EURASIP Journal on Wireless Communications and Networking
a SpringerOpen Journal

RESEARCH Open Access

Sparse channel estimation of MIMO-OFDM systems with unconstrained smoothed l_0-norm-regularized least squares compressed sensing

Xinrong Ye[1,2*], Wei-Ping Zhu[1,3], Aiqing Zhang[1,2] and Jun Yan[1]

Abstract

This paper investigates the sparse channel estimation issue of multiple-input multiple-output orthogonal frequency division multiplexing (MIMO-OFDM) systems. Beginning with the formulation of least squares (LS) solution to sparse MIMO-OFDM channel estimation, a compressed channel sensing (CCS) framework based on the new smoothed l_0-norm-regularized least squares (l_2-Sl_0) algorithm is proposed. Three methods, namely quasi-Newton, conjugate gradient, and optimization in the null and complement spaces of the measurement matrix, are then proposed to solve the l_2-Sl_0 unconstrained optimization problem. Moreover, the two former are also applied to solve the l_2-Sl_0 channel estimation. A number of computer simulation-based experiments are conducted showing a better reconstruction accuracy of the l_2-Sl_0 algorithm as compared with the smoothed l_0-norm (Sl_0) algorithm in the presence of noise. The proposed CCS approach can save nearly 25% pilot signals to maintain the same mean square error (MSE) and bit error rate (BER) performances as given by the conventional LS method.

Keywords: Sparse channel estimation; Smoothed l_0-norm; l_2-norm; MIMO-OFDM

1. Introduction

Coherent detection and equalization in multiple input multiple output orthogonal frequency division multiplexing (MIMO-OFDM [1]) systems require channel state information (CSI) at the receiver. In real wireless environments, however, the CSI is not known. Therefore, channel estimation is of crucial importance to MIMO-OFDM systems. In various wireless propagation environments, the channel may consist of only a few dominant propagation (non-zero) paths, even though it has a large propagation delay. Thus, the channel impulse response has a sparse nature [2-4]. However, conventional methods, such as least squares (LS), ignore this prior information about the unknown channel leading to lower spectral efficiency. Recently, sparse channel estimation with an objective of decreasing the training sequence to improve spectral efficiency is becoming a hot research topic.

Previously reported approaches for sparse channel estimation can broadly be categorized into two types, namely the most significant tap (MST) detection and compressed channel sensing (CCS). The MST detection methods [4-6] used a measure to determine if a channel tap was non-zero ('active'). The disadvantage of this type of methods is that a large number of pilots are needed to render an accurate MST detection and effective channel estimation. The CCS methods are based on the compressed sensing (CS [7]) technology. In [8], the authors formalized the notion of multipath sparsity and proposed the CCS approach. In [9], the orthogonal matching pursuit (OMP) and basis pursuit (BP) algorithms were applied to estimate underwater acoustic channels with large Doppler spread. In [10], the authors proposed an overcomplete basis for doubly selective channels and a metric called localized coherence for selecting training signals to ensure good estimation performance. In [11], a CCS approach for doubly selective channels and a sparsity-enhancing basis expansion with a method for optimizing it were proposed. In [12], two criteria as guiding principles to optimize the pilot pattern for CCS in OFDM systems were proposed. Methods of this type utilize the prior sparse information

* Correspondence: yaya_ye@126.com
[1]Institute of Signal Processing and Transmission, Nanjing University of Posts and Telecommunications, Nanjing 210003, China
[2]College of Physics and Electronic Information, Anhui Normal University, Wuhu 241000, China
Full list of author information is available at the end of the article

of the unknown channel and the advantage of CS and thus can improve the spectral efficiency by reducing the number of pilot symbols to be transmitted.

Different from literatures [9-12] that used the existing sparse reconstruction algorithms for CCS in OFDM or single carrier systems, we aim to exploit a novel reconstruction algorithm for CCS in MIMO-OFDM systems. The proposed smoothed l_0-norm-regularized least squares reconstruction algorithm is named l_2-Sl_0 in this paper, which differs from the smoothed l_0-norm reconstruction algorithm (Sl_0 [13]) in two aspects. First, Sl_0 is a constrained optimization problem which is solved in [13] using the steepest descent approach. However, l_2-Sl_0 is an unconstrained minimization problem which is to be solved in this paper by using three methods, namely quasi-Newton approach, conjugate gradient approach, and optimization in the null and complement spaces of the measurement matrix. Second, unlike the Sl_0 using a fixed step size to control the decrease of the parameter σ, which determines the degree of smoothness and the approximation accuracy of l_0-norm, l_2-Sl_0 uses a variable one. Simulation results show that the proposed l_2-Sl_0 reconstruction approach outperforms the Sl_0 approach in the presence of noise, and at the cost of slightly more computational time, the CCS approach using l_2-Sl_0 in conjunction with conjugate gradient yields a performance slightly better than that of the CCS method using fast iterative shrinkage-thresholding algorithm (FISTA [14]) or orthogonal matching pursuit (OMP [15]) algorithm.

The remainder of the paper is organized as follows: Section 2 formulates the sparse channel estimation problem of MIMO-OFDM systems based on LS, ideal-LS, and compressed sensing. Section 3 presents three sparse reconstruction algorithms using the proposed l_2-Sl_0 objective function, based on which a new CS-based sparse channel estimation approach is developed. Section 4 comprises a number of experiments showing a better reconstruction accuracy of the l_2-Sl_0-based method as compared with the Sl_0 algorithm, and a higher spectrum efficiency of the sparse channel estimation employing l_2-Sl_0 than that using the LS method. Section 5 concludes this paper by highlighting some of the contributions presented.

2. The sparse channel estimation problem of MIMO-OFDM systems

Consider a similar MIMO-OFDM system as described in [16] with N_T transmit and N_R receive antennas. The MIMO channel can be characterized by an array of L-tap finite impulse response (FIR) filters given by a number of $N_R \times N_T$ matrices $\mathbf{H}(n)$, $(n = 0,1,\ldots,L-1)$, whose (i_R,i_T)-th element $h_{i_R,i_T}(l), (0 \le l \le L-1)$ represents the l-th tap of the channel response between the i_R-th receive antenna and the i_T-th transmit antenna. In the case of uniform sampling, a wireless channel can often be modeled as a sparse channel [17-19], i.e., only a few elements are nonzero in $\left[h_{i_R,i_T}(0), h_{i_R,i_T}(1), \ldots, h_{i_R,i_T}(L-1)\right]$. If the length of the cyclic prefix (CP) is not less than the channel length L, the received pilot signal in i_R-th receiver antenna can be written as

$$\mathbf{Y}_{i_R,pilot} = \left[diag\left(\mathbf{X}_{1,pilot}\right)\mathbf{F}_{pilot}, \ldots, diag\left(\mathbf{X}_{N_T,pilot}\right)\mathbf{F}_{pilot}\right]\mathbf{h}_{i_R} + \mathbf{n}_{i_R,pilot}, \tag{1}$$

where $\mathbf{X}_{i_T,pilot} = \left[X_{i_T}(k_1), \ldots, X_{i_T}(k_p)\right]^T$ and $\mathbf{Y}_{i_R,pilot} = \left[Y_{i_R}(k_1), \ldots, Y_{i_R}(k_p)\right]^T$ are the pilot signals in the i_T-th transmit antenna and i_R-th receive antenna, diag($\mathbf{X}_{1,pilot}$) is a diagonal matrix with $\mathbf{X}_{1,pilot}$ as the main diagonal elements, $\mathbf{h}_{i_R} = \left[\mathbf{h}_{i_R,1}^T, \ldots, \mathbf{h}_{i_R,i_T}^T, \ldots, \mathbf{h}_{i_R,N_T}^T\right]^T$ with $\mathbf{h}_{i_R,i_T} = \left[h_{i_R,i_T}(0), \ldots, h_{i_R,i_T}(L-1)\right]^T$, and $\mathbf{n}_{iR,pilot}$ represents the frequency domain noise. Let $\mathbf{F}_L$ be a $K \times L$ matrix formed by the first L columns of a $K \times K$ DFT matrix $\mathbf{F}$, then $\mathbf{F}_{\text{pilot}}$ can be formed by taking only the rows of $\mathbf{F}_L$ associated with the K_P pilot sub-carriers.

By letting $\mathbf{A} = \mathbf{I}_{N_R} \otimes \left[diag\left(\mathbf{X}_{1,pilot}\right)\mathbf{F}_{pilot}, \ldots, diag\left(\mathbf{X}_{N_T,pilot}\right)\mathbf{F}_{pilot}\right]$, $\mathbf{Y}_{pilot} = \left[\mathbf{Y}_{1,pilot}^T, \ldots, \mathbf{Y}_{N_R,pilot}^T\right]^T$, $\mathbf{h} = \left[\mathbf{h}_1^T, \ldots, \mathbf{h}_{N_R}^T\right]^T$, and $\mathbf{n}_{pilot} = \left[\mathbf{n}_{1,pilot}^T, \ldots, \mathbf{n}_{N_R,pilot}^T\right]^T$, where $\otimes$ represents Kronecker product, we can get

$$\mathbf{Y}_{pilot} = \mathbf{A}\mathbf{h} + \mathbf{n}_{pilot}, \tag{2}$$

which can be solved by the conventional LS method, giving $\hat{\mathbf{h}} = \mathbf{A}^{\dagger}\mathbf{Y}_{pilot}$, where † represents the pseudoinverse.

Assuming the positions l_d (d = 0,1,...,D-1, and $l_0 < l_1 < \ldots < l_{D-1}$) of the MST are correctly estimated, Equation 1 can be rewritten as

$$\mathbf{Y}_{i_R,pilot} = \left[diag\left(\mathbf{X}_{1,pilot}\right)\mathbf{W}_{pilot}, \ldots, diag\left(\mathbf{X}_{N_T,pilot}\right)\mathbf{W}_{pilot}\right]\mathbf{z}_{i_R} + \mathbf{n}_{i_R,pilot}, \tag{3}$$

where $\mathbf{z}_{i_R} = \left[\mathbf{z}_{i_R,1}^T, \ldots, \mathbf{z}_{i_R,i_T}^T, \ldots, \mathbf{z}_{i_R,N_T}^T\right]^T$ with $\mathbf{z}_{i_R,i_T} = \left[\mathbf{z}_{i_R,i_T}(0), \ldots, \mathbf{z}_{i_R,i_T}(D-1)\right]^T$, D is the number of nonzero taps, and $\mathbf{W}_{pilot}$ can be formed by taking only the D columns of $\mathbf{F}_{pilot}$ associated with the nonzero tap positions l_d. Let $\tilde{\mathbf{A}} = \mathbf{I}_{N_R} \otimes \left[diag\left(\mathbf{X}_{1,pilot}\right)\mathbf{W}_{pilot}, \ldots, diag\left(\mathbf{X}_{N_T,pilot}\right)\mathbf{W}_{pilot}\right]$ and $\mathbf{z} = \left[\mathbf{z}_1^T, \ldots, \mathbf{z}_{N_R}^T\right]^T$. We can obtain

$$\mathbf{Y}_{pilot} = \tilde{\mathbf{A}}\mathbf{z} + \mathbf{n}_{pilot}. \tag{4}$$

When $\mathbf{n}_{pilot}$ is white noise and the positions l_d of MST are correctly estimated, we can obtain the estimate of the MST as $\hat{\mathbf{z}} = \tilde{\mathbf{A}}^{\dagger}\mathbf{Y}_{pilot}$. We can also obtain the Cramer-Rao bound of the sparse channel estimate $\hat{\mathbf{h}}$ through setting the elements of the positions l_d equal to $\hat{\mathbf{z}}$ and other elements equal to zero [4]. The above method to obtain the Cramer-Rao bound of $\hat{\mathbf{h}}$ is named as ideal-LS for comparison in this paper.

Note that the dimension of $\mathbf{Y}_{pilot}$ is proportional to the number of pilot subcarriers, and Equation 2 is an underdetermined problem when the dimension of $\mathbf{Y}_{pilot}$ is smaller than that of $\mathbf{h}$. Therefore, the sparse channel estimation in MIMO-OFDM systems can be viewed as solving an underdetermined linear inverse problem with sparsity constraint, i.e.,

$$\min_{\mathrm{h}} ||\mathbf{h}||_0 \quad \text{s.t.} \quad \mathbf{Y}_{pilot} = \mathbf{A}\mathbf{h} + \mathbf{n}_{pilot}, \tag{5}$$

where $||\cdot||_0$ represents the number of nonzero components named as l_0-norm.

3. Sparse channel estimation using l_2-Sl_0 reconstruction algorithm

The sparse signal reconstruction problem in CS is to estimate a sparse vector $\mathbf{x} \in \mathbb{C}^N$ from an observed vector $\mathbf{y} \in \mathbb{C}^M$ based on the linear model

$$\mathbf{y} = \mathbf{\Phi}\mathbf{x} + \mathbf{w}, \tag{6}$$

where $\mathbf{w} \in \mathbb{C}^M$ is unknown noise and $\mathbf{\Phi} \in \mathbb{C}^{M \times N}$ is a known measurement matrix, typically with $M \ll N$. This means that the signal $\mathbf{x}$ is 'sensed' by a reduced or 'compressed' number of measurements. Therefore, the signal reconstruction problem can be described as the following constrained minimization problem,

$$\min_{\mathbf{x}} ||\mathbf{x}||_0 \ \text{ s. t. } \ ||\mathbf{\Phi}\mathbf{x} - \mathbf{y}||_2 \leq \varepsilon, \tag{7}$$

where the bound $\varepsilon \geq 0$ is used to allow certain error tolerance. In general, ε is related to the variance of noise $\mathbf{w}$. Unfortunately, the problem in Equation 7 is a NP-hard combinatorial problem, whose computational complexity grows exponentially with the increase of the signal size and becomes prohibitive even for signals of moderate sizes. Consequently, several techniques have been proposed to tackle this difficult problem. One of the approaches is the convex relaxation, such as BP [20], which replaces $||\mathbf{x}||_0$ with $||\mathbf{x}||_1$ to make the problem easier to solve. Another approach, such as matching pursuit (MP [21]) or OMP, is much faster than BP but is a greedy algorithm and does not have provable reconstruction quality at the level of BP method [22]. Different from the above techniques, the smoothed l_0-norm approach [13] is to approximate the discontinuous l_0-norm by a suitable continuous one and then minimize it by an optimization algorithm dedicated to continuous functions. For example, the following continuous function

$$\begin{aligned} F_\sigma(\mathbf{x}) &= \sum_{i=1}^{N} f_\sigma(x_i) \ \text{ with } f_\sigma(x_i) \\ &= 1 - \exp\left(\frac{-x_i^2}{2\sigma^2}\right), \end{aligned} \tag{8}$$

where σ is a small value, has been proposed to approximate $||\mathbf{x}||_0$ in [13]. In other words, the minimum l_0-norm solution is then found by minimizing $F_\sigma(\mathbf{x})$ for a very small value of σ. The parameter σ determines how smooth the function $F_\sigma(\mathbf{x})$ would be and the accuracy of the approximation. Generally speaking, for larger values of σ, $F_\sigma(\mathbf{x})$ is smoother and contains less local minima, but the approximation to l_0-norm is worse. On the other hand, for smaller values of σ, a highly nonsmooth $F_\sigma(\mathbf{x})$ results, which gives a better approximation to l_0-norm but a difficult minimization problem. Consequently, the Sl_0 approach used a 'decreasing' sequence for σ.

The Sl_0 reconstruction algorithm is typically 2 to 3 orders of magnitude faster than BP, while resulting in the same or better accuracy [13]. However, in the presence of noise, the accuracy of Sl_0 algorithm needs to be improved. Therefore, in the next section, we will propose several improved Sl_0 reconstruction algorithms.

3.1 The l_2-Sl_0-BFGS reconstruction algorithm for channel estimation

Like l_1-regularized l_2 approach (l_2-l_1 [14,23,24]) and l_p-regularized l_2 algorithm [25], we use a parameter $\lambda > 0$ to balance the twin objectives of minimizing both error and sparsity, giving the following unconstrained optimization problem:

$$\min_{\mathbf{x}} F(\mathbf{x}) = \frac{1}{2}||\mathbf{\Phi}\mathbf{x} - \mathbf{y}||_2^2 + \lambda \sum_{i=1}^{N} \left[1 - \exp\left(\frac{-x_i^2}{2\sigma^2}\right)\right]. \tag{9}$$

The objective function in Equation 9 remains differentiable, and its gradient can be obtained as

$$\nabla F(\mathbf{x}) = \mathbf{\Phi}^T(\mathbf{\Phi}\mathbf{x} - \mathbf{y}) + \mathbf{g}, \tag{10}$$

where $\mathbf{g} = [g_1, g_2, \ldots, g_N]^T$ with g_i being given by

$$g_i = \lambda\left(x_i/\sigma^2\right)e^{-x_i^2/2\sigma^2}. \tag{11}$$

For a fixed value of σ, the problem in Equation 9 is now solved using a quasi-Newton algorithm where an approximation of the inverse of the Hessian is obtained using the Broyden-Fletcher-Goldfarb-Shanno (BFGS) update formula [26-28]. As such, the algorithm is referred to as the l_2-Sl_0-BFGS algorithm.

The quadratic (l_2-norm) error term $\frac{1}{2}||\mathbf{\Phi x}-\mathbf{y}||_2^2$ in Equation 9 is a convex function, but the convex region of the approximate l_0-norm term $F_\sigma(\mathbf{x}) = \sum_{i=1}^{N}[1-\exp\left(\frac{-x_i^2}{2\sigma^2}\right)]$ depends on parameter is σ. In general, the greater the value of σ, the larger the convex region is. To see this, we compute the gradient of $F_\sigma(\mathbf{x})$, denoted as $\mathbf{g}' = [g_1', g_2', ..., g_N']^T$, whose element is given by

$$g_i' = (x_i/\sigma^2)e^{-x_i^2/2\sigma^2}. \tag{12}$$

Also, the Hessian of $F_\sigma(\mathbf{x})$ is a diagonal matrix as given by

$$\begin{aligned}\nabla^2 F_\sigma(\mathbf{x}) &= diag\{h_{11}, h_{22}, ..., h_{NN}\} \text{ with } h_{ii} \\ &= \left(\frac{1}{\sigma^2}-\frac{x_i^2}{\sigma^4}\right)e^{\frac{-x_i^2}{2\sigma^2}}\end{aligned} \tag{13}$$

Therefore, $F_\sigma(\mathbf{x})$ is convex if and only if

$$|x_i| \le \sigma, \qquad 1 \le i \le N. \tag{14}$$

Since Equation 14 defines an N-dimensional hypercube whose volume is $(2\sigma)^N$, the size of the convex region in the x-space is proportional to σ. On the other hand, in order to better approximate the l_0-norm, σ must be sufficiently small. Consequently, to avoid getting trapped into local minima, we gradually decrease the value of σ, as in the Sl_0 approach. More specifically, for minimum $F(\mathbf{x})$ at σ_i, the initial point is $\mathbf{x}_*(\sigma_{i-1})$ obtained in the previous iteration, which is near the global optimal solution.

Since a broadband wireless channel response $\mathbf{h}$ usually consists of a few dominant propagation paths and Equation 2 has a similar form as Equation 6, the estimation of $\mathbf{h}$ can be viewed as a sparse signal reconstruction in compressed sensing. Thus, we refer to this kind of sparse channel estimation method as CCS. Using Equations 2, 6, and 9, we can obtain the objective function of CCS based on the l_2-Sl_0 reconstruction algorithm,

$$\begin{aligned}\min_{\mathbf{h}} F(\mathbf{h}) = &\frac{1}{2}||\mathbf{Ah}-\mathbf{Y}_{pilot}||_2^2 \\ &+\lambda\sum_{i=1}^{N}\left[1-\exp\left(\frac{-h_i^2}{2\sigma^2}\right)\right].\end{aligned} \tag{15}$$

From the above analysis, the proposed CCS using the l_2-Sl_0-BFGS algorithm can be implemented by the pseudo-code in Algorithm 1.

Algorithm 1 CCS using the l_2-Sl_0-BFGS algorithm

Input: measurement matrix $\mathbf{A}$, measurement value $\mathbf{Y}_{pilot}$, regularization parameter λ, descent factor r, target factor r_J, step size δ_r, and target parameter σ_J.

1) Initialization: $\hat{\mathbf{h}}_0 = \mathbf{A}^\dagger\mathbf{Y}_{pilot}$, $\sigma = \max_i[|\hat{h}_0(i)|]$, and $k = 0$.
2) If $\sigma > \sigma_J$,
   a) Using $\hat{\mathbf{h}}_k$ as an initial point and applying the BFGS algorithm to solve the problem in Equation 15, we obtain solution $\hat{\mathbf{h}}_{k+1}$.
   b) Let $r = r + \delta_r$. If $r \ge r_J$, set $r = r_J$.
   c) Let $\sigma = r\sigma$, $k = k + 1$, and go to step 2).

Else, output estimate $\hat{\mathbf{h}}_k$.

Note that in Algorithm 1, the values of δ_r and r_J are chosen such that $0 < \delta_r < 0.1$ and $0.5 < r_J < 1$. The method of l_2-Sl_0 uses a variable factor $r_i = r_{i-1} + \delta_r$ to control the decrease of the parameter σ. Our idea is to use an 'increasing' step size corresponding to the decreasing values of σ.

3.2 The l_2-Sl_0-CG reconstruction algorithm for channel estimation

The Hessian matrix of the objective function $F(\mathbf{x})$ in Equation 9 can be computed as

$$\nabla^2 F(\mathbf{x}) = \mathbf{\Phi}^T\mathbf{\Phi} + \lambda\nabla^2 F_\sigma(\mathbf{x}), \tag{16}$$

where $\nabla^2 F_\sigma(\mathbf{x})$ is computed using Equation 13. Since the gradient and Hessian matrix of $F(\mathbf{x})$ can be efficiently evaluated using the closed-form formula in Equations 10 and 16, it is convenient to apply the conjugate gradient method to solve the l_2-Sl_0 optimization problem. The algorithm is thus referred to as the l_2-Sl_0-CG algorithm.

In the k-th iteration of the conjugate gradient technique, $\mathbf{x}_k$ is updated as

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k\mathbf{d}_k, \quad (k = 0, 1, ..., L-1). \tag{17}$$

The conjugate direction $\mathbf{d}_k$ is computed as

$$\mathbf{d}_k = \begin{cases} -\mathbf{g}_0, & k = 0 \\ -\mathbf{g}_k + \beta_{k-1}\mathbf{d}_{k-1}, & k = 1, 2, ..., L-1 \end{cases} \text{ with}$$

$$\beta_{k-1} = \frac{\mathbf{g}_k^T\mathbf{g}_k}{\mathbf{g}_{k-1}^T\mathbf{g}_{k-1}}, \tag{18}$$

and the k-th step size α_k is computed using

$$\alpha_k = \frac{\mathbf{g}_k^T\mathbf{g}_k}{\mathbf{d}_k^T\mathbf{H}_k\mathbf{d}_k}, \tag{19}$$

Where $\mathbf{g}_k$ is the gradient vector computed using Equation 10 and $\mathbf{H}_k$ is the Hessian matrix obtained using Equation 16 at $\mathbf{x} = \mathbf{x}_k$, respectively. The proposed CCS using l_2-Sl_0-CG algorithm can be implemented by the pseudo-code in Algorithm 2.

Algorithm 2 CCS using l_2-Sl_0-CG algorithm

Input: measurement matrix $\mathbf{A}$, measurement value $\mathbf{Y}_{pilot}$, regularization parameter λ, descent factor r, target factor r_J, step size δ_r, target parameter σ_J, and iteration number L.

1) Initialization: $\hat{\mathbf{h}} = \mathbf{A}^{\dagger}\mathbf{Y}_{pilot}$, $\sigma = \max_i[|\hat{h}(i)|]$, and $\hat{\mathbf{h}}_0 = 0$.

2) For $k = 0,\ldots, L-1$

a) Compute $\mathbf{g}_k$ and $\mathbf{H}_k$ according to Equations 10 and 16 at $\hat{\mathbf{h}}_k$, respectively;

b) If $k = 0$, set $\mathbf{d}_k = -\mathbf{g}_k$; else, set $\mathbf{d}_k = -\mathbf{g}_k + \frac{\mathbf{g}_k^T\mathbf{g}_k}{\mathbf{g}_{k-1}^T\mathbf{g}_{k-1}}\mathbf{d}_{k-1}$;

c) Set $\alpha_k = \frac{\mathbf{g}_k^T\mathbf{g}_k}{\mathbf{d}_k^T\mathbf{H}_k\mathbf{d}_k}$ and compute $\hat{\mathbf{h}}_{k+1} = \hat{\mathbf{h}}_k + \alpha_k\mathbf{d}_k$.

3) Set $r = r + \delta_r$. If $r \geq r_J$, set $r = r_J$.

4) Set $\sigma = r\sigma$. If $\sigma > \sigma_J$, set $\hat{\mathbf{h}}_0 = \hat{\mathbf{h}}_L$ and go to step 2); else, output the estimate $\hat{\mathbf{h}}_L$.

3.3 Signal reconstruction via optimization in null and complement spaces of Φ

Let $\mathbf{\Phi} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$ be the singular value decomposition (SVD) of $\mathbf{\Phi}$ where $\mathbf{U}_{M\times M}$ and $\mathbf{V}_{N\times N}$ are unitary matrices, and $\mathbf{\Sigma} = [\mathbf{S}, \mathbf{0}]_{M\times N}$ with $\mathbf{S} = \mathrm{diag}(s_1, \ldots, s_M)$ being a diagonal matrix composed by the singular values of $\mathbf{\Phi}$. Let $\mathbf{V} = [\mathbf{V}_r, \mathbf{V}_n]$, where the columns of $\mathbf{V}_n$ span the null space of $\mathbf{\Phi}$ and the columns of $\mathbf{V}_r$ span the orthogonal complement of the null space. Using $\mathbf{V}_n$ and $\mathbf{V}_r$, a signal $\mathbf{x}$ of length N can be expressed as

$$\mathbf{x} = \mathbf{V}_r\boldsymbol{\alpha} + \mathbf{V}_n\boldsymbol{\beta}, \tag{20}$$

Where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are vectors of length M and $N-M$, respectively. Applying the SVD of $\mathbf{\Phi}$, the l_2-norm term in Equation 9 can be simplified as [25]

$$\frac{1}{2}||\mathbf{\Phi}\mathbf{x}-\mathbf{y}||_2^2 = \frac{1}{2}||\mathbf{\Sigma}\boldsymbol{\alpha}-\tilde{\mathbf{y}}||_2^2 = \frac{1}{2}\sum_{i=1}^{M}(s_i\alpha_i-\tilde{y}_i)^2, \tag{21}$$

Where S_i is the i-th singular value of $\mathbf{\Phi}$, α_i and $\tilde{y}_i$ are the i-th component of $\boldsymbol{\alpha}$ and $\tilde{\mathbf{y}} = \mathbf{U}^T\mathbf{y}$, respectively. Using Equations 20 and 21, the optimization problem in Equation 9 can be recast as

$$\min_{\boldsymbol{\alpha},\boldsymbol{\beta}} \quad F(\boldsymbol{\alpha},\boldsymbol{\beta}) = \frac{1}{2}\sum_{i=1}^{M}(s_i\alpha_i-\tilde{y}_i)^2 + \lambda\sum_{j=1}^{N}\left[1-\exp\left(\frac{-\left(V_{r,j}\boldsymbol{\alpha}+V_{n,j}\boldsymbol{\beta}\right)^2}{2\sigma^2}\right)\right], \tag{22}$$

Where $V_{r,j}$ and $V_{n,j}$ are the j-th row of $\mathbf{V}_r$ and that of $\mathbf{V}_n$, respectively.

An iterative algorithm to solve the optimization problem in Equation 22 is proposed as follows. In the k-th iteration of the optimization process, signal $\mathbf{x}^{(k)}$ is updated as

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mu^{(k)}\mathbf{d}^{(k)}, \tag{23}$$

where

$$\mathbf{x}^{(k)} = \mathbf{V}_r\boldsymbol{\alpha}^{(k)} + \mathbf{V}_n\boldsymbol{\beta}^{(k)}, \tag{24}$$

$$\mathbf{d}^{(k)} = \mathbf{V}_r\mathbf{d}_r^{(k)} + \mathbf{V}_n\mathbf{d}_n^{(k)}$$

and the step size $\mu^{(k)} > 0$ is determined by the inexact line search method of Roger Fletcher [26]. Assuming that the updating vectors $\mathbf{d}_r^{(k)}$ and $\mathbf{d}_n^{(k)}$ are written as

$$\mathbf{d}_r^{(k)} = \left[d_{r,1}^{(k)}, d_{r,2}^{(k)}, \ldots, d_{r,M}^{(k)}\right]^T, \tag{25}$$

$$\mathbf{d}_n^{(k)} = \left[d_{n,1}^{(k)}, d_{n,2}^{(k)}, \ldots, d_{n,N-M}^{(k)}\right]^T$$

which are determined by minimizing $F(\boldsymbol{\alpha},\boldsymbol{\beta})$ along each of the directions defined by the column vectors of $[\mathbf{V}_r,\mathbf{V}_n]$. Therefore, $\mathbf{d}_r^{(k)}$ and $\mathbf{d}_n^{(k)}$ become descent directions of $F(\boldsymbol{\alpha},\boldsymbol{\beta})$, and in the case of real $\mathbf{\Phi}$ and $\mathbf{x}$, $d_{r,i}^{(k)}$ can be calculated via iteration as

$$\left(d_{r,i}^{(k)}\right)^{(p)} = \frac{s_i\tilde{y}_i - s_i^2\alpha_i - \frac{\lambda}{\sigma^2}\sum_{j=1}^{N}\left[\exp\left(\frac{-\left(x_j + v_r^{(j,i)}\left(d_{r,i}^{(k)}\right)^{(p-1)}\right)^2}{2\sigma^2}\right)x_j v_r^{(j,i)}\right]}{s_i^2 + \frac{\lambda}{\sigma^2}\sum_{j=1}^{N}\left[\exp\left(\frac{-\left(x_j + v_r^{(j,i)}\left(d_{r,i}^{(k)}\right)^{(p-1)}\right)^2}{2\sigma^2}\right)\left(v_r^{(j,i)}\right)^2\right]}, \quad (1\leq i\leq M), \tag{26}$$

Where x_j is the j-th component of vector $\mathbf{x}^{(k)}$, $v_r^{(j,i)}$ is the (j,i)-th component of matrix $\mathbf{V}_r$, α_i is the i-th component of vector $\boldsymbol{\alpha}^{(k)}$, and $\left(d_{r,i}^{(k)}\right)^{(p)}$ is the p-th iteration value of $d_{r,i}^{(k)}$ with the initialization value $\left(d_{r,i}^{(k)}\right)^{(0)} = 0$.

Similarly, $d_{n,i}^{(k)}$ in Equation 25 is given by

$$\left(d_{n,i}^{(k)}\right)^{(q)} = \frac{-\sum_{j=1}^{N}\left[\exp\left(\frac{-\left(x_j + v_n^{(j,i)}\left(d_{n,i}^{(k)}\right)^{(q-1)}\right)^2}{2\sigma^2}\right) x_j v_n^{(j,i)}\right]}{\sum_{j=1}^{N}\left[\exp\left(\frac{-\left(x_j + v_n^{(j,i)}\left(d_{n,i}^{(k)}\right)^{(q-1)}\right)^2}{2\sigma^2}\right)\left(v_n^{(j,i)}\right)^2\right]}, \quad (1 \le i \le N-M), \tag{27}$$

where $v_n^{(j,i)}$ is the (j,i)-th component of matrix $\mathbf{V}_n$, and $\left(d_{n,i}^{(k)}\right)^{(q)}$ is the q-th iteration value of $d_{n,i}^{(k)}$ with the initialization value $\left(d_{n,i}^{(k)}\right)^{(0)} = 0$. The derivation of Equations 26 and 27 is given in the Appendix. In addition, the computation of $d_{r,i}^{(k)}$ using Equation 26 requires vector $\boldsymbol{\alpha}^{(k)}$ to be computed first as

$$\mathbf{V}_r^T \mathbf{x}^{(k)} = \mathbf{V}_r^T\left(\mathbf{V}_r \boldsymbol{\alpha}^{(k)} + \mathbf{V}_n \boldsymbol{\beta}^{(k)}\right) = \boldsymbol{\alpha}^{(k)}. \tag{28}$$

The reconstruction algorithm via optimization in null and complement spaces of measurement matrix $\boldsymbol{\Phi}$ is referred to hereafter as the l_2-Sl_0-NC algorithm, which can be implemented by the pseudo-code in Algorithm 3.

Algorithm 3 The l_2-Sl_0-NC reconstruction algorithm

Input: measurement matrix $\boldsymbol{\Phi}$, measurement value $\mathbf{y}$, regularization parameter λ, descent factor r, target factor r_J, step size δ_r, and target parameter σ_J.

1) Initialization: $\hat{\mathbf{x}}_0 = \boldsymbol{\Phi}^{\dagger}\mathbf{y}$, $\sigma = \max_i[|\hat{x}_0(i)|]$, $\hat{\mathbf{x}}^{(0)} = 0$, $\delta_1 = 0.001$, and $\delta_2 = 0.01$.

2) Perform SVD of $\boldsymbol{\Phi}$ to get $\boldsymbol{\Phi} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$, and construct $\mathbf{V}_r$ and $\mathbf{V}_n$ using the first M columns and last N-M columns of $\mathbf{V}$, respectively. Set $\tilde{\mathbf{y}} = \mathbf{U}^T\mathbf{y}$ and construct vector $\mathbf{s}$ using the singular values of $\boldsymbol{\Phi}$.

3) If $\sigma > \sigma_J$,

a) Set $\Delta x = \Delta d_r = \Delta d_n = 1$, and $k = 0$.

b) While $\Delta x > \delta_1$, do the following:

① Set $(\mathbf{d}_r^{(k)})^{(0)} = (\mathbf{d}_n^{(k)})^{(0)} = \mathbf{0}$, $p = q = \mathbf{0}$, and compute $\boldsymbol{\alpha}^{(k)}$ according to Equation 28.

② While $\Delta d_r > \delta_2$,

Set $p = p + 1$, and compute $(d_{r,i}^{(k)})^{(p)}$ according to Equation 26 at point $\hat{\mathbf{x}}^{(k)}$, $(i = 1, \ldots, M)$.

Set $\Delta d_r = \sum_{i=1}^{M} |(d_{r,i}^{(k)})^{(p)} - (d_{r,i}^{(k)})^{(p-1)}|$.

③ While $\Delta d_n > \delta_2$,

Set $q = q + 1$ and compute $(d_{n,i}^{(k)})^{(q)}$ according to Equation 27 at point $\hat{\mathbf{x}}^{(k)}$, $(i = 1, \ldots, N-M)$.

Set $\Delta d_n = \sum_{i=1}^{N-M} |(d_{n,i}^{(k)})^{(q)} - (d_{n,i}^{(k)})^{(q-1)}|$.

④ Set $\mathbf{d}_r^{(k)} = (\mathbf{d}_r^{(k)})^{(p)}$, $\mathbf{d}_n^{(k)} = (\mathbf{d}_n^{(k)})^{(q)}$, $\mathbf{d}_k^{(k)} = \mathbf{V}_r\mathbf{d}_r^{(k)} + \mathbf{V}_n\mathbf{d}_n^{(k)}$, and compute the step size $\mu^{(k)}$.

⑤ Compute $\hat{\mathbf{x}}^{(k+1)} = \hat{\mathbf{x}}^{(k)} + \mu^{(k)}\mathbf{d}_k^{(k)}$, $\Delta x = \sum_{i=1}^{N} |\mu^{(k)} d_k^{(k)}(i)|$, and set $k = k + 1$.

c) Set $r = r + \delta_r$. If $r \ge r_J$, set $r = r_J$.

d) Set $\sigma = r\sigma$, $\hat{\mathbf{x}}^{(0)} = \hat{\mathbf{x}}^{(k)}$, and go to step 3).

Else, output the estimate $\hat{\mathbf{x}}^{(k)}$.

4. Simulation results

In this section, the reconstruction performance of the proposed approach (l_2-Sl_0) is evaluated by computer simulations. The spectral efficiency of the CCS using l_2-Sl_0 algorithm is also discussed. More specifically, the l_2-Sl_0 algorithm includes l_2-Sl_0-BFGS, l_2-Sl_0-CG, and l_2-Sl_0-NC in the scenario where $\mathbf{y}$, $\mathbf{\Phi}$,$\mathbf{x}$ are real-valued. While in complex-valued scenarios, l_2-Sl_0 only means l_2-Sl_0-BFGS and l_2-Sl_0-CG. Note that Equations 26 and 27 are obtained only in the case where $\mathbf{y}$, $\mathbf{\Phi}$, $\mathbf{x}$ are real-valued. Namely, the l_2-Sl_0-NC is not suitable to reconstruct complex signals. In all the experiments, the initial value of r is set to 0.5 for both Sl_0 and l_2-Sl_0 algorithms. The values of δ_r and r_J required by the l_2-Sl_0 algorithm are chosen as 0.05 and 0.7, respectively.

In experiments 1 and 2, the signal length and the number of measurements are set to $N = 1{,}000$ and $M = 400$, respectively. A K-sparse source $\mathbf{x}$ was artificially created as follows: (1) set $\mathbf{x}$ to a zero vector of length N, (2) generate a vector $\mathbf{z}$ of length K assuming that each element z_i is a random value drawn from the normal distribution $N(0,1)$ in the real-valued scenario or from $N(0,1/2)+jN(0,1/2)$ in the complex-valued scenario, and (3) randomly select K components of $\mathbf{x}$ and set them to $\mathbf{z}$. Each element of the measurement matrix $\mathbf{\Phi}$ is randomly generated using the normal distribution $N(0,1)$ or $N(0,1)+jN(0,1)$, and each row is normalized to unity. Then, the mixtures are generated using the noisy model $\mathbf{y}=\mathbf{\Phi x}+\mathbf{w}$, where $\mathbf{w}$ is an additive white Gaussian noise with covariance matrix $\sigma_w\mathbf{I}_M$ ($\mathbf{I}_M$ stands for the $M \times M$ identity matrix). To evaluate the estimation accuracy, the signal-to-noise ratio (SNR) defined as $20\log(||\mathbf{x}||_2/||\mathbf{x}-\hat{\mathbf{x}}||_2)$ is used, where $\mathbf{x}$ and $\hat{\mathbf{x}}$ denote the true value and its estimate, respectively.

In experiment 1, we compare the reconstruction performance of l_2-Sl_0 with that of Sl_0. Figures 1 and 2 show the reconstruction SNR at different powers of noise σ_w in real and complex signal scenarios, respectively. For each value of σ_w, the reconstruction SNR is averaged over 100 runs. It is seen that l_2-Sl_0 produces a better SNR than Sl_0, which shows the robustness of l_2-Sl_0 against noise. The objective function of l_2-Sl_0 algorithm in Equation 15 comprises the quadratic error term $\frac{1}{2}||\mathbf{Ah}-\mathbf{Y}_{\text{pilot}}||_2^2$ which permits a small perturbation. Therefore, the l_2-Sl_0 algorithm has a larger capability to reconstruct sparse signal in the presence of noise than Sl_0. For smaller values of σ, $F_\sigma(\mathbf{x})$ contains more local minima. Therefore, the decrease of σ should not be too quick in the Sl_0 and l_2-Sl_0 algorithms. Moreover, unlike Sl_0 using a fixed step size to control the decrease of the parameter σ, l_2-Sl_0 uses a variable one, and the step size δ_r slightly increasing with the reduction of σ may also help the l_2-Sl_0 to improve its estimation accuracy.

In experiment 2, the l_2-Sl_0 algorithms are tested using $N = 1{,}000$, $M = 400$, and various K sparse signals with $\sigma_\omega = 0.01$, to examine the algorithms' performance for signals of different sparsity levels. The results obtained are plotted in Figures 3 and 4 with $\mathbf{y}$, $\mathbf{\Phi}$, $\mathbf{x}$ being real and complex values, respectively. It is observed that the performance of the l_2-Sl_0 algorithm is better than the Sl_0 algorithm in most cases. In real-valued scenario, the l_2-Sl_0-BFGS, l_2-Sl_0-CG, and l_2-Sl_0-NC are comparable for K smaller than 130, but l_2-Sl_0-BFGS performs better for K between 130 and 210. In addition, when K is smaller than 90, the final SNR of the Sl_0 algorithm increases with the rise of sparsity K. This is because the initial estimate $\hat{\mathbf{x}}_0$ is set to the minimum l_2-norm solution of $\mathbf{y} = \mathbf{\Phi x} + \mathbf{w}$, which has few zero elements and is far away from the actual signal with many zero elements, and $\hat{\mathbf{x}}_0$ is gradually close to the actual signal with the rise of sparsity K. However, this phenomena is not obvious in l_2-Sl_0 algorithm, since the initial estimate $\hat{\mathbf{x}}_0$ is set to zeros in l_2-Sl_0-CG and l_2-Sl_0-NC, which is near the actual solution for a small value of K. Because of $\hat{\mathbf{x}}_0$ being set to zeros and the thresholds of δ_1 and δ_2 being not small enough for the value sparsity K above 230, the l_2-Sl_0-NC performs the worst among the algorithms tested.

Next, we investigate the accuracy and the spectral efficiency of CCS method using l_2-Sl_0. We consider a MIMO-OFDM system with two transmit and two receive antennas ($N_T = N_R = 2$). The number of subcarriers is 512, and the QPSK modulation is used. The length of cyclic prefix is 20, which equals the length of wireless channel impulse response. In experiment 3, a Rayleigh channel modeled by a 4-tap MIMO-FIR filter is assumed, in which each tap corresponds to a 2×2 random matrix whose elements are i.i.d. complex Gaussian variables with zero mean and unit variance, and the position l_d of MSTs is {2, 6, 13, 19}. The estimation performance is evaluated in terms of the bit error rate (BER) and mean square error (MSE) given by $MSE(\Delta h) = \frac{\sum_{i=1}^{M}||\hat{\mathbf{h}}_i-\mathbf{h}_i||_2^2}{\sum_{i=1}^{M}||\mathbf{h}_i||_2^2}$, where M represents the number of simulations and $\mathbf{h}_i$ and $\hat{\mathbf{h}}_i$ represent the actual and the estimated channels from the i-th simulation, respectively.

In experiment 3, we investigate the performance and required computational time of the CCS using l_2-Sl_0-BFGS, l_2-Sl_0-CG, and Sl_0 reconstruction algorithms with 30 pilot signals in each transmit antenna. The simulation consists of 2,000 Monte Carlo runs. Moreover, their performance is compared with those of the CCS using OMP and FISTA. OMP is the most popular one in the type of greedy reconstruction algorithm, and FISTA is the most fast one in the type of l_2-l_1 reconstruction algorithm. Figures 5 and 6 show the MSE and BER plots

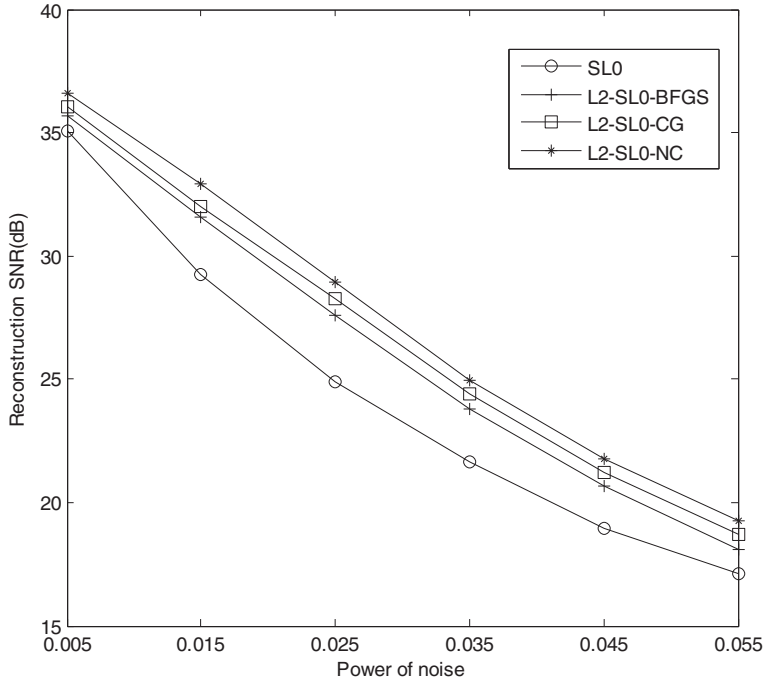

Figure 1 Reconstruction SNR versus noise power in real-valued scenario.

resulting from the above five CCS methods and the conventional LS method, respectively. As can be seen, the CCS method using l_2-Sl_0-CG only needs 30 pilot signals to obtain the approximate performance of the LS method using 40 pilot signals which implies that the CCS using l_2-Sl_0-CG can save nearly 25% pilot signals. This merit of CCS is due to the prior sparse information of the wireless channel utilized and the efficient reconstruction of sparse signals from a very limited number of measurements allowed by CS. In addition, the CCS applying l_2-Sl_0-CG or l_2-Sl_0-BFGS outperforms the CCS using Sl_0 more obviously than that in experiment 1,

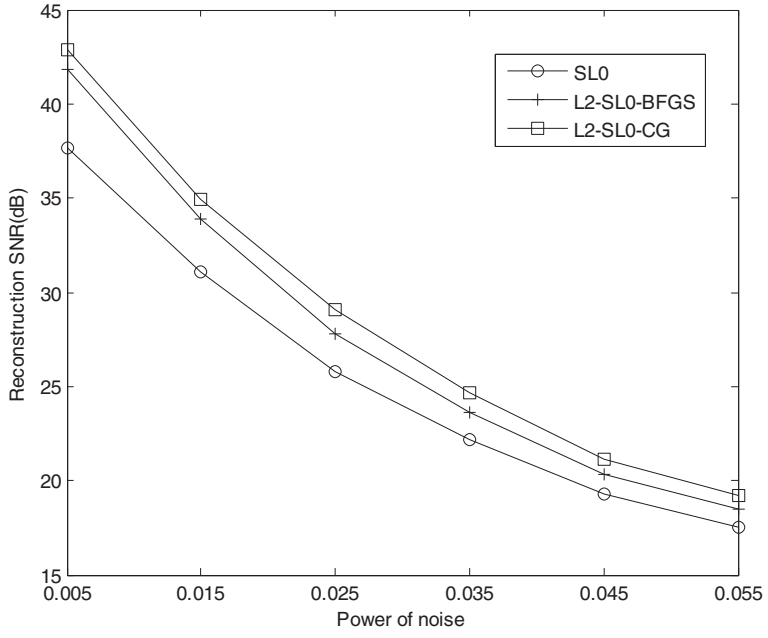

Figure 2 Reconstruction SNR versus noise power in complex-valued scenario.

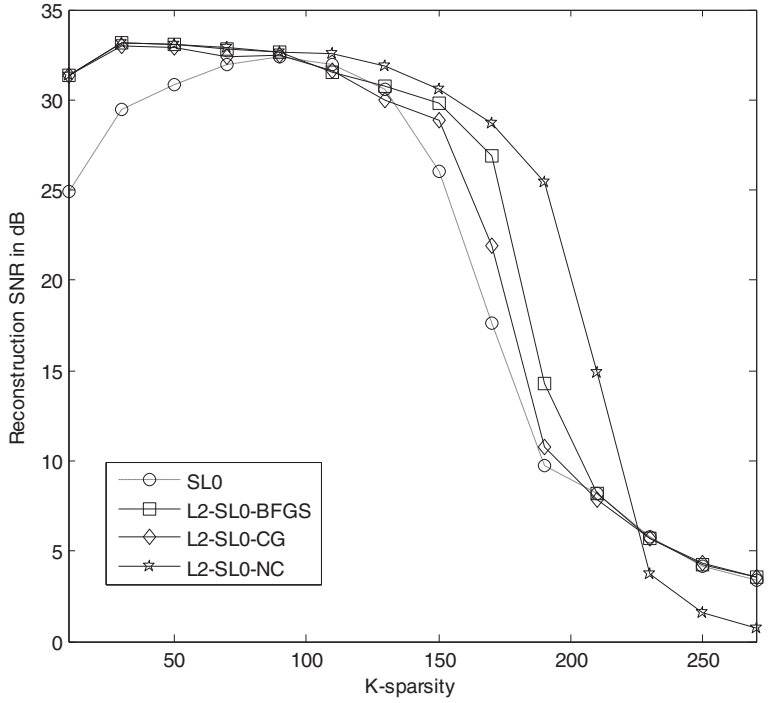

Figure 3 Reconstruction SNR at various sparsity levels over 100 runs in real-valued scenario.

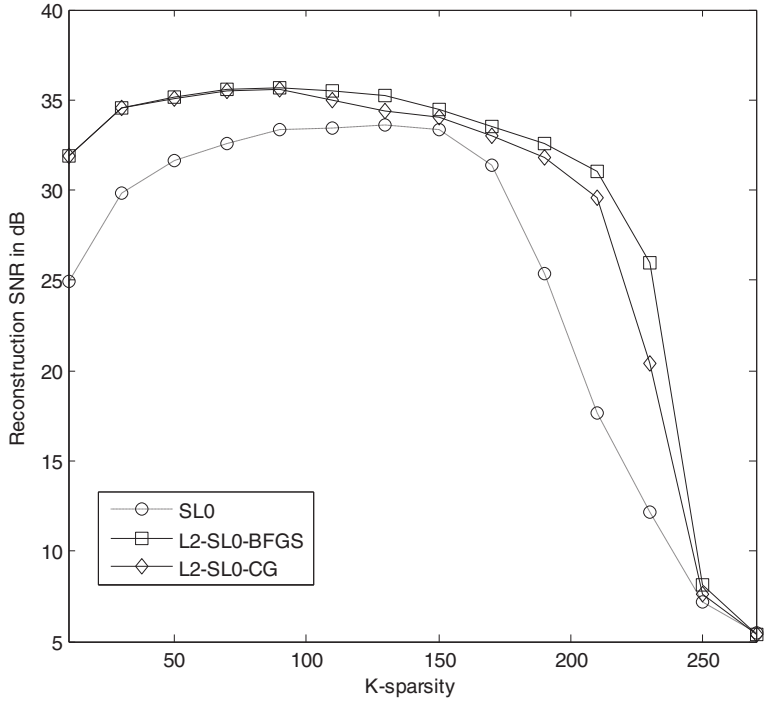

Figure 4 Reconstruction SNR at various sparsity levels over 100 runs in complex-valued scenario.

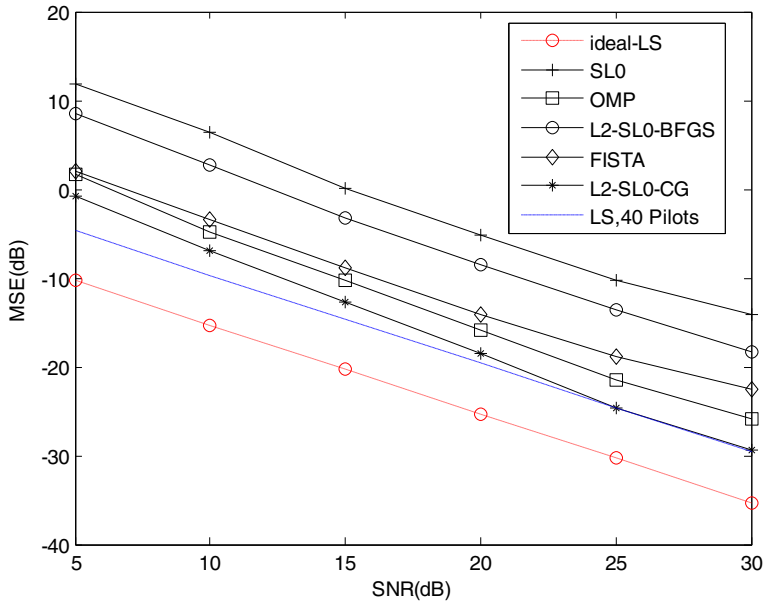

Figure 5 MSE versus SNR.

which shows that the l_2-Sl_0 has a larger capability to reconstruct sparse signal than Sl_0 in the case when each row of measurement matrix is not normalized to unity. Since the l_2-Sl_0 algorithm is halted after a fixed number of iterations, furthermore, the fixed number does not depend on the sparsity of the signal directly; it is convenient to set the number in practical applications.

We use the CPU time as a measure of complexity. The simulations are performed in MATLAB R2009b environment using an Intel Core i3, 2.53-GHz processor with 2 GB of memory, and under Microsoft Windows XP operating system. The results shown in Figure 7 indicate that the CCS using l_2-Sl_0 requires more computational time than that using other algorithms tested. The l_2-Sl_0 algorithm needs an iterative process to find the optimal solution at each value of σ; therefore, the running time of l_2-Sl_0 is longer than that of others tested. However, at the cost of slightly more computational time, the CCS using l_2-Sl_0-CG yields slightly better performance than the CCS using OMP or FISTA, and the threshold value

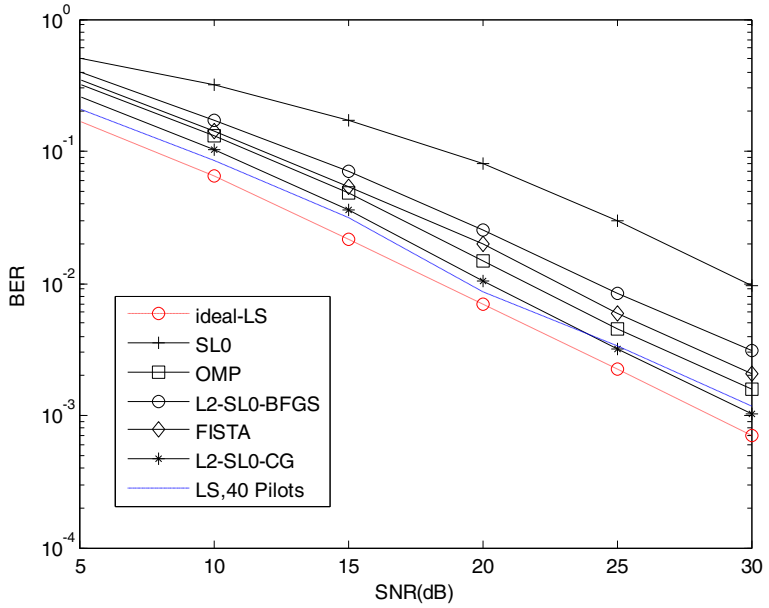

Figure 6 BER versus SNR.

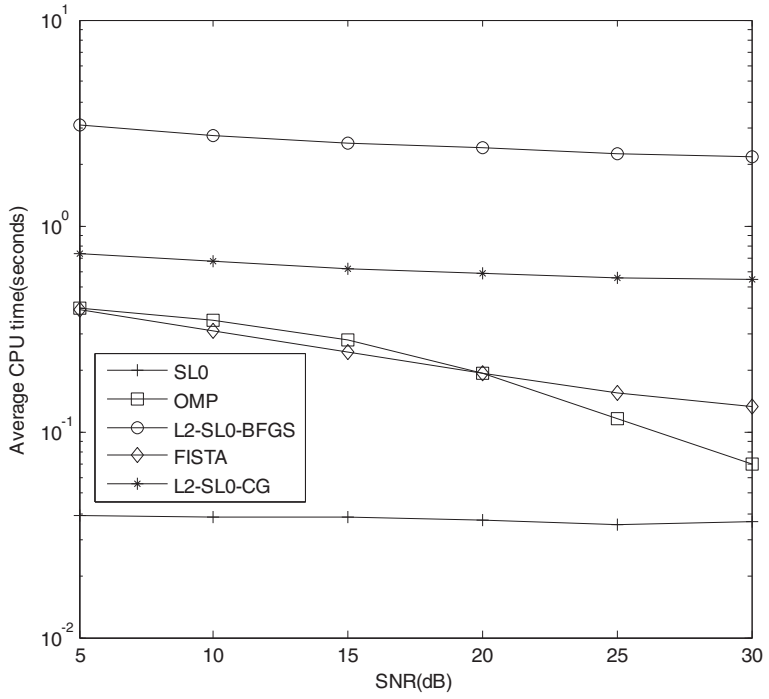

Figure 7 Computation time of each method.

for termination iteration in the l_2-Sl_0 algorithm is easier to be set. More specifically, it is shown in Algorithm 2 that l_2-Sl_0-CG applies a constant value L to stop the iteration, and the constant value is independent of the sparsity of signal and the power of noise. However, the valid threshold values for termination iteration in the OMP and FISTA algorithms always depend on the power of noise or the sparsity of signal, which are both quite difficult to estimate beforehand in practical applications.

In experiment 4, we investigate the BER of the CCS using 30 pilot signals in each transmit antenna under different channel sparsities, namely for different numbers of MSTs. Moreover, the position l_d of MST is randomly selected in each Monte Carlo simulation. Figure 8 shows the BER plots of CCS using l_2-Sl_0-CG and l_2-Sl_0-BFGS algorithms. The figure shows that a better BER performance can be expected in general for less number of MSTs. In addition, when the length of channel response is 20, the CCS using l_2-Sl_0-CG and that using l_2-Sl_0-BFGS are found to yield acceptable BERs for up to 8 and 4 MSTs, respectively.

5. Conclusion

In this paper, a new approach for sparse channel estimation of MIMO-OFDM systems based on compressed sensing has been presented. The new approach uses a smoothed l_0-norm-regularized least squares (l_2-Sl_0) objective function and solves the optimization problem by three reconstruction algorithms: quasi-Newton, conjugate gradient (CG), and optimization in the null and complement spaces of measurement matrix (ONCS). The better reconstruction accuracy of the l_2-Sl_0 as compared with the Sl_0 algorithm and the higher spectrum efficiency of the CCS using l_2-Sl_0-CG or l_2-Sl_0-BFGS as compared with the conventional LS method have been shown by computer simulations.

Appendix

Derivation of Equations 26 and 27

Suppose that $\mathbf{e}_i$ is the i-th column of an $M \times M$ identity matrix, and the vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ in Equation 22 are fixed. At point $\boldsymbol{\alpha}$, a line search along direction $\mathbf{e}_i$ is carried out by solving the one-dimensional optimization problem

$$\begin{aligned}\min_{d_{r,i}} \quad & F(\boldsymbol{\alpha}+d_{r,i}\mathbf{e}_i,\boldsymbol{\beta}) = \frac{1}{2}\left|\left|\mathbf{s}(\boldsymbol{\alpha}+d_{r,i}\mathbf{e}_i)-\tilde{\mathbf{y}}\right|\right|_2^2 \\ & +\lambda\sum_{j=1}^{N}\left[1-\exp\left(-\frac{\left(x_j+d_{r,i}v_r^{(j,i)}\right)^2}{2\sigma^2}\right)\right] \\ = & \frac{1}{2}\left(s_i(\alpha_i+d_{r,i})-\tilde{y}_i\right)^2+\frac{1}{2}\sum_{\substack{k=1\\k\neq i}}^{M}\left(s_k\alpha_k-\tilde{y}_k\right)^2 \\ & +\lambda\sum_{j=1}^{N}\left[1-\exp\left(-\frac{\left(x_j+d_{r,i}v_r^{(j,i)}\right)^2}{2\sigma^2}\right)\right],\end{aligned} \tag{29}$$

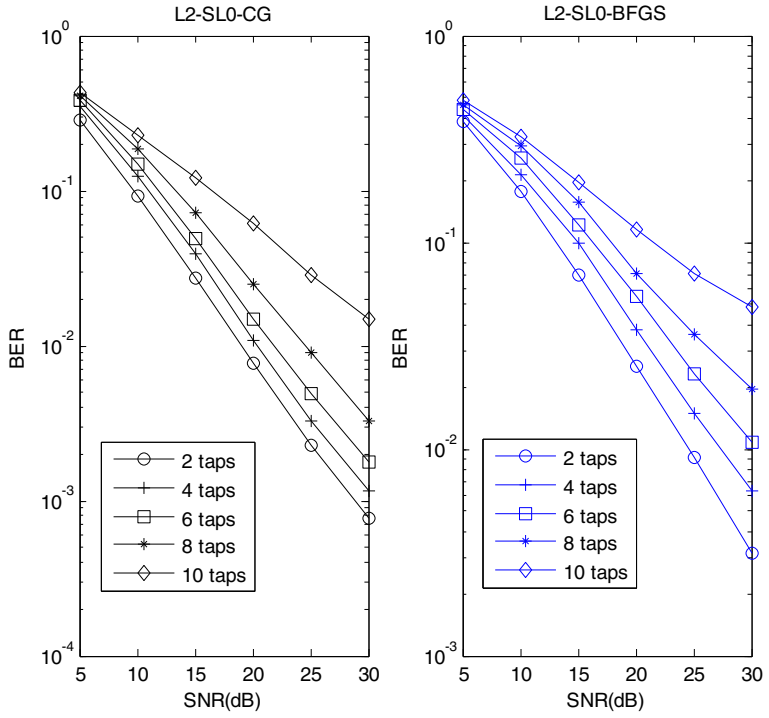

Figure 8 BER of CCS using l_2-Sl_0 at different channel sparsities.

Where x_j is the j-th component of vector $\mathbf{x}$, and $v_r^{(j,i)}$ is the (j,i)-th component of matrix $\mathbf{V}_r$. By equating the derivative $\partial F(\boldsymbol{\alpha}+d_{r,i}\mathbf{e}_i,\boldsymbol{\beta})/\partial d_{r,i}$ to zero, for real $\boldsymbol{\Phi}$ and $\mathbf{x}$, we can obtain

$$d_{r,i}=\frac{s_i\tilde{y}_i-s_i^2\alpha_i-\frac{\lambda}{\sigma^2}\sum_{j=1}^{N}\left[\exp\left(\frac{-\left(x_j+v_r^{(j,i)}d_{r,i}\right)^2}{2\sigma^2}\right)x_jv_r^{(j,i)}\right]}{s_i^2+\frac{\lambda}{\sigma^2}\sum_{j=1}^{N}\left[\exp\left(\frac{-\left(x_j+v_r^{(j,i)}d_{r,i}\right)^2}{2\sigma^2}\right)\left(v_r^{(j,i)}\right)^2\right]}. \tag{30}$$

Note that $d_{r,i}$ can be solved via iterations with the initial value of $d_{r,i}$ being set to zero in the right side of (30). In a similar manner, $d_{n,i}$ can be obtained as

$$d_{n,i}=\frac{-\sum_{j=1}^{N}\left[\exp\left(\frac{-\left(x_j+v_n^{(j,i)}d_{n,i}\right)^2}{2\sigma^2}\right)x_jv_n^{(j,i)}\right]}{\sum_{j=1}^{N}\left[\exp\left(\frac{-\left(x_j+v_n^{(j,i)}d_{n,i}\right)^2}{2\sigma^2}\right)\left(v_n^{(j,i)}\right)^2\right]}, \tag{31}$$

where $v_n^{(j,i)}$ is the (j,i)-th component of matrix $\mathbf{V}_n$.

Competing interests
The authors declare that they have no competing interests.

Acknowledgements
We express our thanks to the anonymous reviewers for their valuable comments to improve the quality and the presentation of this paper. This work is supported by the National Natural Science Foundation of China under grant nos. 61372122 and 61302104 and the Basic Research Program of Jiangsu Province under grant no. BK2011756.

Author details
[1]Institute of Signal Processing and Transmission, Nanjing University of Posts and Telecommunications, Nanjing 210003, China. [2]College of Physics and Electronic Information, Anhui Normal University, Wuhu 241000, China. [3]Department of Electrical and Computer Engineering, Concordia University, Montreal QCH3G1M8, Canada.

Received: 19 July 2013 Accepted: 21 November 2013
Published: 10 December 2013

References

1. G Stuber, JR Barry, SW Mclaughlin, Y Li, MA Ingram, TG Pratt, Broadband MIMO-OFDM wireless communications. Proc IEEE **92**(2), 241–294 (2004)
2. MR Raghavendra, K Giridhar, Improving channel estimation in OFDM systems for sparse multipath channels. IEEE Signal Process Lett **12**(1), 52–55 (2005)
3. JK Hwang, RL Chung, MF Tsai, JH Deng, Highly efficient sparse multipath channel estimator with Chu-sequence preamble for frequency-domain MIMO DFE receiver. IEICE Trans Commun **E90B**(8), 2103–2110 (2007)
4. C Carbonelli, S Vedantam, U Mitra, Sparse channel estimation with zero-tap detection. IEEE Trans Wireless Commun **6**(5), 1743–1763 (2007)
5. F Wan, W-P Zhu, MNS Swamy, Semiblind most significant tap detection for sparse channel estimation of OFDM systems. IEEE Trans Circuits Syst I: Reg Papers **57**(3), 703–713 (2010)
6. F Wan, W-P Zhu, MNS Swamy, Semiblind sparse channel estimation for MIMO-OFDM systems. IEEE Trans Vehicular Technol **60**(6), 2569–2582 (2011)
7. D Donoho, Compressed sensing. IEEE Trans Inf Theory **52**(4), 1289–1306 (2006)

8. WU Bajwa, J Haupt, AM Sayeed, R Nowak, Compressed channel sensing: a new approach to estimating sparse multipath channels. IEEE Trans on Signal Processing **98**(6), 1058–1076 (2010)
9. CR Berger, S Zhou, JC Preisig, P Willet, Sparse channel estimation for multicarrier underwater acoustic communication: from subspace methods to compressed sensing. IEEE Trans Signal Process **58**(3), 1708–1721 (2010)
10. M Sharp, A Scaglione, A useful performance metric for compressed channel sensing. IEEE Trans Signal Process **59**(6), 2982–2988 (2011)
11. G Taubock, F Hlawatsch, D Eiwen, H Rauhut, Compressive estimation of doubly selective channels in multicarrier systems: leakage effects and sparsity-enhancing processing. IEEE J Sel Top Signal Process **4**(2), 255–271 (2010)
12. X He, R Song, W Zhu, Optimal pilot pattern design for compressed sensing-based sparse channel estimation in OFDM systems. Circuits Syst Signal Process **31**, 1379–1395 (2012)
13. H Mohimani, M Babaie-Zadeh, C Jutten, A fast approach for overcomplete sparse decomposition based on smoothed l_0 norm. IEEE Trans Signal Process **57**(1), 289–301 (2009)
14. A Beck, M Teboulle, A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM J Imaging Sci **2**(1), 183–202 (2009)
15. JA Tropp, AC Gilbert, Signal recovery from random measurements via orthogonal matching pursuit. IEEE Trans Inf Theory **53**(12), 4655–4666 (2007)
16. X Ye, W Zhu, A Zhang, Q Meng, Sparse channel estimation in MIMO-OFDM systems based on an improved sparse reconstruction by separable approximation algorithm. Journal of Information and Computational Science **10**(2), 609–619 (2013)
17. WU Bajwa, A Sayeed, R Nowak, Sparse multipath channels: modeling and estimation, in *Proceedings of the 13th IEEE Digital Signal Processing Workshop* (Marco Island). 4–7 Jan 2009
18. G Gui, F Adachi, Improved least mean square algorithm with application to adaptive sparse channel estimation. EURASIP J Wirel Commun Netw **2013**, 204 (2013)
19. G Gui, A Mehbodniya, F Adachi, Bayesian sparse channel estimation and data detection for OFDM communication systems, in *2013 IEEE 78th Vehicular Technology Conference (VTC2013-Fall)* (, Las Vegas). 2–5 Sept 2013
20. SS Chen, DL Donoho, MA Saunders, Atomic decomposition by basis pursuit. SIAM J Scientif Comput **20**(1), 33–61 (1999)
21. S Mallat, Z Zhang, Matching pursuits with time-frequency dictionaries. IEEE Trans Signal Process **41**(12), 3397–3415 (1993)
22. W Dai, O Milenkovic, Subspace pursuit for compressive sensing signal reconstruction. IEEE Trans Inf Theory **55**(5), 2230–2249 (2009)
23. M Zibulevsky, M Elad, L1-L2 optimization in signal and image processing. IEEE Signal Process Mag **27**(5), 76–88 (2010)
24. SJ Wright, RD Nowak, MAT Figueiredo, Sparse reconstruction by separable approximation. IEEE Trans Signal Process **57**(7), 2479–2493 (2009)
25. JK Pant, W-S Lu, A Antoniou, Recovery of sparse signals from noisy measurements using an l_p regularized least-squares algorithm, in *IEEE Pacific Rim Conference on communications, computers and signal processing* (University of Victoria, Canada), pp. 48–53. 23–26 Aug 2011
26. A Antoniou, W-S Lu, *Practical Optimization: Algorithms and Engineering Applications* (Springer, New York, 2007)
27. JK Pant, W-S Lu, A Antoniou, Reconstruction of sparse signals by minimizing a re-weighted approximate l_0-norm in the null space of the measurement matrix, in *Proceedings of the Midwest Symposium on Circuits and Systems* (Seattle), pp. 430–433. 1–4 Aug 2010
28. JK Pant, W-S Lu, A Antoniou, Unconstrained regularized l_p norm based algorithm for the reconstruction of sparse signals, in *Proceedings of the IEEE International Symposium on Circuits and Systems* (Brazil), pp. 1740–1743. 15–18 May 2011

doi:10.1186/1687-1499-2013-282
Cite this article as: Ye *et al.*: Sparse channel estimation of MIMO-OFDM systems with unconstrained smoothed l_0-norm-regularized least squares compressed sensing. *EURASIP Journal on Wireless Communications and Networking* 2013 **2013**:282.