposed algorithm outperforms others in signal reconstruction accuracy.

Acknowledgment

This work was supported by the NSFC-Grants 61273195, CPSF-Grant 2012M511386 and ZSFC-Grant Y13F010050.

References

[1] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489-509, Feb., 2006.

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

[3] E. J. Candès and T. Tao, "Near optimal signal recovery fom random projection: universal encoding strategies?," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406 - 5425, Dec., 2006.

[4] M. F. Duarte and Y. C. Eldar, "Structured compressed sensing: from theory to applications," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4053 - 4085, Sept., 2011.

[5] D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthonormal) dictionaries via l_1 minimization," *Proc. Nat. Acad. Sci.*, vol. 100, no. 5, pp. 2197 - 2202, Mar. 2003.

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

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

[8] L. Zelnik-Manor, K. Rosenblum, and Y. C. Eldar, "Sensing matrix optimization for block-sparse decoding," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4300-4312, 2011.

[9] T. Strohmer and R. W. Heath, "Grassmannian frames with applications to coding and communication," Appl. Comp. Harmonic Anal., vol. 14, no. 3, pp. 257-275, May 2003.

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

[11] M. Yaghoobi, L. Daudet, and M. E. Davies, "Parametric disctionary design for sparse coding," *IEEE Trans. Signal Process.*, vol. 57, no. 12, pp. 4800-4810, 2009.