Group-Sparse Channel Estimation using Bayesian Matching Pursuit for OFDM Systems

Yi Liu[1,2], Wenbo Mei[1] and Huiqian Du[1]
[1]School of Information and Electronics, Beijing Institute of Technology
Beijing, 100081, China
[e-mail: duhuiqian@bit.edu.cn]
[2] Department of Information Technology, College of National Defense Information Science
Wuhan, 430010, China
[e-mail: rfly84@126.com]
*Corresponding author: Huiqian Du

Received May 26, 2014; revised September 30, 2014; accepted December 13, 2014; published February 28, 2015

Abstract

We apply the Bayesian matching pursuit (BMP) algorithm to the estimation of time-frequency selective channels in orthogonal frequency division multiplexing (OFDM) systems. By exploiting prior statistics and sparse characteristics of propagation channels, the Bayesian method provides a more accurate and efficient detection of the channel status information (CSI) than do conventional sparse channel estimation methods that are based on compressive sensing (CS) technologies. Using a reasonable approximation of the system model and a skillfully designed pilot arrangement, the proposed estimation scheme is able to address the Doppler-induced inter-carrier interference (ICI) with a relatively low complexity. Moreover, to further reduce the computational cost of the channel estimation, we make some modifications to the BMP algorithm. The modified algorithm can make good use of the group-sparse structure of doubly selective channels and thus reconstruct the CSI more efficiently than does the original BMP algorithm, which treats the sparse signals in the conventional manner and ignores the specific structure of their sparsity patterns. Numerical results demonstrate that the proposed Bayesian estimation has a good performance over rapidly time-varying channels.

Keywords: Bayesian matching pursuit, orthogonal frequency division multiplexing, group-sparse channel estimation, doubly selective channels, inter-carrier interference

This work was supported by Natural Science Foundation of China (No.61077022).

http://dx.doi.org/10.3837/tiis.2015.02.006 **ISSN : 1976-7277**

1. Introduction

Channel estimation for high mobility wireless communication scenarios is a challenging task in orthogonal frequency division multiplexing (OFDM) systems because the severe Doppler frequency shift destroys the orthogonality of subcarriers, which causes strong inter-carrier interference (ICI) to pilot symbols. This interference results in a sharp decline of the estimation performance. To address the ICI, traditional estimation methods [1] such as the least square (LS) method require a large number of pilot symbols, thereby leading to a reduced bandwidth efficiency.

Recently, the compressive sensing (CS [2, 3])-based channel estimation [4-6] has gained a fast-growing interest. Compared with traditional pilot-aided methods, the CS-based channel estimation is considered as a more promising technology because it is able to take advantage of the inherent sparsity of the propagation channel. This superiority enables the latter to detect the channel status information (CSI) with considerably reduced pilot symbols in high mobility communication environments. Moreover, in the recent work [7, 8], experimental results demonstrated that CS-based methods are clearly superior to traditional estimation methods over some real-world multipath channels (e.g., the underwater acoustic channel) in OFDM systems.

As an extension of CS, the Bayesian CS (BCS) [9-11] technologies such as the Bayesian matching pursuit (BMP) [12-14] are based on prior statistics of the sparse signal. Using the additional priori statistical knowledge, the Bayesian recovery method outperforms the conventional sparse recovery methods in many applications. In the field of channel estimation, some statistical information of the channel can also be obtained in some scenarios, providing the potential for a more precise estimation of the channel or a further reduction of pilot symbols. Based on the available channel statistics, the recent work presented in [6] developed a statistical basis optimization procedure, which enables channel coefficients to be represented by a sparser basis expansion model. Simulation results in this literature showed that the statistically optimized basis can achieve an improvement of the estimation performance for doubly selective channels. However, until now there has been little CS-based work that considers the direct application of BCS approaches to the channel estimation. Therefore, we believe that the statistical information of the channel is not effectively exploited in the existing sparse channel estimation methods.

Compared with the related work mentioned above, the uniqueness of the approach presented in this paper can be summarized as follows. First, using the channel basis expansion function in the delay-Doppler domain, we depict the group-sparsity (i.e., block sparsity) [15, 16] and statistical properties of doubly selective channels. Based on these priors, we represent the CSI by a statistical group-sparsity model, which enables us to formulate the channel estimation as a problem of reconstructing the statistical group-sparsity signal. Then, we propose a BMP-based method to reconstruct the CSI for rapidly time-varying OFDM systems. In this estimation scheme, we adopt an approximate model of the fast fading transmission system. This approximation allows us to handle the Doppler-induced ICI with a relatively low computational complexity. Moreover, because the BMP algorithm treats the signal as being sparse in the conventional sense and ignores the specific structure of their sparsity patterns, we make some small changes to BMP, which enable the algorithm to make good use of the group-sparse structure of the channel. Simulation results demonstrate that these modifications lead to a further reduction of the computational cost. Finally, we also analyze the influence of

different pilot patterns on the estimation performance and then present an approximately optimal pilot placement for the proposed estimation scheme.

The remainder of the paper is organized as follows. We introduce some fundamentals about the BMP algorithm and the OFDM system in the next section. The sparse and statistical properties of the doubly selective channel are described in Section 3. In Section 4, we propose a low-complexity estimation method which is based on the modified BMP algorithm. The pilot design is also studied. Section 5 shows the experimental results. Conclusions are finally drawn in Section 6.

2. Preliminaries

2.1 BMP Method

CS and other sparse recovery methods focus on the classical linear measurement model defined as follows

$$\mathbf{Y} = \mathbf{\Phi u} + \mathbf{Z} \tag{1}$$

where $\mathbf{u} \in \mathbb{C}^M$ is an unknown K-sparse vector (has only K nonzero entries, $K \ll M$), $\mathbf{Y} \in \mathbb{C}^N$ ($K < N < M$) is a known measurement signal, $\mathbf{\Phi} \in \mathbb{C}^{N \times M}$ is the measurement matrix and $\mathbf{Z} \sim \mathrm{CN}(0, \sigma_Z^2 \mathbf{I}_N)$ is an additive Gaussian noise. It has been proved in [2] that $\mathbf{u}$ can be reconstructed from $\mathbf{Y}$ by conventional sparse recovery methods [2, 3] if the measurement matrix satisfies the so-called restricted isometry property (RIP) [17].

In some scenarios, the sparse signal exhibits an additional structure in the form that the nonzero elements occur in groups, not arbitrary places. By taking advantage of the structured sparsity, the group-sparse recovery methods [15, 16] can further compress the information of sparse signals; thus yield an improved recovery performance compared with traditional CS algorithms.

As another special case of CS technologies, the BMP algorithm [12-14] and other BCS methods are based on the additional statistical information of the sparse signal. Generally, $\mathbf{u}$ is considered as a mixed Bernoulli-Gaussian process in these approaches. More specifically, let S be the support of $\mathbf{u}$, then the elements of S are assumed to be sampled from $\{1, 2, \ldots, M\}$ according to the Bernoulli distribution with success probability P. Under this assumption, we obtain the probability of S:

$$p(\mathrm{S}) = P^{|\mathrm{S}|}(1-P)^{M-|\mathrm{S}|} \tag{2}$$

It is observed that the sparsity of $\mathbf{u}$ is controlled by P ($\mathrm{E}[K] = PM$). Defining $\mathbf{u}_{\mathrm{S}}$ to be the subvector of $\mathbf{u}$, which is indexed by S, we further assume that the entries of $\mathbf{u}_{\mathrm{S}}$ are drawn from an Gaussian process with zero-mean and known variance σ^2. Then, the minimum mean-square error (MMSE) estimate of $\mathbf{u}$ from $\mathbf{Y}$ is given by

$$\hat{\mathbf{u}}_{\mathrm{MMSE}} = \mathrm{E}[\mathbf{u} \mid \mathbf{Y}] = \sum_{\mathrm{S}} \mathrm{E}[\mathbf{u} \mid \mathbf{Y}, \mathrm{S}] p(\mathrm{S} \mid \mathbf{Y}) \tag{3}$$

and the maximum *a posteriori* (MAP) estimate is expressed as

$$\hat{\mathbf{u}}_{\mathrm{MAP}} = \mathrm{E}[\mathbf{u} \mid \mathbf{Y}, \mathrm{S}_{\mathrm{MAP}}] \tag{4}$$

where

$$\mathrm{S}_{\mathrm{MAP}} = \arg\max_{\mathrm{S}} p(\mathrm{S} \| \mathbf{Y}) \tag{5}$$

To calculate the two estimates, the expectation $\mathrm{E}[\mathbf{u} \mid \mathbf{Y}, \mathrm{S}]$ and the posterior $p(\mathrm{S} \| \mathbf{Y})$ must be evaluated. We first consider the calculation of the former. Because the relationship between **u** and **Y** is linear, the linear MMSE (LMMSE) estimate can be used to evaluate the expectation [14]:

$$\mathrm{E}[\mathbf{u} \mid \mathbf{Y}, \mathrm{S}] = \sigma^2 \mathbf{\Phi}_{\mathrm{S}}^H \left(\sigma_{\mathbf{Z}}^2 \mathbf{I}_N + \sigma^2 \mathbf{\Phi}_{\mathrm{S}} \mathbf{\Phi}_{\mathrm{S}}^H \right)^{-1} \mathbf{Y} \tag{6}$$

where $\mathbf{\Phi}_{\mathrm{S}}$ is the submatrix that consists of the columns of $\mathbf{\Phi}$, indexed by S, $\mathbf{I}_N$ is an $N \times N$ identity matrix, and $(\cdot)^H$ denotes the conjugate transpose matrix.

Next, using the Bayesian rule, we have

$$p(\mathrm{S} \| \mathbf{Y}) = \frac{p(\mathbf{Y} \| \mathrm{S}) p(\mathrm{S})}{p(\mathbf{Y})} \tag{7}$$

where the factor $p(\mathbf{Y})$ can be ignored since it is common for each potential support, and $p(\mathrm{S})$ is obtained from (2). Then, we focus on the prior probability $p(\mathbf{Y} \| \mathrm{S})$. Because the entries of $\mathbf{u}_{\mathrm{S}}$ are drawn from $\mathcal{CN}(0, \sigma^2)$, we have that $\mathbf{Y} \| \mathrm{S}$ is $\mathcal{CN}(0, \mathbf{\Sigma}_{\mathrm{S}})$, where

$$\mathbf{\Sigma}_{\mathrm{S}} = \sigma_{\mathbf{Z}}^2 \mathbf{I}_N + \sigma^2 \mathbf{\Phi}_{\mathrm{S}} \mathbf{\Phi}_{\mathrm{S}}^H \tag{8}$$

So,

$$p(\mathbf{Y} \| \mathrm{S}) = \frac{1}{(\sqrt{2\pi})^N \left| \mathbf{\Sigma}_{\mathrm{S}} \right|^{1/2}} \exp\left(-\frac{1}{2} \mathbf{Y}^H \mathbf{\Sigma}_{\mathrm{S}}^{-1} \mathbf{Y} \right) \tag{9}$$

Note that if the dimension of **u** is large, there are quite many possible supports. In this case, computing the posterior and expectation for all these supports requires an unacceptable computational cost. To simplify the estimation, the computation should be restricted to the dominant supports, i.e., the supports which have significant posteriors. Defining the dominant support set as S^*, we approximate the MMSE estimate as follows

$$\hat{\mathbf{u}}_{\mathrm{AMMSE}} = \sum_{\mathrm{S} \in \mathrm{S}^*} \mathrm{E}[\mathbf{u} \mid \mathbf{Y}, \mathrm{S}] p(\mathrm{S} \| \mathbf{Y}) \tag{10}$$

which is called the AMMSE estimate. Note that the MAP estimate can be treated as an extreme case of AMMSE, where the size of S^* is reduced to one. Therefore, we only analyze the AMMSE estimate in the rest of the paper. The search procedure of S^* is presented in Section 4.2.

2.2 OFDM System

We focus on the discrete time-varying transmission model

$$y[n] = \sum_{l=0}^{L-1} h[n,l]\, x[n-l] + z[n] \tag{11}$$

where $h[n, l]$ denotes the channel impulse response (L is the maximum multipath delay), $x[n]$ is the input signal, $y[n]$ is the output signal and $z[n]$ is the Gaussian noise with zero-mean and variance $\sigma_{\mathbf{z}}^2$. Over a block of N symbols, the linear system can be written in matrix notation as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{z} \tag{12}$$

where the vectors $\mathbf{x} = (x[0],\ldots, x[N-1])^T$, $\mathbf{y} = (y[0],\ldots, y[N-1])^T$, $\mathbf{z} = (z[0],\ldots, z[N-1])^T$ and $\mathbf{H} \in \mathbb{C}^{N\times N}$ represents the channel matrix:

$$\mathbf{H} = \begin{bmatrix} h[0,0] & 0 & \ldots & 0 & h[0,L-1] & \ldots & \ldots & h[0,1] \\ h[1,1] & h[1,0] & 0 & \ldots & 0 & h[1,L-1] & \ldots & h[1,2] \\ & & & \vdots & & & & \\ h[L-1,L-1] & \ldots & h[L-1,0] & 0 & \ldots & \ldots & \ldots & 0 \\ & & & \vdots & & & & \\ 0 & \ldots & \ldots & 0 & h[N-1,L-1] & \ldots & \ldots & h[N-1,0] \end{bmatrix} \tag{13}$$

We construct an OFDM system with N subcarriers. Let $\mathbf{Q} \in \mathbb{C}^{N\times N}$ be the standard discrete Fourier transform (DFT) matrix ($Q(p,q) = 1/\sqrt{N}e^{-j2\pi(p-1)(q-1)/N}$), and $\mathbf{X}$, $\mathbf{Y}$ be the transmitted and received data vectors, respectively. Then, we have $\mathbf{x} = \mathbf{Q}^H\mathbf{X}$ and $\mathbf{Y} = \mathbf{Q}\mathbf{y}$. To remove the inter-symbol interference (ISI) from the OFDM system, we assume that the length of the cyclic-prefix (CP) is longer than L. Then, over a block of an OFDM symbol, we can express the overall system model as follows

$$\mathbf{Y} = \mathbf{G}\mathbf{X} + \mathbf{Z} \tag{14}$$

where the noise vector is $\mathbf{Z} = \mathbf{Q}\mathbf{z}$ and the overall system matrix is $\mathbf{G} = \mathbf{Q}\mathbf{H}\mathbf{Q}^H$. The system model can also be written in the following form

$$Y_k = G_k X_k + \sum_{m=1, m\neq k}^{N} G_{k,m} X_m + Z_k \tag{15}$$

where the non-diagonal entries of $\mathbf{G}$ represent the ICI response. Inserting $\mathbf{G} = \mathbf{Q}\mathbf{H}\mathbf{Q}^H$ into (15), we have

$$\begin{aligned} Y_k &= \sum_{m=1}^{N}\left(\frac{1}{N}\sum_{n=0}^{N-1}\sum_{l=0}^{L-1}h[n,l]e^{-j2\pi(nk+ml-nm-l)/N}\right)X_m + Z_k \\ &= \sum_{n=0}^{N-1}\sum_{l=0}^{L-1}h[n,l]\left(\frac{1}{N}\sum_{m=1}^{N}X_m e^{-j2\pi(nk+ml-nm-l)/N}\right) + Z_k \end{aligned} \tag{16}$$

We define the impulse response vector $\mathbf{h} \in \mathbb{C}^{NL}$ as follows

$$\begin{aligned} &\mathbf{h} = (\mathbf{h}_0, \ldots, \mathbf{h}_{L-1})^T, \\ &\mathbf{h}_l = (h[0,l], \ldots, h[N-1,l]) \end{aligned} \tag{17}$$

and an $N \times NL$ matrix $\mathbf{A}$ as follows

$$A_{k,q} = \frac{1}{N}\sum_{m=1}^{N}X_m e^{-j2\pi(nk+ml-nm-l)/N} \tag{18}$$

where $n = (q-1) \bmod (N)$, and $l = (q-1-n)/N$. Then, (16) can be expressed in matrix notation as

$$\mathbf{Y} = \mathbf{A}\mathbf{h} + \mathbf{Z} \tag{19}$$

The system model is changed as a function of **h**, which is convenient for us to address the channel estimation problem.

3. Sparsity and Statistical Characteristics

We construct a doubly selective channel which consists of *S* propagation paths. The channel impulse response is expressed as

$$h(t,\tau) = \sum_{s=1}^{S}\alpha_s \delta(\tau - \tau_s)e^{j2\pi f_s t} \tag{20}$$

where $\delta(\cdot)$ is the Dirac function, α_s, τ_s and f_s denote the magnitude, multipath delay and Doppler frequency, respectively. Assuming that the sampling period is *T*, we approximately consider that τ_s / T are integers for $s = 1, \ldots, S$. In this case, the discrete impulse response is given by

$$h[n,l] = \sum_{s=1}^{S}\alpha_s \delta(l - \tau_s/T)e^{j2\pi f_s nT} \tag{21}$$

We first investigate the sparse structure of the channel. To describe the sparsity in the delay-Doppler domain, we use the delay-Doppler basis expansion:

$$u[l,d] = \frac{1}{\sqrt{N}}\sum_{n=0}^{N-1}h[n,l]e^{-j2\pi dn/N} = \sum_{s=1}^{S}u_s[l,d] \tag{22}$$

where

$$u_s[l,d]=\frac{1}{\sqrt{N}}\alpha_s\delta(l-\tau_s/T)\sum_{n=0}^{N-1}e^{-j2\pi(d-f_sNT)n/N} \\ l=0,\ldots,L-1,\ d=0,\ldots,N-1 \tag{23}$$

Now we focus on the following function

$$E_s[d]=\left|\sum_{n=0}^{N-1}e^{-j2\pi(d-f_sNT)n/N}\right|=\left|\frac{\sin(\pi(d-f_sNT))}{\sin(\pi(d-f_sNT)/N)}\right| \tag{24}$$

which describes the energy distribution of $u_s[l,d]$ in the frequency domain. Generally, the normalized Doppler frequency is $|f_sNT| \ll 1$; thus the frequency offset is constrained in a small band around the zero point [6-8, 18]. An example of $E_s[d]$ is shown in **Fig. 1**.

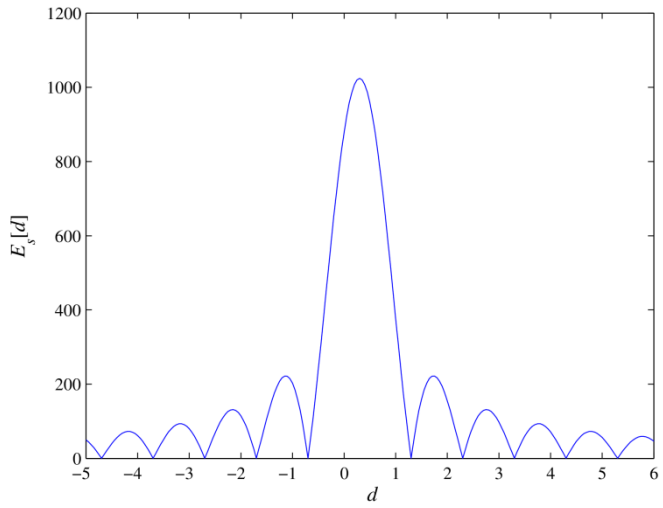

Fig. 1. An example of $E_s[d]$, N = 1024, $f_sNT = 0.3$

As observed in **Fig. 1**, the energy of $E_s[d]$ is concentrated in the subspace $d \in [-D, D]$ for an appropriately chosen D ($[-D,-1]$ is the circular shift of $[N-D, N-1]$). Because $\delta(l-\tau_s/T)=0$ if $l \neq \tau_s/T$, the dominant components of $u_s[l,d]$ are mainly concentrated in the group $\{u_s[\tau_s/T,0],\ldots,u_s[\tau_s/T,D],\ u_s[\tau_s/T,N-D],\ldots,u_s[\tau_s/T,N-1]\}$. For all S paths, the significant elements of $u[l,d]$ cluster in S such groups. Ignoring the small-magnitude elements, we approximately consider that $u[l,d]$ has only S nonzero groups, which only occur in the subset $[0,L-1]\times([0,D]\cup[N-D,N-1])$. The number of valid channel coefficients is decreased from NL to $(2D+1)L$.

We define a $(2D+1)L$-length vector as follows:

$$\begin{aligned}&\mathbf{u}=(\mathbf{u}_0,\ldots,\mathbf{u}_{L-1})^T,\\&\mathbf{u}_l=(u[l,0],\ldots,u[l,D],u[l,N-D],\ldots,u[l,N-1])\end{aligned}\tag{25}$$

Then, only S of the groups $(\mathbf{u}_0,\ldots,\mathbf{u}_{L-1})$ are nonzero. If $S \square L$, $\mathbf{u}$ is approximately an group-sparse vector over this group partition.

Next, we investigate the channel statistical properties. Without loss of generality, we assume that for $s=1,\ldots,S$, τ_s and f_s are uniformly sampled from $\{0,T,\ldots,(L-1)T\}$ and $[-f_{\max}, f_{\max}]$, respectively, and $\alpha_s \square \mathbb{CN}(0,\sigma^2)$ is a Gaussian process. Because $\mathbf{u}_l$ is nonzero only if $l=\tau_s/T$, we can approximately consider that each group of $\mathbf{u}$ has the probability $P=S/L$ to be nonzero (i.e., a Bernoulli process). We assume that $\mathbb{S}$ is a possible support of $\mathbf{u}$, which corresponds to a certain group-sparse pattern. Then, the probability of $\mathbb{S}$ is given by

$$p(\mathbb{S})=P^{|\mathbb{S}|/(2D+1)}(1-P)^{L-|\mathbb{S}|/(2D+1)}\tag{26}$$

Moreover, the entries within each nonzero group of $\mathbf{u}$ are Gaussian. Therefore, $\mathbf{u}$ can be considered as a group-version of the mixed Bernoulli-Gaussian process.

4. BMP-Based Channel Estimation

In this section, we propose a BMP-based estimation method to reconstruct the CSI. In the proposed estimation, we make some changes to BMP, which enable the algorithm to address the group-sparse channel coefficients in a more efficient manner. Furthermore, a discussion about the pilot pattern is also presented.

4.1 Proposed estimation scheme

We consider the representation of (19) in the delay-Doppler domain. From (22), the relationship between $\mathbf{h}$ and $\mathbf{u}$ is obtained as

$$\mathbf{h}=(\mathbf{I}_L\otimes(\mathbf{FQ})^H)\mathbf{u}\tag{27}$$

where $\otimes$ denotes the Kronecker product, and $\mathbf{F}\in\square^{(2D+1)\times N}$ denotes the uniform down-sampling operation from $[1,N]$ to $[1,D+1]\cup[N-D+1,N]$. Using $\mathbf{u}$ to replace $\mathbf{h}$, we transform (19) into the form

$$\mathbf{Y}=\mathbf{\Phi u}+\mathbf{Z}\tag{28}$$

where the $N\times(2D+1)L$ matrix $\mathbf{\Phi}=\mathbf{A}(\mathbf{I}_L\otimes(\mathbf{FQ})^H)$. We define a $(2D+1)L$-length vector

$$\begin{aligned}&\mathbf{m}=(\mathbf{m}_0,\ldots,\mathbf{m}_{L-1}),\\&\mathbf{m}_i=\left(Ni+1,\ldots,Ni+D+1,Ni+N-D+1,\ldots,Ni+N\right)\end{aligned}\tag{29}$$

and, subsequently, $n=(m_q-1)\bmod(N)$ and $l=(m_q-1-n)/N$ for $q=1,\ldots,(2D+1)L$. Then, the entries of $\mathbf{\Phi}$ are expressed as follows

$$\Phi_{k,q} = \frac{1}{\sqrt{N}} e^{-j2\pi(k-n-1)l/N} X_{(k-n-1)\bmod(N)+1} \tag{30}$$

It is seen that $X_{(k-n-1)\bmod(N)+1} \neq X_k$ if $n \neq 0$. Essentially, these symbols are the direct neighbors on both sides of X_k. Hence, the corresponding entries $\Phi_{k,q}$ describe the interference of these neighboring symbols (i.e., the ICI response) on Y_k. Assume that there are N_P pilots in the N-symbol block, and the pilot tones belong to the subset $\mathrm{P} = \{k(1),\ldots,k(N_P)\}$. Selecting the measurements at the pilot positions, we form the measurement model

$$\begin{pmatrix} Y_{k(1)} \\ \vdots \\ Y_{k(N_P)} \end{pmatrix} = \begin{pmatrix} \Phi_{k(1),1} & \cdots & \Phi_{k(1),(2D+1)L} \\ & \vdots & \\ \Phi_{k(N_P),1} & \cdots & \Phi_{k(N_P),(2D+1)L} \end{pmatrix} \mathbf{u} + \begin{pmatrix} Z_{k(1)} \\ \vdots \\ Z_{k(N_P)} \end{pmatrix} \tag{31}$$

or

$$\mathbf{Y}_{\mathrm{P}} = \mathbf{\Phi}_{\mathrm{P}} \mathbf{u} + \mathbf{Z}_{\mathrm{P}} \tag{32}$$

In the pth row of $\mathbf{\Phi}_{\mathrm{P}}$, the entries $\Phi_{k(p),q}$ corresponding to $(k(p)-n-1)\bmod(N)+1 \notin \mathrm{P}$ (n was defined in (30)) are unknown because we know only the pilot symbols, not data symbols. These unknown entries represent the ICI from the data symbols. To construct the measurement equation without the pre-estimation of $\mathbf{\Phi}_{\mathrm{P}}$, we should ignore these unknown components. The approximate matrix is defined as follows

$$\tilde{\Phi}_{\mathrm{P}}[p,q] = \begin{cases} \Phi_{\mathrm{P}}[p,q], & (k(p)-n-1)\bmod(N)+1 \in \mathrm{P} \\ 0, & \text{otherwise} \end{cases} \tag{33}$$

Using $\mathbf{w} = (\mathbf{\Phi}_{\mathrm{P}} - \tilde{\mathbf{\Phi}}_{\mathrm{P}})\mathbf{u} + \mathbf{Z}_{\mathrm{P}}$ to replace $\mathbf{Z}_{\mathrm{P}}$, we change (32) into the form

$$\mathbf{Y}_{\mathrm{P}} = \tilde{\mathbf{\Phi}}_{\mathrm{P}} \mathbf{u} + \mathbf{w} \tag{34}$$

To reconstruct **u** from (34), we use the modified BMP algorithm, which is described in Section 4.2. With the estimate of **u**, the impulse response vector **h** is obtained via (27).

4.2 Modified BMP Algorithm

To compute the AMMSE estimate of **u**, we must look for appropriate supports to form the dominant set S^* in (10) at first. In [12, 13], a fast greedy approach was proposed to search the best supports of different sizes. Here, to make full use of the group-sparsity of **u**, we make some modifications to this greedy approach. And, we will show that the modified algorithm can significantly reduce the computational complexity.

Assume that the sparsity of **u** is unknown. Then, all potential support sizes must be taken into consideration. Since **u** is a group-sparse vector, we investigate only the group-sparse samples instead of all sparse ones. Assuming the maximum possible group-sparsity of **u** is $S_{\max}$, we consider the group-sparsity that belongs to $[1, S_{\max}]$. In other words, the support sizes of **u** are selected from the subset $\{2D+1, 2(2D+1),\ldots, S_{\max}(2D+1)\}$. For each potential size,

we choose the support $\mathbb{S}$ which has the highest posterior probability. For example, when searching the optimal support of size $2D+1$ (i.e., one nonzero group), we evaluate $p(\mathbb{S}_1\|\mathbf{Y})$ for $\mathbb{S}_1=\mathbb{U}_0, \ldots, \mathbb{U}_{L-1}$, where $\mathbb{U}_l$ is the support corresponding to $\mathbf{u}_l$ ($\mathbf{u}_l$ was defined in (25)). The support which maximizes $p(\mathbb{S}_1\|\mathbf{Y})$ is considered optimal for this support size. We assume that the optimal support is $\mathbb{S}_1=\mathbb{U}_i$. When evaluating supports of size $2(2D+1)$, we choose $\mathbb{U}_j\neq\mathbb{U}_i$, which enables $\mathbb{S}_2=\mathbb{U}_i\cup\mathbb{U}_j$ to maximize $p(\mathbb{S}_2\|\mathbf{Y})$. When the support size further increases, the greedy search continues in the same manner. Finally, there are $S_{\max}$ supports selected to form $\mathbb{S}^*$.

With the dominant set, the AMMSE estimate in (10) can be calculated using the fast recursive method proposed in [12]. In contrast to AMMSE, the MAP estimate in (4) needs less steps since it considers only the optimal support instead of all dominant supports. However, the reduced complexity probably leads to performance degradation. Detailed steps of the modified algorithm are presented in **Table 1**.

Table 1. Modified BMP algorithm

Input: $\mathbf{Y}$, $\mathbf{\Phi}$, σ^2, $\sigma_{\mathbf{z}}^2$, $S_{\max}$, P

Initialize: $\mathcal{B}_0=\{\mathbb{U}_0,\ldots,\mathbb{U}_{L-1}\}$, $\mathbb{S}_0=\{\}$, $\mathbb{S}^*=\{\}$, $i=1$

Support search:

While $i\le S_{\max}$, do

1. $\Omega=\left\{\mathbb{S}_{i-1}\cup\lambda_1,\mathbb{S}_{i-1}\cup\lambda_2,\ldots,\mathbb{S}_{i-1}\cup\lambda_{|\mathcal{B}_{i-1}|}\,\middle|\,\lambda_k\in\mathcal{B}_{i-1}\right\}$;
2. $\mathbb{S}_i=\arg\max_{\mathbb{S}\in\Omega} p(\mathbb{S}\|\mathbf{Y})$;
3. $\mathbb{S}^*=\{\mathbb{S}^*,\mathbb{S}_i\}$;
4. $\mathcal{B}_i=\mathcal{B}_0\setminus\mathbb{S}_i$;
5. $i=i+1$.

End while

$\mathbb{S}_{\text{MAP}}=\arg\max_{\mathbb{S}\in\mathbb{S}^*} p(\mathbb{S}\|\mathbf{Y})$

Compute: $\hat{\mathbf{u}}_{\text{AMMSE}}$, $\hat{\mathbf{u}}_{\text{MAP}}$

Output: $\mathbb{S}^*$, $\mathbb{S}_{\text{MAP}}$, $\hat{\mathbf{u}}_{\text{AMMSE}}$, $\hat{\mathbf{u}}_{\text{MAP}}$

Using the modified BMP algorithm, in each iteration of the support search procedure, we only need to calculate the posterior around L times. Compared with the modified algorithm, the original BMP algorithm requires a far larger number of calculations (about $(2D+1)L$ times) because the latter cannot use the group-sparse structure of $\mathbf{u}$. Furthermore, the iteration number of the modified BMP algorithm is also decreased from $(2D+1)S_{\max}$ to $S_{\max}$. Hence, the total number of operations is reduced by $(2D+1)^2$ times. Obviously, the modified BMP algorithm has a far lower computational complexity.

To improve the accuracy of the AMMSE estimate, we should expand the dominant set $\mathbb{S}^*$ to involve more potential supports. This can be done by repeating the support search step several times. In each repetition, the supports which already belong to $\mathbb{S}^*$ should not be chosen again. If we repeat the search procedure R times, the size of $\mathbb{S}^*$ is increased from $S_{\max}$ to $RS_{\max}$.

In Table 1, $S_{\max}$, P, σ^2 and σ_z^2 can be estimated from some prior channel information or set as predefined values. The initial estimates of these hyperparameters should be refined during the repeated search procedure. In the rth repetition ($r>1$), assuming that $\mathbb{S}_{\text{MAP}}^{(r-1)}$ is the optimal support chosen in the previous round and $\hat{\mathbf{u}}_{\text{MAP}}^{(r-1)}$ is the corresponding MAP estimate, we update the hyperparameters as follows:

$$S_{\max} = \frac{\left|\mathbb{S}_{\text{MAP}}^{(r-1)}\right|}{2D+1} \tag{35}$$

$$P = \frac{S_{\max}}{L} \tag{36}$$

$$\sigma^2 = \text{var}(\hat{\mathbf{u}}_{\text{MAP}}^{(r-1)}) \tag{37}$$

$$\sigma_{\mathbf{z}}^2 = \text{var}(\mathbf{Y}_{\text{P}} - \tilde{\mathbf{\Phi}}_{\text{P}}\hat{\mathbf{u}}_{\text{MAP}}^{(r-1)}) \tag{38}$$

4.3 Pilot Design

The performance of the proposed estimation scheme is mainly influenced by two issues: one is the approximation error $\mathbf{w}-\mathbf{Z}_{\text{P}} = (\mathbf{\Phi}_{\text{P}} - \tilde{\mathbf{\Phi}}_{\text{P}})\mathbf{u}$, which is equivalent to the data-induced ICI; the other is the reconstruction error. To eliminate the former, the pilots should be put together and far away from the data symbols. This consecutive pilot pattern enables $\tilde{\mathbf{\Phi}}_{\text{P}}$ to be extremely approximate to $\mathbf{\Phi}_{\text{P}}$, thereby minimizing the approximation error. On the other hand, for the sparse signal reconstruction, the pilot symbols are suggested to be set at random places [4, 6]. However, in this case the pilots are very likely to be separated from each other and close to the data symbols.

For the trade-off between the accurate approximation of the system model and the stable reconstruction performance, we adopt the combination of the above two pilot patterns. Specifically, we divide the pilots into several groups and put these groups at random positions. Because each symbol mainly affects its adjacent ones, only the pilots at the borders of the groups are interfered by the unknown data symbols if the group size is sufficiently large. In this case, the data-induced ICI can be significantly reduced. Hence, compared with the random pilot pattern, the proposed combo pilot pattern has an improved robustness to the ICI. On the other hand, compared with the consecutive pilot pattern, the combo pilot pattern is expected to be more appropriate for the sparse signal reconstruction. The actual effects of the three pilot patterns are investigated in the next section.

5. Numerical Simulations

5.1 Simulation Settings

We assume that the OFDM system is divided into 1024 subcarriers, and the CP length ratio is $1/4$. We also simulate a doubly selective channel with ten paths, whose magnitudes $\alpha_1,\ldots,\alpha_{10}$ are drawn from $\mathbb{CN}(0,1)$. The delay-Doppler points $(\tau_s/T, f_s NT)$ are uniformly chosen from the subset $\{0,1,\ldots,255\}\times[-f_{\max}NT, f_{\max}NT]$, where the maximum normalized Doppler frequency $f_{\max}NT$ varies from 0.1 to 0.3. The additive channel noise is a zero-mean Gaussian process. The pilots are independently and randomly sampled from the QPSK

alphabets, and the pilot ratio is $1/8$.

Based on the approximate measurement model (34) with $D=2$, we reconstruct $\mathbf{u}$ using the modified BMP algorithm. The support search is repeated $R=3$ times, and both the AMMSE and MAP estimates are calculated. The initial value of σ^2 is roughly estimated using the LS estimate [1]. The initialization of the other hyperparameters is given by: $S_{\max}=7$, $P=S_{\max}/256$ and $\sigma_{\mathbf{z}}^2=\sigma^2/10$. Moreover, we present the performances of the basis pursuit de-noising (BPDN) algorithm [19] and the orthogonal matching pursuit (OMP) algorithm [20] for comparison. Finally, the runtime of the estimation is also investigated.

5.2 Numerical Results

We first focus on the comparison between the proposed Bayesian approach and the conventional CS-based methods. The combo pilot pattern is adopted with the group size of $2D+1$. **Fig. 2** depicts the mean square error (MSE) performance versus SNR when $f_{\max}NT=0.2$. **Fig. 3** shows the MSE performance versus $f_{\max}NT$ when $\mathrm{SNR}=20\,\mathrm{dB}$. On different conditions, the performances of the two modified-BMP estimators (AMMSE and MAP) are clearly superior to those of BPDN and OMP. The performance advantage demonstrates that the proposed Bayesian method effectively exploits the prior statistics and the group-sparsity of channel coefficients. Compared with the MAP estimate, the AMMSE estimate has an improved performance as expected. Moreover, the gap between the two becomes larger when low SNR or large Doppler spreading occurs, indicating that the AMMSE estimate is more robust to extreme channel conditions. This additional superiority of AMMSE is because that the repeated support search procedure enhances its performance in poor communication environments.

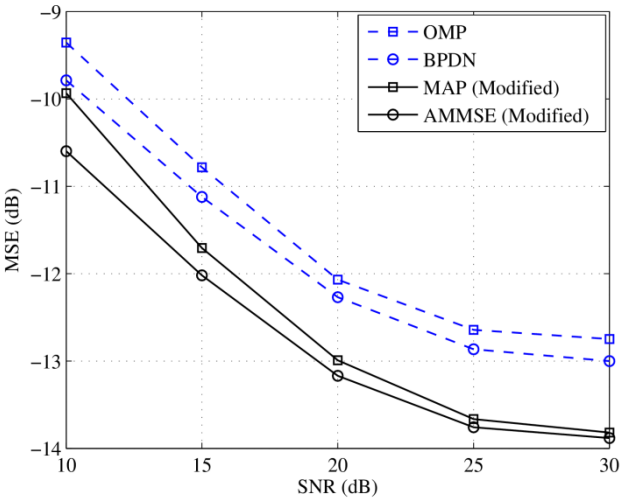

Fig. 2. Performance versus the SNR.

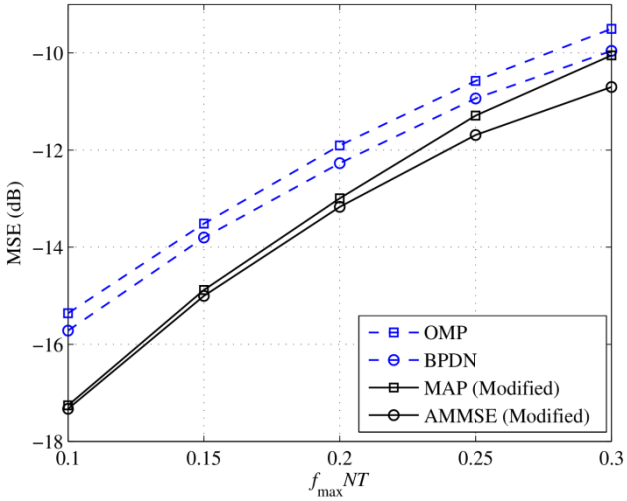

Fig. 3. Performance versus the Doppler frequency.

Next, **Figs. 4** and **5** depict the performance of the modified BMP algorithm for different pilot patterns. Here, we compare the three previously mentioned pilot patterns: Pattern 1 is the consecutive pilot pattern; Pattern 2 is the random pilot pattern; Pattern 3 is the combo pilot pattern. The AMMSE estimate is adopted. It is observed that Pattern 3 outperforms the other two for different SNRs and Doppler frequencies. The good performance of this pattern demonstrates that it effectively reduces both the data-induced ICI and the reconstruction error. From **Fig. 5**, it is seen that the performance of Pattern 2 is close to that of Pattern 3 if the Doppler spreading is small. However, the former is more sensitive to $f_{\max}NT_s$ since it cannot address the severe ICI. Pattern 1 has a much worse performance than the other two pilot patterns do, indicating that it is not applicable for sparse recovery methods. In summary, the combo pilot pattern is very likely to be optimal for the proposed Bayesian estimation scheme over rapidly time-varying channels.

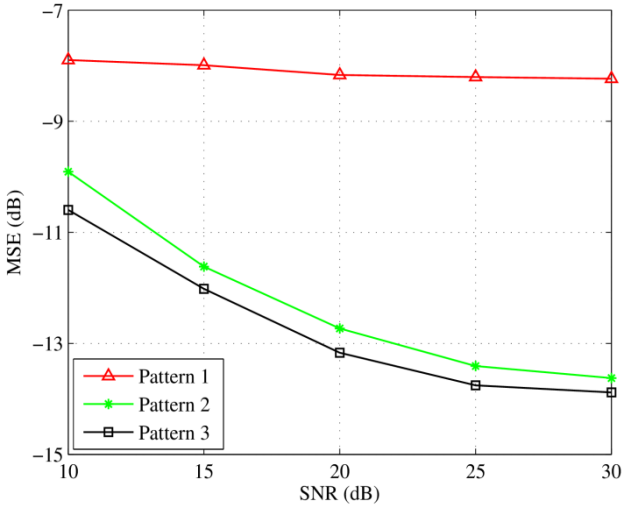

Fig. 4. Performances of different pilot patterns versus the SNR

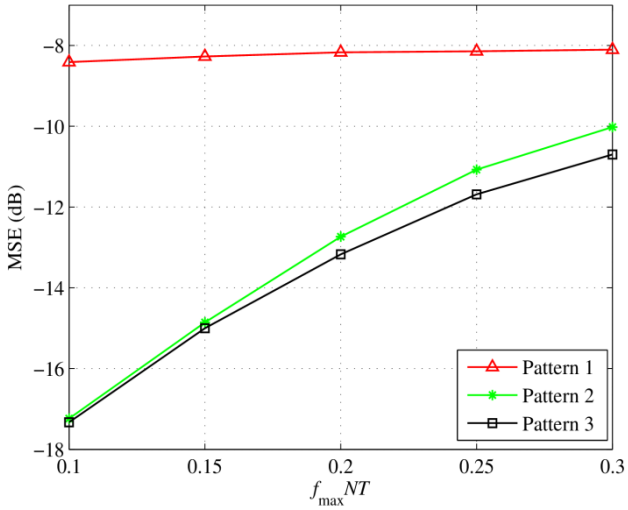

Fig. 5. Performances of different pilot patterns versus the Doppler frequency

Finally, **Figs. 6** and **7** show the mean runtime of these recovery methods. Here we still adopt the AMMSE estimate in the two BMP methods. As can be seen in the plots, the two BMP algorithms need shorter runtime than does the traditional CS-based method because of the efficient use of channel statistics. It is also observed that the modified BMP method is much faster than the original one, demonstrating that the proposed modifications effectively simplify the reconstruction steps.

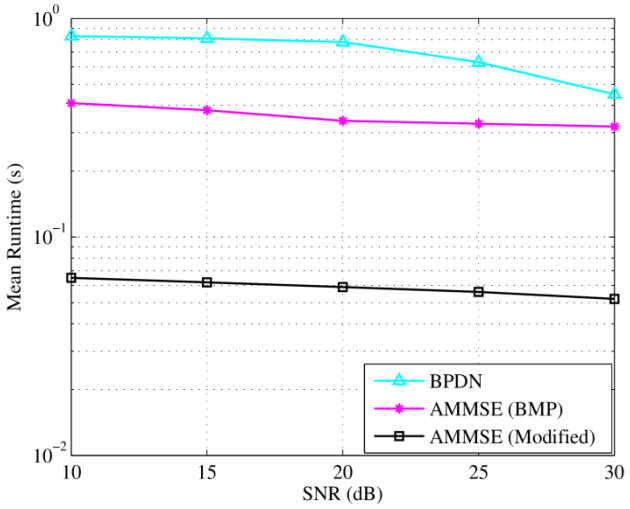

Fig. 6. Mean runtime versus the SNR.

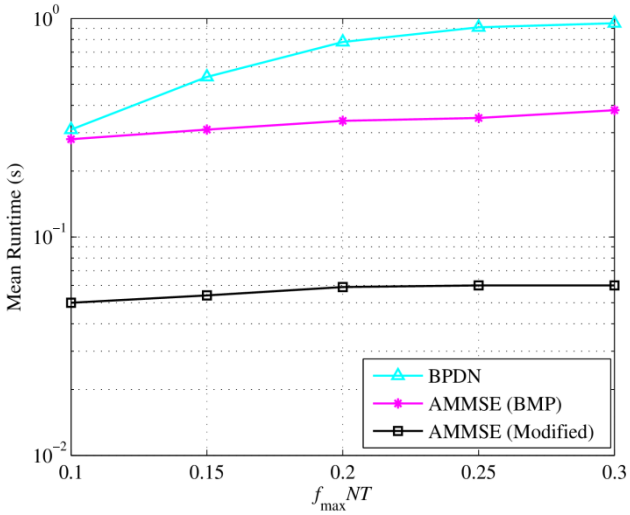

Fig. 7. Mean runtime versus the Doppler frequency.

6. Conclusion

We considered the application of the BMP algorithm to the channel estimation for rapidly time-varying OFDM systems. By exploiting the basis expansion function of the channel, we depicted the group-sparse structure and statistical characteristics of the doubly selective channel. Then, we proposed a BMP-based estimation scheme to reconstruct the CSI. In this method, we adopted both the MMSE and MAP estimates, providing a trade-off between the estimate performance and the computational complexity. Moreover, because the BMP algorithm cannot take advantage of the specific structure of the sparsity pattern, we made some changes to the support search procedure of BMP, which enable the algorithm to address the group-sparse channel coefficients in a more efficient manner. We showed that the modified BMP algorithm has a significantly reduced complexity compared with its original version. Furthermore, a pilot arrangement was designed to work in conjunction with the proposed estimation scheme. Experimental results illustrated the superiority of our method over conventional CS-based approaches.

References

[1] M. K. Ozdemir, H. Arslan, "Channel estimation for wireless OFDM systems," *IEEE Communications Surveys and Tutorials*, vol. 9, no. 2, pp. 18-48, 2007. Article (CrossRef Link)

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

[3] E. J. Candès, M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21-30, 2008. Article (CrossRef Link)

[4] W. U. Bajwa, J. Haupt, A. M. Sayeed, R. Nowak, "Compressed channel sensing: a new approach to estimating sparse multipath channels," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1058-1076, 2010. Article (CrossRef Link)

[5] J. Meng, W. Yin, Y. Li, N. T. Nguyen, Z. Han, "Compressive sensing based high-resolution channel estimation for OFDM system," *IEEE Journal of Selected Topics in Signal Processing*, vol. 6, no. 1, pp. 15-25, 2012. Article (CrossRef Link)

[6] G. Tauböck, F. Hlawatsch, D. Eiwen, H. Rauhut, "Compressive estimation of doubly selective channels in multicarrier systems: Leakage effects and sparsity-enhancing processing," *IEEE Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 255-271, 2010. Article (CrossRef Link)
[7] J. Huang, S. Zhou, J. Huang, C. Berger, P. Willett, "Progressive inter-carrier interference equalization for OFDM transmission over time-varying underwater acoustic channels," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 8, pp. 1524-1536, 2011. Article (CrossRef Link)
[8] C. R. Berger, Z. Wang, J. Huang, S. Zhou, "Application of compressive sensing to sparse channel estimation," *IEEE Communications Magazine*, vol. 48, no. 11, pp. 164-174, 2010. Article (CrossRef Link)
[9] S. Ji, Y. Xue, L. Carin, "Bayesian compressive sensing," *IEEE Transactions on Signal Processing*, vol. 56, no. 6, pp. 2346-2356, 2008. Article (CrossRef Link)
[10] D. Baron, S. Sarvotham, R. G. Baraniuk, "Bayesian compressive sensing via belief propagation," *IEEE Transactions on Signal Processing*, vol. 58, no. 1, pp. 269-280, 2010. Article (CrossRef Link)
[11] S. D. Babacan, R. Molina, A. K. Katsaggelos, "Bayesian compressive sensing using Laplace priors," *IEEE Transactions on Imaging Processing*, vol. 19, no. 1, pp. 53-63, 2010. Article (CrossRef Link)
[12] P. Schniter, L. C. Potter, J. Ziniel, "Fast Bayesian matching pursuit," *Information Theory and Applications Workshop*, pp. 326-333, 2008. Article (CrossRef Link)
[13] A. A. Quadeer, T.Y. Al-Naffouri, "Structure-based Bayesian sparse reconstruction," *IEEE Transactions on Signal Processing*, vol. 60, no. 12, pp. 6354-6367, 2012. Article (CrossRef Link)
[14] E. G. Larsson, Y. Selén "Linear regression with a sparse parameter vector," *IEEE Transactions on Signal Processing*, vol. 55, no. 2, pp. 451-460, 2007. Article (CrossRef Link)
[15] Y. C. Eldar, M. Mishalii, "Robust recovery of signals from a structured union of subspaces," *IEEE Transactions on Information Theory*, vol. 55, no. 11, pp. 5302-5316, 2009. Article (CrossRef Link)
[16] Y. C. Eldar, P. Kuppinger, H. Bolcskei, "Block-sparse signals: uncertainty relations and efficient recovery," *IEEE Transactions on Signal Processing*, vol. 58, no. 6, pp. 3042-3054, 2010. Article (CrossRef Link)
[17] E. J. Candès, J. Romberg, T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489-509, 2006. Article (CrossRef Link)
[18] G. Leus, P. A. van Walree, "Multiband OFDM for Covert Acoustic Communications," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 9, pp. 1662-1673, 2008. Article (CrossRef Link)
[19] D. L. Donoho, Y. Tsaig, "Fast solution of *l*1-norm minimization problems when the solution may be sparse," *IEEE Transactions on Information Theory*, vol. 54, no. 11, pp. 4789-4812, 2008. Article (CrossRef Link)
[20] T. Zhang, "Sparse recovery with orthogonal matching pursuit under RIP," *IEEE Transactions on Information Theory*, vol. 57, no. 9, pp. 6215-6221, 2011. Article (CrossRef Link)

Yi Liu received the B.S. degree in Communication and Electronic System from the Beijing Institute of Technology (BIT), Beijing, China, in 2007. From 2009, he begins to work toward the Ph.D. degree in the successive master-doctor program in BIT. He is currently also an assistant in the Department of Information Technology, College of National Defense Information, Wuhan, China. His research interests include compressive sensing and signal processing for communication.

Wenbo Mei received the B.S. degree in Electronic Engineering and M.S. Degree in Communication and Electronic System from BIT, China, in 1982 and 1993, respectively. He is professor of School of Information and Electronics, BIT. He is also a Visiting Research Fellow at University of Central Lancashire UK. His research interests include wavelet transform, time-frequency analysis, compressive sensing in signal and image processing for Radar, Communication system and channel, MRI, etc.

Huiqian Du received her B.S., M.S and Ph.D. degrees in Electrical Engineering from BIT, Beijing, China, in 1995, 1998 and 2007, respectively. In 1998, she joined Beijing Institute of Technology, where she is currently an associate professor. Her research interests lie in the areas of signal and image processing and include wavelet theory and compressive sensing.