ADAPTIVE SENSING MATRIX DESIGN FOR GREEDY ALGORITHMS IN MMV COMPRESSIVE SENSING

L. Zhang, L. Huang, B. Zhao*

College of Information Engineering
Shenzhen University
Shenzhen, China

B. Li

Nuance Communication Inc.
Montreal, Canada

ABSTRACT

Sensing matrix can be designed with low coherence with the measurement matrix to improve the sparse signal recovery performance of greedy algorithms. However, most of the sensing matrix design algorithms are computationally expensive due to large number of iterations. This paper proposes an iteration-free sensing matrix design algorithm for multiple measurement vectors (MMV) compressive sensing. Specifically, sensing matrix is designed in the sense of the local cumulative cross-coherence (LCCC) of the sensing matrix with respect to the measurement matrix when the number of MMV is sufficient and the sparse signals are of full rank. Experiment results verify the effectiveness of the proposed algorithm in terms of improving the sparse signal recovery performance of greedy algorithms.

***Index Terms*—** Sparse signal recovery, multiple measurement vectors, sensing matrix design, simultaneous orthogonal matching pursuit

1. INTRODUCTION

With multiple measurement vectors (MMV) along time instance, the measurement equation of compressive sensing (CS) can be formulated as

$$\boldsymbol{y}_l = \boldsymbol{\Phi}\boldsymbol{x}_l + \boldsymbol{n}_l, \quad l = 1, 2, \cdots, L \tag{1}$$

where $\boldsymbol{\Phi} \in \mathbb{R}^{M\times N}$ $(M < N)$ is the measurement matrix, and $\boldsymbol{x}_l \in \mathbb{R}^{N\times 1}$, $\boldsymbol{y}_l \in \mathbb{R}^{M\times 1}$ and $\boldsymbol{n}_l \in \mathbb{R}^{M\times 1}$ are the vectors of signal, measurement and noise, respectively. Note that (1) can be compactly represented in the form of matrix

$$\boldsymbol{Y} = \boldsymbol{\Phi}\boldsymbol{X} + \boldsymbol{N} \tag{2}$$

where $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_L]$ is the jointly sparse signal, $\boldsymbol{Y} = [\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_L]$ is the measurement, $\boldsymbol{N} = [\boldsymbol{n}_1, \boldsymbol{n}_2, \cdots, \boldsymbol{n}_L]$ is the measurement noise. In the case of MMV-CS, the vectors $\{\boldsymbol{x}_l\}_{l=1}^{L}$ share the same sparse pattern which means that the matrix $\boldsymbol{X}$ only has a few of rows with non-zero entries.

Usually, the recovery of the sparse signal from its multiple linear measurements can be realized by solving the following optimization problem

$$\min_{\boldsymbol{X}\in\mathbb{R}^{N\times L}} \|\boldsymbol{Y} - \boldsymbol{\Phi}\boldsymbol{X}\|_F^2 + \lambda\mathcal{R}(\boldsymbol{X}) \tag{3}$$

where $\|\cdot\|_F$ indicates the Frobenius norm, $\mathcal{R}(\cdot)$ is an operator that gives the number of non-zero rows of the input signal $\boldsymbol{X}$. The first term in (3) is the data fidelity term and second one forces the recovered signal to be sparse. The $\lambda > 0$ is a regularization parameter which is the trade-off between data fitting and the sparsity of signal.

Various methodologies have been proposed to address the sparse recovery problem, such as simultaneous orthogonal matching pursuit (SOMP) [1–3], Reduce MMV and boost (ReMbo) [4], mixed norm approach [5], rank aware order recursive matching pursuit (RA-ORMP) [6]. Among these methods, the SOMP algorithm has received attention due to its simplicity, low computational complexity and excellent recovery performance. The essential of greedy algorithms is to estimate the support which contains the indexes of the non-zero rows in the sparse signal. It has been proved that the error of the estimated sparse signal achieves the Cramer-Rao bound as long as the support is recovered correctly [1].

The SOMP algorithm recovers the support according to the absolute inner product of the residual signal and each column of $\boldsymbol{\Phi}$ [1, 3]. The methods of designing measurement matrix with low coherence between different columns are developed in [1, 7–9]. Subsequently, the sensing matrix $\boldsymbol{\Psi} \in \mathbb{R}^{M\times N}$ is proposed for support recovery in greedy algorithms and hard thresholding algorithm [10], where $\boldsymbol{\Psi}$, instead of $\boldsymbol{\Phi}$, is exploited for sparse recovery. It is proved that low coherence between $\boldsymbol{\Psi}$ and $\boldsymbol{\Phi}$ can improve the support recovery accuracy. Based on this conclusion, both $\boldsymbol{\Psi}$ and $\boldsymbol{\Phi}$ are constructed in [10, 11]. However, the performance in terms of sparsity level is not generally satisfactory, and there is a great gap from the upper bound provided in [12].

*This work was supported by the Innovative and Entrepreneurial Talent Scheme in Jiangsu Province (2018), in part by the National Natural Science Foundation of China under Grants U1713217 and 61801297.

In this paper, an iteration-free sensing matrix design algorithm for MMV-CS is proposed to improve the sparsity level. In comparison with other sensing matrix design criteria, the coherency between $\boldsymbol{\Psi}$ and $\boldsymbol{\Phi}$, as well as that between $\boldsymbol{\Psi}$ and $\boldsymbol{Y}$ are considered. This is because the measurement matrix $\boldsymbol{Y}$ is with the support information, since they are linear combinations of the measurement columns indexed by the support. Therefore, the proposed method is able to improve the accuracy of the support recovery, which in turn enhances the performance of sparse signal recovery. The proposed method can be described by a quadratic optimization problem, and the closed form solution to this problem is provided. Furthermore, numerical simulations have been conducted to illustrate the performance of the method.

2. COHERENCE MEASUREMENT OF MEASUREMENT MATRIX AND SENSING MATRIX

A generalized parameter to address the coherence of measurement matrices is the cumulative coherence [3]. The kth cumulative coherence is defined as

$$\mu_c(k, \boldsymbol{\Phi}) = \max_{|\Gamma|=k} \sum_{i,j=1,2,\cdots,N, i\notin\Gamma, j\in\Gamma} \left|\boldsymbol{\Phi}_{.i}^T \boldsymbol{\Phi}_{.j}\right|. \tag{4}$$

It is proved in [1] that $\mu_c(k, \boldsymbol{\Phi}) + \mu_c(k-1, \boldsymbol{\Phi}) < 1$ can guarantee the success of both OMP and basis pursuit (BP) algorithms. Due to the significance of the correlation property, many works have been done to design the measurement matrix with small cumulative coherence. In [13], a measurement matrix with low coherence is constructed by shrinkage method. In [14], a gradient based optimization method is proposed to determine the measurement matrix. Other projection matrix or measurement matrix design methods have been suggested in [15,16]. However, the measurement matrix is redundant which makes it impossible to design a measurement matrix with zero coherence coefficient or with all columns orthogonal.

In [10], the concept of sensing matrix $\boldsymbol{\Psi} \in \mathbb{R}^{M\times N}$ is proposed for support recovery in OMP and hard thresholding algorithms. The purpose of designing sensing matrix in CS is to reduce the coherence of sensing and measurement matrices, which can improve the recovery performance. The process of sparse recovery by using sensing matrix can be expressed as

$$\hat{\boldsymbol{X}} = \mathcal{R}(\boldsymbol{Y}, \boldsymbol{\Psi}, \boldsymbol{\Phi}, \cdots). \tag{5}$$

A parameter termed as the cumulative cross-coherence (CCC) is proposed to measure the coherence between $\boldsymbol{\Phi}$ and $\boldsymbol{\Psi}$ [10], defined as

$$\tilde{\mu}_c(k, \boldsymbol{\Psi}, \boldsymbol{\Phi}) = \max_i \max_{|J|=k, i\notin J} \sum_{j\in J} \left|\boldsymbol{\Psi}_{.i}^T \boldsymbol{\Phi}_{.j}\right|. \tag{6}$$

In [10], a sufficient condition for the support recovery with the OMP algorithm is proposed. It has been proved that the smaller CCC between $\boldsymbol{\Psi}$ and $\boldsymbol{\Phi}$ results in higher accuracy of support recovery. Based on this property, the sensing matrices are constructed in [10, 11].

By exploiting the received data to determine the support, a re-weighted algorithm for data-dependent sensing matrix design is proposed in [17], in which the local cumulative cross-coherence (LCCC) is defined

$$\hat{\mu}_c(k, \boldsymbol{\Psi}, \boldsymbol{\Phi}_\Gamma) = \max_{|J|=k, J\subseteq\Gamma} \max_{i\notin J} \sum_{j\in J} \left|\boldsymbol{\Psi}_{.i}^T \boldsymbol{\Phi}_{.j}\right| \tag{7}$$

where Γ is the support of sparse signal.

It can be seen from (7) that $\hat{\mu}_c(k, \boldsymbol{\Psi}, \boldsymbol{\Phi}_\Gamma)$ represents the worst case coherence between the columns of the sensing matrix and measurement columns indexed by the support Γ. In other words, $\hat{\mu}_c(k, \boldsymbol{\Psi}, \boldsymbol{\Phi}_\Gamma)$ describes the local coherence between the sensing matrix and the measurement matrix, while $\tilde{\mu}_c(k, \boldsymbol{\Psi}, \boldsymbol{\Phi})$ describes the global coherence of the sensing matrix with the measurement matrix.

3. ADAPTIVE SENSING MATRIX DESIGN

3.1. Sensing Matrix Design with Sufficient MMV

For sufficient multiple measurement vectors, i.e., $L \geq M$, the recovered signal can be expressed as

$$\begin{aligned}
\hat{\boldsymbol{X}}_{i.} =& \boldsymbol{\Psi}_{.i}^T \boldsymbol{Y} = \boldsymbol{\Psi}_{.i}^T (\boldsymbol{\Phi}\boldsymbol{X} + \boldsymbol{N}) = \boldsymbol{\Psi}_{.i}^T \boldsymbol{\Phi}\boldsymbol{X} + \boldsymbol{\Psi}_{.i}^T \boldsymbol{N} \\
=& \boldsymbol{\Psi}_{.i}^T \Bigg[\sum_{j=1}^N \boldsymbol{\Phi}_{.j}\boldsymbol{X}(j,1), \sum_{j=1}^N \boldsymbol{\Phi}_{.j}\boldsymbol{X}(j,2), \cdots, \\
& \sum_{j=1}^N \boldsymbol{\Phi}_{.j}\boldsymbol{X}(j,L)\Bigg] + \boldsymbol{\Psi}_{.i}^T \boldsymbol{N} \\
=& \Bigg[\boldsymbol{\Psi}_{.i}^T \boldsymbol{\Phi}_{.i}\boldsymbol{X}(i,1) + \boldsymbol{\Psi}_{.i}^T \sum_{i\neq j, j=1}^N \boldsymbol{\Phi}_{.j}\boldsymbol{X}(j,1), \cdots, \\
& \boldsymbol{\Psi}_{.i}^T \boldsymbol{\Phi}_{.i}\boldsymbol{X}(i,L) + \boldsymbol{\Psi}_{.i}^T \sum_{i\neq j, j=1}^N \boldsymbol{\Phi}_{.j}\boldsymbol{X}(j,L)\Bigg] \\
& + \boldsymbol{\Psi}_{.i}^T \boldsymbol{N}.
\end{aligned} \tag{8}$$

It follows from (8) that in order to exactly recover the jointly sparse signal $\boldsymbol{X}$, the terms $\boldsymbol{\Psi}_{.i}^T \boldsymbol{\Phi}_{.i}\boldsymbol{X}(i,1), \cdots, \boldsymbol{\Psi}_{.i}^T \boldsymbol{\Phi}_{.i}\boldsymbol{X}(i,L)$ for $i = 1, 2, \cdots, N$ should be kept distortionless for $\boldsymbol{\Psi}_{.i}^T \boldsymbol{\Phi}_{.i} = 1$, while other terms should be minimized. Given the measurements $\boldsymbol{Y}$, the sensing matrix can be designed as follows

$$\begin{aligned}
&\min_{\boldsymbol{\Psi}_{.i}\in\mathbb{R}^{M\times 1}} \left\|\boldsymbol{\Psi}_{.i}^T \boldsymbol{Y}\right\|_2^2 \\
&\text{s.t.} \quad \boldsymbol{\Psi}_{.i}^T \boldsymbol{\Phi}_{.i} = 1.
\end{aligned} \tag{9}$$

The optimization problem in (9) is a quadratic programming problem with a linear constraint. Its closed form solution is

$$\boldsymbol{\Psi}_{.i} = \frac{\boldsymbol{R}^{-1}\boldsymbol{\Phi}_{.i}}{\boldsymbol{\Phi}_{.i}^T \boldsymbol{R}^{-1}\boldsymbol{\Phi}_{.i}} \tag{10}$$

where $\boldsymbol{R} = \frac{1}{L}\boldsymbol{Y}\boldsymbol{Y}^T$. To ensure that $\boldsymbol{Y}$ is with full row rank, the sample size needs to be no less than measurement size, i.e., $L \geq M$.

Proposition 1 *For the SOMP algorithm, the sensing matrix $\boldsymbol{\Psi}$ designed by (10) provides a decreased LCCC and the bound is*

$$0 \leq \hat{\mu}_c(K, \boldsymbol{\Psi}, \boldsymbol{\Phi}_\Gamma) \leq \tilde{\mu}_c(K, \boldsymbol{\Psi}, \boldsymbol{\Phi}). \tag{11}$$

Proof: For brevity, only the proof of Proposition 1 in the noisy case is provided, since the proof in the noiseless case can be derived with the similar procedure. For a given vector, the minimization of (9) is equivalent to the form of the ℓ_2-norm minimization, therefore, the objective function can be rewritten as

$$\begin{aligned}
\left\|\boldsymbol{\Psi}_{\cdot i}^T \boldsymbol{Y}\right\|_2 &= \left\|\boldsymbol{\Psi}_{\cdot i}^T \left(\boldsymbol{\Phi}_\Gamma \boldsymbol{X}^\Gamma + \boldsymbol{N}\right)\right\|_2 \\
&\leq \left\|\boldsymbol{\Psi}_{\cdot i}^T \boldsymbol{\Phi}_\Gamma \boldsymbol{X}^\Gamma\right\|_2 + \left\|\boldsymbol{\Psi}_{\cdot i}^T \boldsymbol{N}\right\|_2 \\
&\leq \left\|\boldsymbol{\Psi}_{\cdot i}^T \boldsymbol{\Phi}_\Gamma\right\|_2 \left\|\boldsymbol{X}^\Gamma\right\|_F + \left\|\boldsymbol{\Psi}_{\cdot i}^T\right\|_2 \|\boldsymbol{N}\|_F \\
&= \sigma_s \left\|\boldsymbol{\Psi}_{\cdot i}^T \boldsymbol{\Phi}_\Gamma\right\|_2 + \sigma_n \\
&\leq \sigma_s \left\|\boldsymbol{\Psi}_{\cdot i}^T \boldsymbol{\Phi}_\Gamma\right\|_1 + \sigma_n \\
&= \sigma_s \left(\sum_{j\in\Gamma} \left|\boldsymbol{\Psi}_{\cdot i}^T \boldsymbol{\Phi}_{\cdot j}\right|\right) + \sigma_n
\end{aligned} \tag{12}$$

where $i \in \Gamma^c$, $\|\boldsymbol{X}^\Gamma\|_F = \sigma_s$ and $\|\boldsymbol{N}\|_F = \sigma_n$. In the second inequation of (12), $\|\boldsymbol{\Psi}_{\cdot i}\|_2 = 1$ is applied. Therefore, the minimization of (9) is equivalent to the minimization of $f(i) = \sum_{j\in\Gamma} |\boldsymbol{\Psi}_{\cdot i}^T \boldsymbol{\Phi}_{\cdot j}|$. Refer to the definition of $\hat{\mu}_c(K, \boldsymbol{\Psi}, \boldsymbol{\Phi}_\Gamma)$ in (7), it can be rewritten as follows

$$\hat{\mu}_c(K, \boldsymbol{\Psi}, \boldsymbol{\Phi}_\Gamma) = \max_{i\in\Gamma^c} \left(\sum_{j\in\Gamma} \left|\boldsymbol{\Psi}_{\cdot i}^T \boldsymbol{\Phi}_{\cdot j}\right|\right) = \max_{i\in\Gamma^c} f(i). \tag{13}$$

As $f(i)$ for any $i \in \Gamma^c$ is minimized by the optimization problem (9), it suggests that $\hat{\mu}_c(K, \boldsymbol{\Psi}, \boldsymbol{\Phi}_\Gamma)$ gets decreased.

Next, we consider the lower bound of the LCCC. According to (12), one has

$$\sigma_s \left\|\boldsymbol{\Psi}_{\cdot i}^T \boldsymbol{\Phi}_\Gamma\right\|_2 + \sigma_n \leq \sigma_s \left(\sum_{j\in\Gamma} \left|\boldsymbol{\Psi}_{\cdot i}^T \boldsymbol{\Phi}_{\cdot j}\right|\right) + \sigma_n. \tag{14}$$

Based on the concept of subspace, it is easy to understand that the vector $\boldsymbol{\Psi}_{\cdot i}$ lies in the null space of $\boldsymbol{\Phi}_\Gamma$ such that $\sum_{j\in\Gamma} |\boldsymbol{\Psi}_{\cdot i}^T \boldsymbol{\Phi}_{\cdot j}|$ gets the lower bound. Therefore, the inner product of $\boldsymbol{\Psi}_{\cdot i}$ and $\boldsymbol{\Phi}_{\cdot j}$ for $j \in \Gamma$ is 0, which means the lower bound of LCCC is decreased to 0. Furthermore, the upper bound of $\hat{\mu}_c(K, \boldsymbol{\Phi}, \boldsymbol{\Psi}_\Gamma)$ can be derived as follows

$$\begin{aligned}
\hat{\mu}_c(K, \boldsymbol{\Phi}, \boldsymbol{\Psi}_\Gamma) &= \max_{i\in\Gamma^c} \sum_{j\in\Gamma} \left|\boldsymbol{\Psi}_{\cdot i}^T \boldsymbol{\Phi}_{\cdot j}\right| \\
&= \max_{i\in\Gamma^c} \left\|\boldsymbol{\Psi}_{\cdot i}^T \boldsymbol{\Phi}_\Gamma\right\|_1 \\
&= \left\|\boldsymbol{\Psi}_{\Gamma^c}^T \boldsymbol{\Phi}_\Gamma\right\|_{\infty,\infty} \\
&\leq \max_{i\in J^c} \sum_{j\in J, |J|=K} \left|\boldsymbol{\Psi}_{\cdot i}^T \boldsymbol{\Phi}_{\cdot j}\right| \\
&= \tilde{\mu}_c(K, \boldsymbol{\Phi}, \boldsymbol{\Psi})
\end{aligned} \tag{15}$$

where $\|\cdot\|_{\infty,\infty}$ is the maximum ℓ_1-norm of the row of a matrix, J is an arbitrary subset of $\Omega = \{1, 2, \cdots, N\}$ with $|J| = K$. From (15) it can be seen that the upper bound of $\hat{\mu}_c(K, \boldsymbol{\Phi}, \boldsymbol{\Psi}_\Gamma)$ is $\tilde{\mu}_c(K, \boldsymbol{\Phi}, \boldsymbol{\Psi})$. Therefore, the bound of LCCC is provided as (11). ∎

Remarks: 1) The subspace method is presented in the proof of Proposition 1. The $\boldsymbol{\Psi}_i$ lies in the null space, which is orthogonal to the signal space that consists of the columns of measurement matrix $\boldsymbol{\Phi}$ indexed by the support set Γ. According to the definition of the LCCC, it is decreased by the obtained sensing matrix. 2) In [17], it has been proved that the smaller LCCC is, the better performance of sparse recovery can be achieved. Therefore, Proposition 1 indicates that the designed sensing matrix is able to improve the performance of sparse recovery.

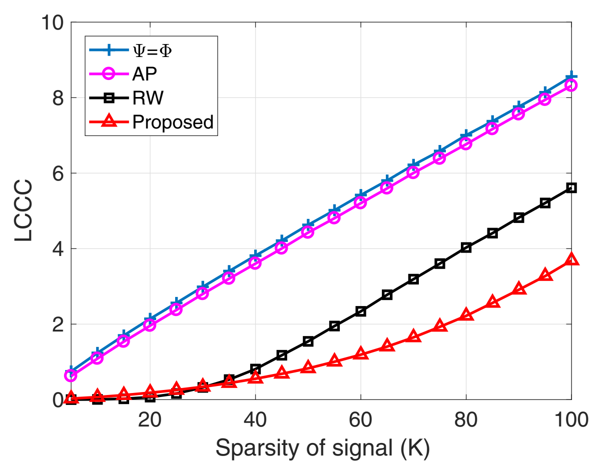

Fig. 1: LCCC versus sparsity of signal with SNR $=$ 20dB and $L = 500$.

4. NUMERICAL SIMULATIONS

In order to validate the effectiveness of the proposed method, numerical simulations have been conducted. The simulation

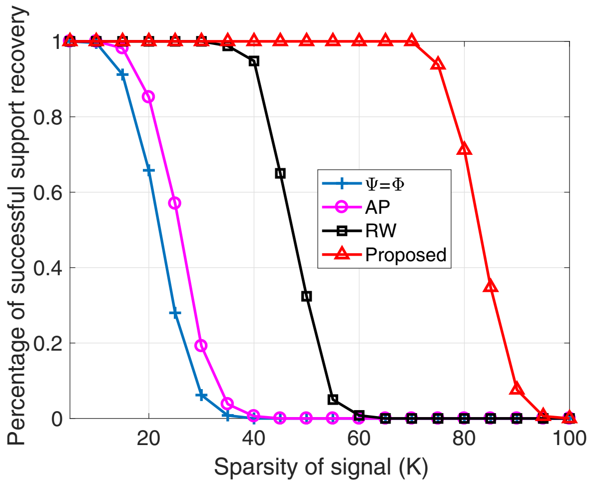

Fig. 2: Percentage of successful recovery versus sparsity of signal with $\mathrm{SNR} = 20\mathrm{dB}$ and $L = 500$.

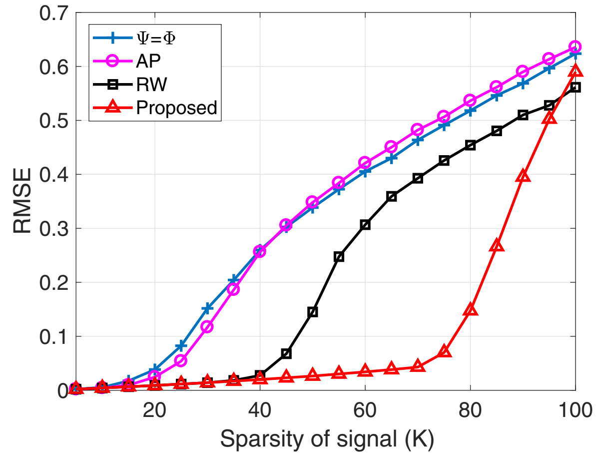

Fig. 3: RMSE versus sparsity of signal with $\mathrm{SNR} = 20\mathrm{dB}$ and $L = 500$.

settings are provided as follows. The sparse signal is generated by Gaussian distribution with mean one and variance 0.1. The sparsity of the signal, i.e., K, varies from 5 to 100. The sizes of the sensing matrix and measurement matrix are both 128×256. The entries of the measurement matrix are drawn from Gaussian distribution with zero mean and unit variance. In order to evaluate the performances of the proposed approach, 500 independent trials are carried out at each specific case. The percentage of successful support recovery and the root mean square error (RMSE) of recovered signal are both calculated. For the purpose of comparison, the Alternating Projection (AP) algorithm [10] and Re-weighted (RW) algorithm [17] for sensing matrix design are performed. The conventional approach $\boldsymbol{\Psi} = \boldsymbol{\Phi}$ is also conducted.

When $\mathrm{SNR} = 20\mathrm{dB}$ and $L = 500$, the LCCC versus sparsity of signal is shown in Fig.1. As $K > 30$, the proposed method exhibits excellent performance in terms of minimizing LCCC, which ensures the support recovery for the SOMP algorithm.

Under the same simulation conditions, the performance of support recovery is evaluated and the corresponding result is depicted in Fig.2. It can be seen that the proposed method can correctly recovery the support with high probability particularly when K reaches 70.

The RMSEs for different sparsities of the signals are evaluated, and the simulation results are shown in Fig.3. It can be seen that all the three methods achieve excellent RMSE when the sparsity of the signal is less than 20. However, the AP and RW algorithms cannot work successfully when K increases to 40. This is because the SOMP algorithm cannot select the correct indices of the measurement matrix due to the coherence among the columns of the measurement matrix. The proposed method exhibits superior performance when $K \leq 70$ since it is able to provide smaller LCCC between $\boldsymbol{\Psi}$ and $\boldsymbol{\Phi}$ than the other two schemes.

5. CONCLUSIONS

In this paper, an iteration-free sensing matrix design algorithm for MMV-CS is proposed. In order to improve the performance of sparse signal recovery, the coherency of $\boldsymbol{\Psi}$ and $\boldsymbol{\Phi}$ as well as that of $\boldsymbol{\Psi}$ and $\boldsymbol{Y}$ are exploited. Comparing with the existing methods of sensing matrix design, the proposed algorithm is iteration-free and is able to further enhance the recovery performance. Simulation results confirm the superiority of the proposed approach.

6. REFERENCES

[1] J. A. Tropp, "Greed is good: algorithmic results for sparse approximation," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2231–2242, Oct. 2004.

[2] J. Chen and X. Huo, "Theoretical results on sparse representations of multiple-measurement vectors," *IEEE Trans. Signal Process.*, vol. 54, no. 12, pp. 4634–4643, Dec. 2006.

[3] J. A. Tropp, A. C. Gilbert, and M. J. Strauss, "Algorithms for simultaneous sparse approximation. Part I: Greedy pursuit," *Signal Process.*, vol. 86, no. 3, pp. 572–588, Mar. 2006.

[4] M. Mishali and Y. C. Eldar, "Reduce and boost: Recovering arbitrary sets of jointly sparse vectors," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 4692–4702, Oct. 2008.

[5] J. A. Tropp, "Algorithms for simultaneous sparse approximation. Part II: Convex relaxation," *Signal Process.*, vol. 86, no. 3, pp. 589–602, Mar. 2006.

[6] M. E. Davies and Y. C. Eldar, "Rank awareness in joint sparse recovery," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 1135–1146, Feb. 2012.

[7] G. Davis, S. Mallat, and M. Avellaneda, "Adaptive greedy approximations," *Constr. Approx.*, vol. 13, no. 1, pp. 57–98, Mar. 1997.

[8] J. J. Fuchs, "Extension of the pisarenko method to sparse linear arrays," *IEEE Trans. Signal Process.*, vol. 45, no. 10, pp. 2413–2421, Oct. 1997.

[9] D. Donoho and M. Elad, "Optimally sparse representation in general (non-orthogonal) dictionaries via ℓ_1 minimization," *P. Natl. Acad. Sci.*, vol. 100, no. 5, pp. 2197–2202, Mar. 2003.

[10] K. Schnass and P. Vandergheynst, "Dictionary preconditioning for greedy algorithms," *IEEE Trans. Signal Process.*, vol. 56, no. 5, pp. 1994–2002, May. 2008.

[11] B. Li, Y. Shen, and J. Li, "Dictionaries construction using alternating projection method in compressive sensing," *IEEE Signal Process. Lett.*, vol. 18, no. 11, pp. 663–666, Nov. 2011.

[12] J. M. Kim, O. K. Lee, and J. C. Ye, "Compressive MUSIC: Revisiting the link between compressive sensing and array signal processing," *IEEE Trans. Inf. Theory*, vol. 58, no. 1, pp. 278–301, Jan. 2012.

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

[14] V. Abolghasemi, S. Ferdowsi, and S. Sanei, "A gradient-based alternating minimization approach for optimization of the measurement matrix in compressive sensing," *Signal Process.*, vol. 92, no. 4, pp. 999–1009, Apr. 2012.

[15] G. Li, Z. Zhu, D. Yang, L. Chang, and H. Bai, "On projection matrix optimization for compressive sensing systems," *IEEE Trans. Signal Process.*, vol. 61, no. 11, pp. 2887–2898, June. 2013.

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

[17] A. Huang, G. Gui, Q. Wan, and A. Mehbodniya, "A reweighted algorithm for designing data dependent sensing dictionary," *Int. J. Phys. Sci.*, vol. 6, no. 3, pp. 386–390, Feb. 2011.