Adaptive Selective Compressive Sensing based Signal Acquisition Oriented toward Strong Signal Noise Scene

Fangqing Wen[1,2], Gong Zhang[1,2], De Ben[1]
[1] Key Laboratory of Radar Imaging and Microwave Photonics, Ministry of Education,
Nanjing, 210016, China
[e-mail: wfqzhc@163.com]
[2] College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics,
Nanjing, 210016, China
[e-mail: gzhang@nuaa.edu.cn]
*Corresponding author: Gong Zhang

Received October 27, 2014; revised January 29, 2015; accepted July 20, 2015;
published September 30, 2015

Abstract

This paper addresses the problem of signal acquisition with a sparse representation in a given orthonormal basis using fewer noisy measurements. The authors formulate the problem statement for randomly measuring with strong signal noise. The impact of white Gaussian signals noise on the recovery performance is analyzed to provide a theoretical basis for the reasonable design of the measurement matrix. With the idea that the measurement matrix can be adapted for noise suppression in the adaptive CS system, an adapted selective compressive sensing (ASCS) scheme is proposed whose measurement matrix can be updated according to the noise information fed back by the processing center. In terms of objective recovery quality, failure rate and mean-square error (MSE), a comparison is made with some nonadaptive methods and existing CS measurement approaches. Extensive numerical experiments show that the proposed scheme has better noise suppression performance and improves the support recovery of sparse signal. The proposed scheme should have a great potential and bright prospect of broadband signals such as biological signal measurement and radar signal detection.

Keywords: Signal acquirsion, Adaptive selective compressive sensing, signal noisy suppression, measurement matrix design

This work was supported by the Chinese National Natural Science Foundation under contract No. 61071163, 61201367, 61271327, 61471191, Funding for Outstanding Doctoral Dissertation in NUAA (BCXJ14-08), Funding of Jiangsu Innovation Program for Graduate Education Province (KYLX_0277), the Fundamental Research Funds for the Central Universities (NP2015504), and partly funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions (PADA). We express our thanks to Prof. Xing Zhang who checked our manuscript.

http://dx.doi.org/10.3837/tiis.2015.09.015 **ISSN : 1976-7277**

1. Introduction

As an alternative paradigm to the Shannon-Nyquist sampling theorem, compressive sensing (CS) enables sparse signal to be acquired by sub-Nyquist analog-to-digital converters (ADC), thus launch a revolution in signal collection, transmission and processing. The CS theory points out that if the signal is compressible or sparse in a transform domain, it can be recovered exactly with high probability from fewer measurements via l_1 -norm optimization [1]. Rather than the classical Shannon-Nyquist sampling theorem, which requires sampling signals at twice the bandwidth, CS promises to reduce the sampling bandwidth, which depends on the sparsity of the signal. Compared with the traditional radio frequency (RF) signal acquisition system, the sampling front-end of CS operates at a lower speed and then lowers the cost of the front-end sensor (such as size, weight and power consumption). The parts with intensive computation of the acquisition process are removed from the front-end sensor and are transferred to a central processing back-end. Due to the potential use in signal processing applications, CS has attracted vast interests in signal acquisition [2], radar detection [3], cognitive radio [4] and Massive antenna arrays [5].

CS has been considered from an adaptive perspective in [6]-[10]. The parameterized Bayesian model [6] is proposed in to dynamically determine whether a sufficient number of CS measurements have been performed. In [7], an empirical Bayesian model based multitask learning algorithm is developed to improve the performance of the inversion. An analogous work has been done in localization in wireless LANs [8]. These Bayesian methods have been demonstrated to achieve better recovery performance. However, they often require fewer noisy observations to recover sparse signals than nonadaptive competitors in practice. In [9], an adaptive optimal measurement matrix design has been studied in CS-based multiple-input multiple-output (MIMO) radar to improve the detection accuracy. In [10], an adaptive CS radar scheme is proposed where the transmission waveform and measurement matrix can be updated by the target scene information fed back by the recovery algorithm, which achieves better detection performance than the traditional CS radar system.

Generally speaking, the measurement noise in CS can be classified into two categories in terms of the generation mechanism [11]. The first category is the signal noise, i.e., jammers and interference in the transmission channel. The second category is the processing noise caused by the processing and acquisition hardware, i.e., the quantization error in the acquisition system. Most of the previous literatures focus on CS acquisition and recovery with the processing noise. The recent works in CS show that the measurement process would causes the noise folding phenomenon [12], which implies that the noise in the signal eventually is amplified by the measuring process. The study of has raised concerns from some scholars [13][14]. In [13], the authors evaluate the performance of the CS based wideband radio receiver in both signal noise and processing noise environments, and some effective suggestions are given for the CS receiver evaluation. In [14], an enhanced l_1 minimization recovery algorithm is developed for signal noise suppression, which has been proven that the algorithm providing relatively simple and precise theoretical guarantees. All the above studies can be summarized as optimization methods after sparse signals and noise have been acquired. Once the acquisition system faced with strong signal noise scene, the benefit from these optimization methods may be diminished. Nonetheless, signal noise suppression has not been taken into account from the perspective of measurement matrix optimization in the signal acquisition current system.

In this paper, we provide a new insight into sparse signal acquisition oriented toward strong signal noise scene. The mechanism of how does the signal noise exacerbates the recovery performance is investigated. An adapted selective compressive sensing (ASCS) scheme is proposed for signal noise suppression in the acquisition system. The measurement matrix can be adapted according to the noise strength so as to selective measure the signal noise, thus provides fewer noisy measurements. For robust noise priori estimation, the multiple measurement vectors (MMV) [15] model is used. A method of joint projection filtering in the compressive domain and the subspace estimation are proposed in this paper. We evaluate the performance via simulations and compare the proposed scheme with a non-adaptive implementation.

The rest of the paper is organized as follows. In section 2, we present the signal model and analyze the impact of measurement process on signal noise. Section 3 provides the proposed ASCS scheme. The simulation results are given in section 4. Finally, conclusions are given in Section 5.

Notation: Lower case and capital letters in bold denote, respectively, vectors and matrices. The superscript $(\bullet)^T$, $(\bullet)^H$, $(\bullet)^{-1}$ and $(\bullet)^{\dagger}$ represent the operators of transpose, Hermitian transpose, inverse and pseudo-inverse, respectively; The subscript $\|\bullet\|_{i\cdot}$ and $\|\bullet\|_{\cdot j}$ accounts for the i-th row and j-th column of a matrix; $\|\bullet\|_1$, $\|\bullet\|_2$ and $\|\bullet\|_F$ separately denote the l_1-norm, l_2-norm and Frobenius norm.

2. Signal Model

2.1 Compressive Sensing

The CS theory states that if a signal is compressible or sparse in a transform domain, it can be recovered exactly with high probability from much fewer samples than that required by the traditional Shannon-Nyquist sampling theorem. Without loss of generality, for any $\boldsymbol{x} \in \mathbb{C}^{N\times 1}$, if there exist unique coefficients $\{s_n\}_{n=0}^{N}$ such that

$$\boldsymbol{x} = \sum_{n=0}^{N-1} s_n \boldsymbol{\varphi}_n = \boldsymbol{\Psi s} \tag{1}$$

where $\boldsymbol{\Psi}$ denote an $N \times N$ orthogonal transform basis with the n-th column given by $\boldsymbol{\varphi}_n \in \mathbb{C}^{N\times 1}$, and $\boldsymbol{s} = [s_1, s_2, \cdots, s_N]^T$ is a complex-valued vector with length of N. The signal $\boldsymbol{x}$ is called K-sparse if no more than K elements of its sparse representation $\boldsymbol{s}$ are nonzero, i.e. $\|\boldsymbol{s}\|_0 = K$ with $K \ll N$. The support of $\boldsymbol{x}$ is

$$\mathrm{supp}(\boldsymbol{x}) = \{n \mid |s_n| > 0\} \subseteq [1, N] \tag{2}$$

In order to recover $\boldsymbol{x}$ one must identify $\mathrm{supp}(\boldsymbol{x})$. Therefore, a natural recovery strategy for signal recovery is support identification.

Now we consider a linear projection operator that computes $M\,(K < M < N)$ inner products between $\boldsymbol{x}$ and a set of vectors $\{\phi_m\}_{m=1}^{M}$

$$y_m = \langle \phi_m, \boldsymbol{x} \rangle = \phi_m^T \boldsymbol{x} \tag{3}$$

We collect the measurements and form a vector $\mathbf{y} = [y_1, y_2, \cdots, y_M]^T$. By arranging the projection operators $\{\phi_m^T\}_{m=1}^{M}$ as rows of an $M \times N$ measurement matrix $\mathbf{\Phi}$, the noisy measurement process in (3) can be represented as

$$\boldsymbol{y} = \mathbf{\Phi}\boldsymbol{x} + \boldsymbol{e} = \mathbf{\Phi}\mathbf{\Psi}\boldsymbol{s} + \boldsymbol{e} = \mathbf{\Theta}\boldsymbol{s} + \boldsymbol{e} \tag{4}$$

where $\boldsymbol{e} = [e_1, e_2, \cdots, e_M]^T$ represents the noisy environment effects with each entry e_m being zero mean Gaussian random variable with variance σ_e^2. As M is typically much smaller than N, the matrix $\mathbf{\Theta} = \mathbf{\Phi}\mathbf{\Psi}$ represents a dimensionality reduction since it maps $\mathbb{R}^N$ into $\mathbb{R}^M$. (4) is turned to be an underdetermined system. The sparse solutions $\hat{\boldsymbol{s}}$ to the linear inverse problem from (4) can be formulated as the following convex problem

$$\hat{\boldsymbol{s}} = \arg\min \|\boldsymbol{s}\|_0 \quad s.t. \quad \|\boldsymbol{y} - \mathbf{\Theta}\boldsymbol{s}\|_2^2 \le \sigma_e^2 \tag{5}$$

In general, this problem is NP-hard. [16] states that the l_0-norm optimization in (5) can be approximated by the l_1-norm relaxation with a bounded error under certain conditions

$$\hat{\boldsymbol{s}} = \arg\min \|\boldsymbol{s}\|_1 \quad s.t. \quad \|\boldsymbol{y} - \mathbf{\Theta}\boldsymbol{s}\|_2^2 \le \sigma_e^2 \tag{6}$$

To ensure stable recovery of sparse vector s by l_1-norm minimization, the matrix $\mathbf{\Theta}$ need satisfying the restricted isometry property (RIP) [17] of the order K with a very small constant δ_K, so that

$$(1-\delta_K)\|\boldsymbol{s}\|_2^2 \le \|\mathbf{\Theta}\boldsymbol{s}\|_2^2 \le (1+\delta_K)\|\boldsymbol{s}\|_2^2 \tag{7}$$

In other word, $\mathbf{\Theta}$ acts as an approximate isometry on the set of vectors that are K-sparse. Note that Gaussian matrices, Bernoulli matrix and uniformly random partial Fourier matrix provide reasonable constants for RIP. A typical means of solving (6) is through an unconstrained l_1-norm regularized formulation

$$\hat{\boldsymbol{s}} = \arg\min \left\{ \|\boldsymbol{y} - \mathbf{\Theta}\boldsymbol{s}\|_2^2 + \eta \|\boldsymbol{s}\|_1 \right\} \tag{8}$$

where η is a tradeoff parameter balancing the estimation quality. The basic framework in (8) can be solved by techniques such as greedy algorithms [18] and Bayesian algorithms.

2.2 Noise Folding in CS

The basic CS model in (4) is adequate when faced with the measured error or noise. However, in many practical scenarios, the signal itself is contaminated by the signal noise, which is not accounted for in (4). In [11], the authors present a generalized CS model

$$\boldsymbol{y} = \mathbf{\Phi}\tilde{\boldsymbol{x}} + \boldsymbol{e} = \mathbf{\Phi}\mathbf{\Psi}(\boldsymbol{s} + \boldsymbol{n}) + \boldsymbol{e} = \mathbf{\Theta}(\boldsymbol{s} + \boldsymbol{n}) + \boldsymbol{e} \tag{9}$$

where $\boldsymbol{n}$ stands for the white signal noise with variance σ_n^2, and $\boldsymbol{e}$ represents the processing noise. Basically, this is equivalent to stating that $\tilde{\boldsymbol{x}} = \mathbf{\Psi}(\boldsymbol{s} + \boldsymbol{n})$ is only approximately sparse. The noise situation in (9) is subtly different from the basic setting because the signal noise has been acted upon by the matrix $\mathbf{\Theta}$, and it is possible that $\mathbf{\Theta}\boldsymbol{n}$ could be potentially rather large. Our chief interest here is to understand how $\boldsymbol{n}$ impacts the recovery performance.

Before establishing our main result concerning white signal noise, some useful assumptions are suggested for our deduction. We suppose that the measurement matrix $\mathbf{\Theta} \in \mathbb{C}^{M \times N}$ fulfills the RIP of the order K and constant δ_K. Furthermore, we suppose that:

1). each row $\boldsymbol{\Theta}_{m\bullet}\,(m=1,2,\ldots,M)$ of $\boldsymbol{\Theta}$ is orthogonal to others, i.e., $\boldsymbol{\Theta}_{i\bullet}\boldsymbol{\Theta}_{j\bullet}^{H}=0\ \ (i\neq j)$, and each column of $\boldsymbol{\Theta}_{\bullet n}\,(n=1,2,\ldots,N)$ is normalized to 1, namely $\|\boldsymbol{\Theta}_{\bullet n}\|_2=1$.

2). each row has the same norm. Since $\|\boldsymbol{\Theta}\|_F=\sqrt{\sum_{m=1}^{M}\sum_{n=1}^{N}|\boldsymbol{\Theta}_{m,n}|^2}=\sqrt{\sum_{n=1}^{N}\|\boldsymbol{\Theta}_{\bullet,n}\|_2^2}=\sqrt{\sum_{m=1}^{M}\|\boldsymbol{\Theta}_{m,\bullet}\|_2^2}$, with the hypothesis in **1)**, we have $\|\boldsymbol{\Theta}_{m,n}\|_2^2={N}/{M}$.

3). acquisition noise $\boldsymbol{e}$ is ignored in our discussion, i.e. $\boldsymbol{e}=\mathbf{0}$.

In our formulation, we use the notation $\lambda_j(\boldsymbol{\Theta})$ to denote the j-th largest eigenvalue of $\boldsymbol{\Theta}$, and we use $s_j(\boldsymbol{\Theta})$ to denote the j-th largest singular value of $\boldsymbol{\Theta}$, thus we obtain $s_j(\boldsymbol{\Theta})=\lambda(\boldsymbol{\Theta}^H\boldsymbol{\Theta})$. To establishing our main result concerning white signal noise, a useful lemma is firstly cited, which has been proven in [19].

Lemma (Lemma 7.1 of [19]). Suppose that $\boldsymbol{\Theta}$ is a $M\times N$ matrix and let Λ be a set of indices with $\|\Lambda\|_0\le K$. If $\boldsymbol{\Theta}$ satisfies the RIP of order K and constant δ_K, for $k=1,2,\ldots,K$ we have

$$\frac{1}{\sqrt{1+\delta_K}}\le s_k(\boldsymbol{\Theta}_\Lambda^{\dagger})\le\frac{1}{\sqrt{1-\delta_K}} \tag{10}$$

We begin by noting that $\boldsymbol{\Theta}\boldsymbol{\Theta}^H=\dfrac{N}{M}\mathbf{I}_M$, the expectation of the measured noise power is

$$E\left\{\boldsymbol{\Theta}\boldsymbol{n}(\boldsymbol{\Theta}\boldsymbol{n})^H\right\}=\boldsymbol{\Theta}E\left\{\boldsymbol{n}\boldsymbol{n}^H\right\}\boldsymbol{\Theta}^H=\frac{N\sigma_n^2}{M}\mathbf{I}_M \tag{11}$$

which establishes $E\left\{\|\boldsymbol{\Theta}\boldsymbol{n}\|_2^2\right\}=E\left\{Tr\left\{\boldsymbol{\Theta}\boldsymbol{n}(\boldsymbol{\Theta}\boldsymbol{n})^H\right\}\right\}=Tr\left\{\dfrac{N\sigma_n^2}{M}\mathbf{I}_M\right\}=N\sigma_n^2$. From the RIP, we can get $\dfrac{\|\boldsymbol{\Theta}\boldsymbol{s}\|_2^2}{\|\boldsymbol{s}\|_2^2}\approx 1$, which implies that the sparse signal power hardly changed during the measurement process. In order to quantify the impaction of signal noise to the random measurement process, we defined the impaction factor $Gain_{noise}$ as ratio of the recovered noise power $E\left\{\|\hat{\boldsymbol{s}}-\boldsymbol{s}\|_2^2\right\}$ to the power of the noise component that attached to the sparse signal $E\left(\|\boldsymbol{n}_\Lambda\|_2^2\right)$, therefore $Gain_{noise}=\dfrac{E\left\{\|\hat{\boldsymbol{s}}-\boldsymbol{s}\|_2^2\right\}}{E\left\{\|\boldsymbol{n}_\Lambda\|_2^2\right\}}$. Let Λ to be the indices set with the elements represent the indexes corresponding to the location of nonzero elements in $\boldsymbol{s}$, i.e. $\Lambda=\mathrm{supp}(\boldsymbol{x})$. The least-squares optimal recovery of $\boldsymbol{s}$ restricted to the index set Λ is given by

$$\hat{\boldsymbol{s}}=\boldsymbol{s}+\boldsymbol{\Theta}_\Lambda^{\dagger}\boldsymbol{\Theta}\boldsymbol{n} \tag{12}$$

Since $\boldsymbol{\Theta}\boldsymbol{n}$ is a white Gaussian process, we have

$$\begin{aligned}E\left\{\|\hat{\boldsymbol{s}}-\boldsymbol{s}\|_2^2\right\}&=E\left\{\|\boldsymbol{\Theta}_\Lambda^{\dagger}\boldsymbol{\Theta}\boldsymbol{n}\|_2^2\right\}=E\left\{Tr\left\{\boldsymbol{\Theta}_\Lambda^{\dagger}\boldsymbol{\Theta}\boldsymbol{n}\left(\boldsymbol{\Theta}_\Lambda^{\dagger}\boldsymbol{\Theta}\boldsymbol{n}\right)^H\right\}\right\}\\&=Tr\left\{\boldsymbol{\Theta}_\Lambda^{\dagger}E\left\{\boldsymbol{\Theta}\boldsymbol{n}(\boldsymbol{\Theta}\boldsymbol{n})^H\right\}\left(\boldsymbol{\Theta}_\Lambda^{\dagger}\right)^H\right\}=\frac{N\sigma_n^2}{M}Tr\left\{\boldsymbol{\Theta}_\Lambda^{\dagger}\left(\boldsymbol{\Theta}_\Lambda^{\dagger}\right)^H\right\}\\&=\frac{N\sigma_n^2}{M}\sum_{k=1}^{K}s_k\left(\boldsymbol{\Theta}_\Lambda^{\dagger}\right)^2\end{aligned} \tag{13}$$

Combining (13) with (10) yields $E\left\{\|\hat{\boldsymbol{s}}-\boldsymbol{s}\|_2^2\right\}=\frac{NK\sigma_n^2}{M}$. In the event that the noise $\boldsymbol{n}$ is a white random vector, there exists $E\left\{\|\boldsymbol{n}_\Lambda\|_2^2\right\}=K\sigma_n^2$, thus

$$Gain_{noise}=\frac{E\left\{\|\hat{\boldsymbol{s}}-\boldsymbol{s}\|_2^2\right\}}{E\left\{\|\boldsymbol{n}_\Lambda\|_2^2\right\}}=\frac{N}{M} \tag{14}$$

From which we observe that the noise added to the signal itself can be highly amplified by the measurement process as $M = N$. In the literature, this effect is known as noise folding.

3. Adaptive Compressive Sensing

Although prior research have validated the benefits of exploiting RIP in measurement design [9][10], such as improving the recovery probability, decreasing the recovery error and so on, these benefits diminished when faced with strong signal noise scene. From the above analysis, the expected $Gain_{noise}$ closely related to parameters M and N, which account for the number of rows and columns in $\mathbf{\Theta}$. Generally, M is related to the RIP condition (which is bound by K, N and δ_K), and N in $Gain_{noise}$ is related to the measured support of noise in $\mathbf{\Theta}$. However, only $\mathbf{\Theta}_\Lambda$ contribute to the sparse vector $\boldsymbol{s}$ and $\mathbf{\Theta}_\Lambda^\perp$ entirely measured the signal noise. In the traditional Shannon-Nyquist sampling system, to avoid the noise off the passband, an antialiasing filter is applied before the sampling process. Inspired by the necessity of antialiasing filtering in bandpass signal sampling, a selective measuring scheme is proposed in this paper. The measurement matrix would only sense the interested spectrum, where most likely the sparse spectrum lying.

The measurement matrix in our scheme is modified into

$$\mathbf{\Phi}=\Im_\Omega(\mathbf{A})\mathbf{\Psi}^{-1} \tag{15}$$

where $\mathbf{A}\in\mathbb{C}^{M\times N}$ is a random matrix, Ω is an index set. $\Im_\Omega(\mathbf{A})$ is defined as a selective operation which setting the n-th $(n\in\Omega)$ column of $\mathbf{A}$ to zeroes, act as an antialiasing filter in our scheme.

3.1 Projection Filtering in the Compressed Domain

The core of the proposed scheme is estimation of the index set Ω, where most likely the noise spectrum lying. It is necessary for us to extract the information that each vector $\mathbf{\Theta}_{\bullet n}$ hide in $\boldsymbol{y}$. The simplest way is to projection $\boldsymbol{y}$ to each vector $\mathbf{\Theta}_{\bullet n}$. However, due to the nonorthogonality between the columns of the matrix $\mathbf{\Theta}$, the projection results would interfere with each other, and the low SNR scene increased the difficulty of information extraction. To minimize the projection interference, a set of projection filters are applied. The output of the n-th $(n=1,2,\ldots,N)$ filter is formed as

$$z_n=\langle\boldsymbol{h}_n,\boldsymbol{y}\rangle=\boldsymbol{h}_n^H\boldsymbol{y} \tag{16}$$

The output energy is defined as $E\left\{\|z_n\|_2^2\right\}=\boldsymbol{h}_n^H\mathbf{R}_y\boldsymbol{h}_n$ with the correlation matrix of the measured signal $\mathbf{R}_y=E\left\{\boldsymbol{y}\boldsymbol{y}^H\right\}$. Our objective function is minimized the output energy. In order to avoid the trivial solution such as $\mathbf{h}_n = \mathbf{0}$, a set of linear constraints are added to the objective function, which can be expressed as

$$\min E\left\{\left\|z_n\right\|_2^2\right\} \quad s.t. \quad \boldsymbol{h}_n \boldsymbol{\Theta}_{\bullet n} = 1 \tag{17}$$

The minimum output energy can be achieved with a proper choice of $\boldsymbol{h}_{nopt}$. We can solve the general constrained minimization problem of equation (17) to obtain $\boldsymbol{h}_{nopt}$ by applying Lagrange multiplier method, which resulting in the following unconstrained objective function

$$f\left(\boldsymbol{h}_{nopt}\right) = \boldsymbol{h}_n^H \mathbf{R}_y \boldsymbol{h}_n + \lambda\left(1 - \boldsymbol{h}_n \boldsymbol{\Theta}_{\bullet n}\right) \tag{18}$$

To force the gradient of the objective function to be zero, i.e., $\frac{\partial f\left(\boldsymbol{h}_{nopt}\right)}{\partial \boldsymbol{h}_n} = 0$, we obtain $\lambda = \frac{1}{\boldsymbol{\Theta}_{\bullet n}^H \mathbf{R}_y^{-1} \boldsymbol{\Theta}_{\bullet n}}$. The optimal values for $\boldsymbol{h}_{nopt}$ that minimize the objective function can be evaluated as follows

$$\boldsymbol{h}_{nopt} = \frac{R_y^{-1} \boldsymbol{\Theta}_{\bullet n}}{\boldsymbol{\Theta}_{\bullet n}^H \mathbf{R}_y^{-1} \boldsymbol{\Theta}_{\bullet n}} \tag{19}$$

The optimal output energy of z_n is $E\left\{\left\|z_{nopt}\right\|_2^2\right\} = \left(\boldsymbol{\Theta}_{\bullet n}^H \mathbf{R}_y^{-1} \boldsymbol{\Theta}_{\bullet n}\right)^{-1}$ and the desired output of the filter banks $\boldsymbol{z} = \left[z_1, z_2, \mathrm{L}\ , z_N\right]$ are

$$\boldsymbol{z} = \mathbf{Q}^H \boldsymbol{y} \tag{20}$$

with $\mathbf{Q} = \mathbf{D}\mathbf{R}_y^{-1}\boldsymbol{\Theta}$ and $\mathbf{D} = diag\left(d_1, d_2, \mathrm{L}\ , d_N\right)$ denotes a diagonal matrix with principal diagonal elements being $d_1, d_2, \mathrm{L}\ , d_N$ in turn, where $d_n = \boldsymbol{\Theta}_{\bullet n}^H \mathbf{R}_y^{-1} \boldsymbol{\Theta}_{\bullet n}$. Note that in the ideal case the matrix $\mathbf{Q}$ can be estimated precisely. Therefore, the Lagrange method converges to the optimal solution in a single iteration, as expected for a quadratic objective function.

3.2 Noise Information Estimation Using Subspace Method

The model in (9) is a typical single measurement vector (SMV) model. When a sequence of measurement vectors are available, (9) can be extended to the multiple measurement vectors (MMV) model, which provides informative coupling between the vectors. The noisy MMV problem can be stated as solving the following underdetermined systems of equations

$$\boldsymbol{y}_l = \boldsymbol{\Theta}\boldsymbol{s}_l + \boldsymbol{e}_l \quad , \quad l = 1, 2, \cdots, L \tag{21}$$

where L is the number of measurement vectors. Since the matrix $\boldsymbol{\Theta}$ is common to each of the representation problem, (21) can be rewritten as

$$\mathbf{Y} = \boldsymbol{\Theta}\mathbf{S} + \mathbf{E} \tag{22}$$

where $\mathbf{Y} = \left[\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_L\right]$, $\mathbf{S} = \left[\boldsymbol{s}_1, \boldsymbol{s}_2, \cdots, \boldsymbol{s}_L\right]$ and $\mathbf{E} = \left[\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_L\right]$. Additional assumptions are that the solution vectors $\boldsymbol{s}_1, \boldsymbol{s}_2, \cdots, \boldsymbol{s}_L$ are sparse and have the same sparsity profile. It is equal to state that $\mathbf{S}$ is an unknown source matrix with nonzero rows representing the targets. In many applications, such as wireless communication and radar detection, the spectrum that signals occupied is slowly time-varying, hence the common sparsity assumption is valid.

The presence of multiple measurements can be helpful in estimating the set $\boldsymbol{\Omega}$. With multiple measurements, the desired output in (25) can be represented as

$$\mathbf{Z} = \mathbf{Q}^H \mathbf{Y} = \mathbf{Q}^H \mathbf{\Theta} \mathbf{S} + \mathbf{N} = \mathbf{H}\mathbf{S} + \mathbf{N} \tag{23}$$

where $\mathbf{N} = \mathbf{Q}^H \mathbf{E}$ and $\mathbf{H} = \mathbf{Q}^H \mathbf{\Theta}$. The covariance matrix of the filtered signal is $\mathbf{R}_Z = E\{\mathbf{Z}\mathbf{Z}^H\}$. The eigenvalue decomposition of $\mathbf{R}_Z$ is

$$\mathbf{R}_Z = \mathbf{U}\mathbf{\Sigma}\mathbf{U}^H = \sum_{j=1}^{N} \lambda_j \boldsymbol{u}_j \boldsymbol{u}_j^H = \mathbf{U}_s \mathbf{\Sigma}_s \mathbf{U}_s^H + \mathbf{U}_n \mathbf{\Sigma}_n \mathbf{U}_n^H \tag{24}$$

where $\mathbf{\Sigma} = diag(\lambda_1, \lambda_2, \cdots, \lambda_N)$, the eigenvalues are complied with $\lambda_1 \geq \cdots \geq \lambda_K > \lambda_{K+1} = \cdots = \lambda_N = \sigma_N^2$. The eigenvectors $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_k$ corresponding to the K larger eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_N$ construct signal subspace $\mathbf{U}_s = [\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_k]$, with $\mathbf{\Sigma}_s = [\lambda_1, \lambda_2, \cdots, \lambda_K]$. Similarly, the later $N-K$ eigenvalue are depending on the noise and their numeric values are σ_N^2. The eigenvectors $\boldsymbol{u}_{K+1}, \boldsymbol{u}_{K+2}, \cdots, \boldsymbol{u}_N$ corresponding to $\lambda_{K+1}, \lambda_{K+2}, \cdots, \lambda_N$ construct noise subspace $\mathbf{U}_n = [\boldsymbol{u}_{K+1}, \boldsymbol{u}_{K+2}, \cdots, \boldsymbol{u}_N]$, and $\mathbf{\Sigma}_n = [\lambda_{K+1}, \lambda_{K+2}, \cdots, \lambda_N]$. Let Λ stands for the index set corresponding to the K nonzero rows of $\mathbf{S}$, we have

$$\mathbf{R}_z \mathbf{U}_n = \mathbf{H}_\Lambda \mathbf{R}_{S_\Lambda} \mathbf{H}_\Lambda^H \mathbf{U}_n + \sigma_n^2 \mathbf{U}_n = \sigma_n^2 \mathbf{U}_n \tag{25}$$

where $\mathbf{R}_{S_\Lambda} = E\{S_\Lambda S_\Lambda^H\}$. It can be seen from (25) that $\mathbf{H}_\Lambda \mathbf{R}_{S_\Lambda} \mathbf{H}_\Lambda^H \mathbf{U}_n = \mathbf{0}$. Since $\mathbf{R}_{S_\Lambda}$ is a non-singular matrix, we get $\mathbf{H}_\Lambda^H \mathbf{U}_n = \mathbf{0}$, thus $\mathbf{U}_n^H \mathbf{H}_\Lambda = \mathbf{0}$. This indicates that the column vectors in $\mathbf{H}_\Lambda$ is orthogonal to the subspace of the noise. The spectrum function of sparse location can be deduced

$$f_n = \frac{1}{\mathbf{H}_{\bullet n}^H \mathbf{U}_n \mathbf{U}_n^H \mathbf{H}_{\bullet n}} \quad , \quad n = 1, 2, \cdots, N \tag{26}$$

With the change of n, there would be K large values in (26), which correspond to the sparse position. The peak values are obvious with high SNR, but this superiority dwindles under the condition of low SNR. However, the non-orthogonality between $\mathbf{H}_\Lambda$ and $\mathbf{U}_n$ barely affected. Hence the index corresponding to the smallest $N-K$ values could be treated as the positions of the noise, which should be ignored by the measurement process for noise suppression. In order to avoid causing any confusion with strong signal noise level, we consider the index corresponding to the smallest $P(2K < \mathrm{P} < \mathrm{N})$ values in (26) seemed a high possibility to be noise. (26) also can be expressed as

$$g_n = 1/ f_n = Tr\{\mathbf{P}_n \mathbf{U}_n \mathbf{U}_n^H\} \quad , \quad n = 1, 2, \cdots, N \tag{27}$$

where $\mathbf{P}_n = \mathbf{H}_{\bullet n} \left[\mathbf{H}_{\bullet n}^H \mathbf{H}_{\bullet n}\right]^{-1} \mathbf{H}_{\bullet n}^H$, which represents the projection matrix of $\mathbf{H}_{\bullet n}$.

3.3 Signal Reconstruction

Recovery of the signal from the linear projections can be accomplished by solving (8). A variety of optimization algorithms are available for the recovery problem, such as Orthogonal Matching Pursuit (OMP), Compressive Sampling Matched Pursuit (CoSaMP), FOCal Underdetermined System Solver (FOCUSS) and Sparse Bayesian Learning (SBL). The regularized M-FOCUSS [15] is chosen for its perfect compromise between computation complexity and reconstruction accuracy, which can be summarized as the following iteration steps

$$
\begin{aligned}
&\mathbf{W}_{d+1} = diag\left(c_d\left[n\right]^{1-0.5p}\right) \\
&\quad where \quad c_d\left[n\right] = \sqrt{\sum_{l=1}^{L}\left(s_d^{(l)}\left[n\right]\right)^2} \quad , p \in \left[0,2\right] \\
&\mathbf{B}_{d+1} = \mathbf{\Theta}_{d+1}^{H}\left(\mathbf{\Theta}_{d+1}\mathbf{\Theta}_{d+1}^{H} + \beta\mathbf{I}\right)^{-1}\mathbf{Y} \\
&\quad where \quad \mathbf{\Theta}_{d+1} = \mathbf{\Theta}\mathbf{W}_{d+1} \quad with \quad \beta \geq 0 \\
&\tilde{\mathbf{S}}_{d+1} = \mathbf{W}_{d+1}\mathbf{B}_{d+1}
\end{aligned} \tag{28}
$$

where β stands for the regularization parameter, the p-norm always set to $p = 0.8$ as suggested by the authors for robust solution.

The regularized M-FOCUSS algorithm can be treat as solving at each iteration a weighted least squares. The initial solution $\tilde{\mathbf{S}}_1$ was firstly set to a nonzero weight matrix, with the iteration of the algorithm, $\tilde{\mathbf{S}}$ would tend to be stable. The algorithm could be terminated once the maximum iteration number reached or a convergence criterion has been satisfied

$$
\frac{\left\|\tilde{\mathbf{S}}_{d+1} - \tilde{\mathbf{S}}_d\right\|_F}{\left\|\tilde{\mathbf{S}}_d\right\|_F} < \varepsilon \tag{29}
$$

where ε is a user-selected parameter. The proposed adaptive CS scheme can be summarized as following

(1). Initialize measurement matrix $\mathbf{\Phi}$ as (15), set $\Omega = \varnothing$. Collect the compressed data $\mathbf{Y}$, and calculate $\mathbf{Z}$ using (23).

(2). Estimate the compressed signal covariance matrix $\mathbf{R}_Z$, then perform an EVD for $\mathbf{R}_Z$, and isolate the subspace of the noise $\mathbf{U}_n$.

(3). Compute the spectrum function in (26) or in (27), select the index corresponding to the P smallest value in (26) or the P largest value in (27).

(4). Update measurement matrix $\mathbf{\Phi}$ using (15).

(5). Measure the signal $\boldsymbol{x}$ using the updated measurement matrix $\mathbf{\Phi}$, recovery the sparse information using the iterations of (28) until (29) being satisfied.

4. Experimental Results and Analysis

Extensive computer experiments have been conducted and a few representative and informative results are presented. We consider a signal sparse in Fourier domain. Unless specifically stated otherwise, the following conditions are applied. We set $N = 150$ and $K = 3$, the compressive measured dimension is $M = 50$, the dimension of the multiple measurement vectors is $L = 10$, and the selective parameter is set to $P = 50$. In our simulation, the SNR is defined as SNR = $20\log\left(\left\|\mathbf{S}\right\|_2 / \left\|\mathbf{N}\right\|_2\right)$, where $\mathbf{N}$ stands for the signal noise matrix. The proposed adaptive method is compared to the adaptive compressive sensing (ACS) method in [10] (using M-FOCUSS algorithm for sparse reconstruction) and traditional nonadaptive scheme with the recovery algorithms OMP, M-FOCUSS and MSBL. To assess the optimization performance of the proposed scheme, 1000 Monte Carlo simulations are conducted. In each trial the initial measurement matrix was created with columns uniformly drawn from the surface of a unit hypersphere, and the source matrix $\mathbf{S} \in \mathbb{C}^{N\times L}$ was randomly generated with K nonzero rows (i.e., sources). In each trial the indexes of the sources were randomly chosen. Two measures were applied for performance assessment, the first one is the failure rate defined in [20], and a failed trial was recognized if the indexes of estimated sources with the largest norms were not the same as the true indexes. The second one is the mean square error (MSE) defined as $\|\hat{\mathbf{s}}-\mathbf{s}\|_2^2 / \|\mathbf{s}\|_2^2$, where $\hat{\mathbf{s}}$ represents the reconstructed sources $\mathbf{s}$.

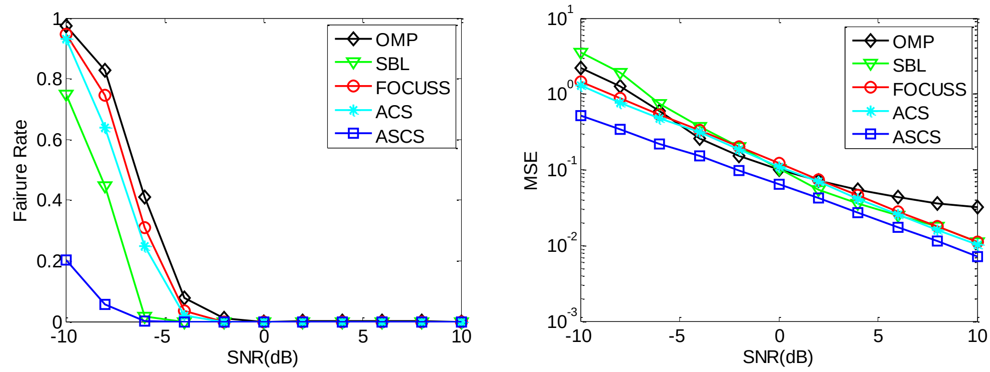


Fig. 1. Performance comparison with various SNR

We explored the recovery performance with different signal noise levels. **Fig. 1** depicts the performance curve, from which we conclude that the adaptive scheme outperform the ACS method and the nonadaptive one with the same noise environment. With the increasing SNR, as expected, both schemes would achieve better performance. But meanwhile we noticed when SNR ≤ -4dB, the benefit from multiple measurement vectors diminished, and the failure rate deteriorate sharply. One obvious observation is that the proposed adaptive scheme would achieve lower failure rate with extreme noise conditions. According to the RIP in CS, once $M \geq C\mu(\mathbf{\Theta}) K \log N$ (C is a constant and $\mu(\mathbf{\Theta})$ is defined as the maximum absolute value of the normalized inner product between all columns in $\mathbf{\Theta}$), one could accurately recovery the sparse vector with high probability. In our setup, the RIP is satisfied, and therefore further optimization for the measurement matrix couldn't improve the performance significantly. However, our ASCS scheme would suppress the signal noise, hence provides high-precision recovery performance.

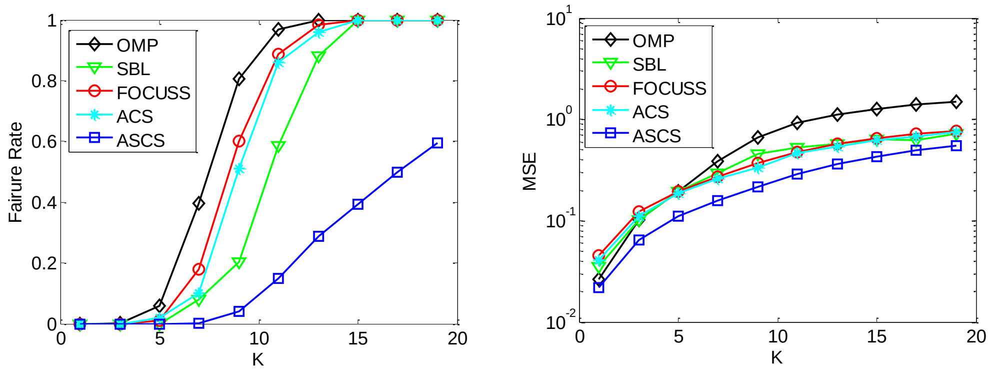


Fig. 2. Performance comparison with different sparsity K

Fig. 2 depicts simulation results with different signal sparsity, the SNR is set to 0dB. As this figure shows, the increasing of sparsity K leads to the decreasing of the recovery performance, but the proposed adaptive method still achieves better performance with respect to failure rate and MSE. This phenomenon can be explained as follows. The configured parameters in our simulation is only robust for K $\leq$ 6 according to the RIP[15]. In the case of K $\geq$ 7, the RIP diminished, thus the recovery algorithm fail to recovery S with high probability. The proposed adaptive method enables fewer signal noise being measured through the measurement process, therefore the adaptive method performs better than the nonadaptive ones with the same configuration.

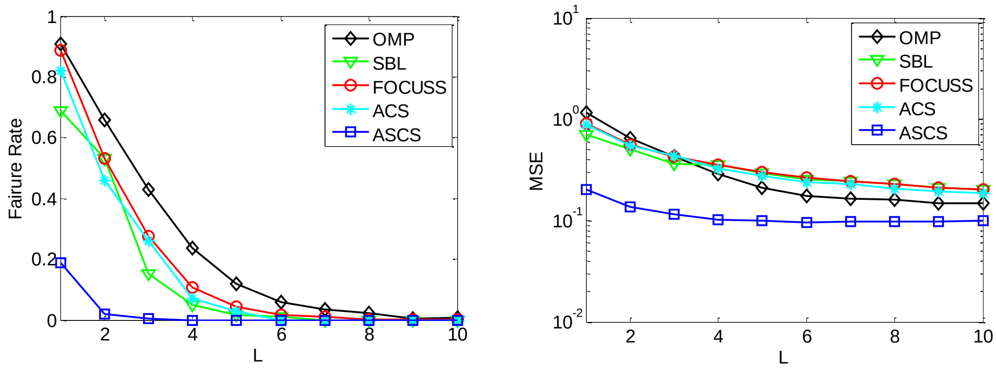


Fig. 3. Performance under different number of measurements L

Fig. 3 shows that the failure rate decreases exponentially with the number of the measurement vectors L, and the increasing L narrows the performance gap between the adaptive scheme and the nonadaptive one. In practical applications, under the common sparsity assumption of source S, we cannot obtain many measurement vectors, as the sparsity profile of practical signals is time-varying, such as frequency hopping system. So the common sparsity assumption is valid for only a small L in the MMV model. Future research will pay much attention to this problem.

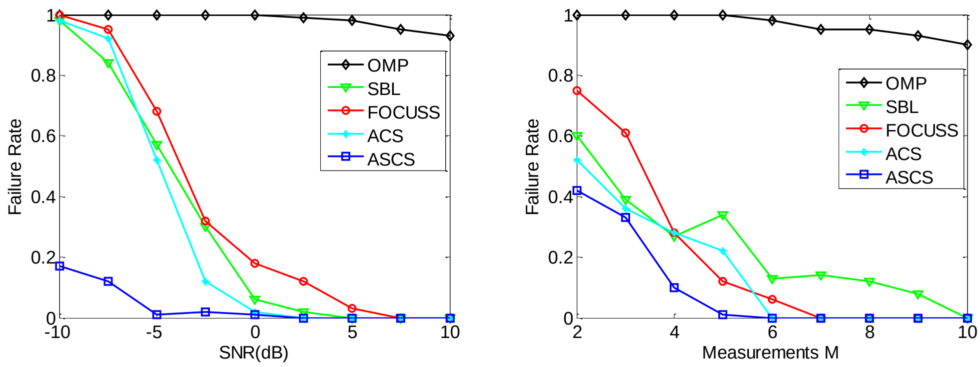


Fig. 4. Performance comparison with applied in spatial compressive sensing MIMO radar

Finally, we investigated the application of the proposed method for diection-of-arrival (DOA) estimation in spatial CS based multip-input and multip-output (MIMO) radar system [21]. In this application, the MIMO radar system is configurated with 10 transmit antennas, 10 receiver antennas, the snapshots number is 5, and 3 targets are located in the far field with DOA $\theta = [15, 40, 65]$. Unlike the Fourier basis that used in the above simulation, sparse dictionary in the application is consist of a series of interesting steering vectors with angel range from 0° -90° and resolution is 0.25°. **Fig. 4** depicts the performance comparson with different SNR and different measurements. As shown in the figures, the OMP method owns high failure rate, this is caused by the severe mutual coherence between the atoms of the dictionary. The greedy OMP algorithm ensures the residual is orthogonal to the atoms that chosen in the last iteration, which may destroy the information hiding in the residual when updating the new residual. Thanks for the noise suppression function, the proposed scheme could provide almost precise estimation results.

5. Conclusion

In this paper, we proposed an ASCS scheme for signal noise suppression in CS based signal acquisition system. A computational framework for the measurement matrix design is investigated, which transforms the measurement matrix design into the noise priori estimation. A two-step process is developed for locating the noise spectrum precisely. A set of projection filter banks are firstly used for minimizing the projection interferences. A subspace method is then applied for the noise information estimation. Simulation results demonstrated the effectiveness of the proposed scheme. From the view point of future implementation, measurement noise should be taken into consideration in the system, and more efficient algorithms have to be developed for source pre-estimation with low SNR. On the other hand, how to deal with the real world signal (e.g., image, video, or audio) is a problem need for further study.

References

[1] D.L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289-1306, April 2006. Article (CrossRef Link)

[2] F. Wen, Y. Tao, G. Zhang, "Analog to Information Conversion using Multi-Comparator based Integrate-and-Fire Sampler," *Electronics Letters*, vol. 51, no. 3, pp. 246-247, February 2015. Article (CrossRef Link)

[3] Y.J. Li, R.F. Song, "A new compressive feedback scheme based on distributed compressed sensing for time-correlated mimo channel," *KSII Transactions on Internet and Information Systems*, vol.6, no.2, pp.580-592, Feb. 2012. Article (CrossRef Link)

[4] H. Anh, I.Koo, "Primary user localization using Bayesian compressive sensing and path-loss exponent estimation for cognitive radio networks," *KSII Transactions on Internet and Information Systems*, vol.7, no.10, pp.2338-2356, OCT. 2013. Article (CrossRef Link)

[5] H.Q. Gao, R.F. Song, "Distributed compressive sensing based channel feedback scheme for massive antenna arrays with spatial correlation," *KSII Transactions on Internet and Information Systems*, vol.8, no.1, pp.108-122, Jan 2014. Article (CrossRef Link)

[6] S.H. Ji, X. Ya, L. Carin, "Bayesian Compressive Sensing," *IEEE Transactions on Signal Processing*, vol.56, no.6, pp.2346-2356, June 2008. Article (CrossRef Link)

[7] S.H. Ji; D. Dunson, L. Carin, "Multitask Compressive Sensing," *IEEE Transactions on Signal Processing*, vol.57, no.1, pp.92-106, Jan. 2009. Article (CrossRef Link)

[8] R.P. Li; Z.F. Zhao; Y. Zhang, J. Palicot, H.G. Zhang, "Adaptive multi-task compressive sensing for localization in wireless local area networks," *IET Communications*, vol.8, no.10, pp.1736-1744, July 2014. Article (CrossRef Link)

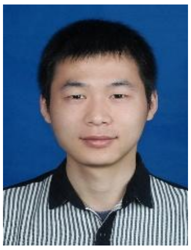
[9] Y. Yao; A.P. Petropulu; H.V. Poor, "Measurement matrix design for compressive sensing-based mimo radar," *IEEE Transactions on Signal Processing*, vol.59, no.11, pp.5338-5352, Nov. 2011. Article (CrossRef Link)

[10] J.D. Zhang, D.Y. Zhu, G. Zhang, "Adaptive compressed sensing radar oriented toward cognitive detection in dynamic sparse target scene," *IEEE Transactions on Signal Processing*, vol.60, no.4, pp.1718-1729, April 2012. Article (CrossRef Link)

[11] M.A. Iwen, A.H. Tewfik, "Adaptive strategies for target detection and localization in noisy environments," *IEEE Transactions on Signal Processing*, vol.60, no.5, pp.2344-2353, May 2012. Article (CrossRef Link)

[12] E. Arias-Castro, Y.C. Eldar, "Noise Folding in Compressed Sensing," *IEEE Signal Processing Letters*, vol.18, no.8, pp.478-481, Aug. 2011. Article (CrossRef Link)

[13] M.A. Davenport, J.N. Laska, J. Treichler, R.G. Baraniuk, "The pros and cons of compressive sensing for wideband signal acquisition: noise folding versus dynamic range," *IEEE Transactions on Signal Processing*, vol.60, no.9, pp.4628-4642, Sept. 2012. Article (CrossRef Link)
[14] A. Marco, F. Massimo, P. Steffen, "Damping noise-folding and enhanced support recovery in compressed sensing," *arXiv preprint arXiv*: 1307-5725, 2013. Article (CrossRef Link)
[15] S.F. Cotter, B.D. Rao, K. Engan, K. Kreutz-Delgado, "Sparse solutions to linear inverse problems with multiple measurement vectors," *IEEE Transactions on Signal Processing*, vol.53, no.7, pp.2477-2488, July 2005. Article (CrossRef Link)
[16] E.J. Candes, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Mathematique*, vol.346, no.9, pp.589-592, 2008. Article (CrossRef Link)
[17] E.J. Candes, T. Tao, "Decoding by linear programming," *IEEE Transactions on Information Theory*, vol.51, no.12, pp.4203-4215, Dec. 2005. Article (CrossRef Link)
[18] J.A. Tropp, A.C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol.53, no.12, pp.4655-4666, Dec. 2007. Article (CrossRef Link)
[19] M.A. Davenport, "Random observations on random observations: sparse signal acquisition and processing," *PhD thesis*, Rice University, 2010. Article (CrossRef Link)
[20] D.P. Wipf, B.D. Rao, "An empirical Bayesian strategy for solving the simultaneous sparse approximation problem," *IEEE Transactions on Signal Processing*, vol.55, no.7, pp.3704-3716, July 2007. Article (CrossRef Link)
[21] M. Rossi, A. M. Haimovich, Y. C. Eldar, "Spatial compressive sensing for MIMO radar", *IEEE Transactions on Signal Processing*, vol.62, no.2, pp.419-430, Jan.15, 2014. Article (CrossRef Link)



Fangqing Wen was born in 1988. He received the B.S. degree in electronic engineering from Hubei University of Automotive Technology, Shiyan, China, 2011. From 2011 to 2013, he was pursuing the post-graduate study in College of Electronics and Information Engineering, Nanjing University of Aeronautics and Astronautics (NUAA). He is a Ph.D. candidate in College of Electronics and Information Engineering, NUAA, His research interests are compressive sensing, wireless communication and signal processing.
E-mail: wfqsyyy@163.com

Gong Zhang was born in 1964. He received the Ph.D. degree in electronic engineering from the Nanjing University of Aeronautics and Astronautics (NUAA), Nanjing, China, in 2002. From 1990 to 1998, he was a Member of Technical Staff at No724 Institute China Shipbuilding Industry Corporation (CSIC), Nanjing. Since 1998, he has been with the College of Electronic and Information Engineering, NUAA, where he is currently a Professor. His research interests include radar signal processing and classification recognition. Dr. Zhang is a member of Committee of Electromagnetic Information, Chinese Society of Astronautics (CEI-CSA) and a Senior Member of the Chinese Institute of Electronics (CIE).
E-mail: wfqzhc@163.com

Ben De was born in 1938. He received the bachelor degree in radar engineering from the Harbin Institute of Technology, Harbin, China, in 1963. Since 1964, he was a member of technical staff at Nanjing Research Institute of Electronics Technology, Nanjing. He is currently a Professor in College of Electronic and Information Engineering, NUAA. His research interests include radar system and radar signal processing. Mr. Ben is a member of Academician of China Engineering Academy.
E-mail: wfqzhc@163.com