Hindawi
Wireless Communications and Mobile Computing
Volume 2019, Article ID 6071672, 10 pages
https://doi.org/10.1155/2019/6071672

Research Article

Compressed Sensing-Based Sparsity Adaptive Doubly Selective Channel Estimation for Massive MIMO Systems

Weisi Kong, Hui Li, and Wenjie Zhang

School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710129, China

Correspondence should be addressed to Weisi Kong; kongsi0627@163.com

Received 19 January 2019; Revised 21 February 2019; Accepted 5 March 2019; Published 20 March 2019

Academic Editor: Juan F. Valenzuela-Valdés

By exploiting the sparsity of the channel in the delay and angle domains, compressed sensing (CS) algorithms can be used for channel estimation of massive multiple-input multiple-output (MIMO) systems to reduce pilot overhead. Due to the Doppler frequency shift, however, the intercarrier interference (ICI) and the rapid change of the channel state result in the poor estimation effect of doubly selective (DS) channel. In this paper, we propose the block sparsity adaptive matching pursuit (B-SAMP) algorithm to solve this problem. Firstly, the complex exponential basis expansion model (CE-BEM) is used to convert numerous channel tap coefficients into BEM parameter vectors and then the sparsity adaptive channel estimation scheme based on compressed sensing is proposed. Specifically, the ICI-free model is obtained by using the proposed equally placed pilot group scheme, and the B-SAMP algorithm is proposed by using the spatio-temporal common sparsity of the channel to complete the estimation of DS channel. Finally, a linear smoothing method is used to reduce the error caused by CE-BEM, thereby further improving the accuracy of the estimation. The simulation results show that the proposed method not only improves the estimation accuracy compared with the existing scheme but also requires fewer pilots.

1. Introduction

Mobile communications will be upgraded to a new level with the coming of 5G era. Massive multiple-input multiple-output (MIMO) technology as one of the key technologies in 5G mobile communications has many advantages such as improving system spectrum efficiency and transmission reliability [1]. Meanwhile, massive MIMO faces challenges such as high-dimensional channel state information (CSI), pilot pollution and the scheduling of numerous accessing users, etc. [2, 3]. For the doubly selective (DS) channel generated by the fast movement of the user, the Doppler shift is generated in the frequency domain, and the channel state changes rapidly [4]. In this case, the channel estimation becomes more complicated. Therefore, we mainly investigate the DS channel estimation problem of massive MIMO systems in this paper.

Compressed sensing (CS) technology can be used to reduce the rank of channel estimation by exploiting the sparse characteristics of the channel in some variation domains, which can reduce pilot overhead and estimation complexity in massive MIMO systems [5, 6]. The channel is sparse in the delay domain due to the limited number of significant channel gains in multipath propagation environments, and this sparsity is used to perform time domain channel estimation under time-domain synchronous orthogonal frequency division multiplexing (TDS-OFDM) transmission scheme in [7, 8]. In order to improve the accuracy of signal recovery and achieve channel estimation in complex scenarios, [9, 10] realizes time-frequency joint channel estimation by utilizing pseudonoise (PN) sequences and pilots. Time domain channel estimation can greatly reduce the overhead of the pilot. However, the interference cancellation of training sequences of different antennas will be difficult with the increase of the number of antennas. Therefore, [11] completes the channel estimation by proposing a nonorthogonal pilot scheme, and the adaptive structured subspace pursuit algorithm is proposed by utilizing the spatiotemporal common sparsity of the channel. The problem of multiuser channel estimation and feedback is further studied in [12, 13]. These papers improve the CS algorithm by using the sparse characteristics

of the channel under slow time-varying channels where the channel remains constant for a single or multiple OFDM symbol time.

The channel is sparse due to the nature of the transmission even under DS channels [14]. However, the Doppler shift caused by the rapid movement of the user leads to the rapid changes of the channel gain in the time domain and also destroys the orthogonality of the subcarriers to generate intercarrier interference (ICI) [4]. In this case, there are many channel coefficients that need to be estimated due to the rapid change of channel status. In addition, numerous pilots need to be placed in order to ensure that ICI can be observed, which will greatly reduce the spectral efficiency. Therefore, even using CS it is challenging to estimate the channel with sufficient accuracy under reasonable model [4].

Fortunately, the basis expansion model (BEM) can be used to simplify the estimation process because the channel changes are relatively smooth in the time domain [15]. The BEM expresses the channel gain of N discrete time points in a linear combination of Q basis functions. Especially the complex exponential basis expansion model (CE-BEM) can further make the frequency-domain channel matrix strictly banded and obtain the ICI-free model by combining specific pilot schemes [16]. After using CE-BEM, the parameters to be estimated for a single symbol of N_t antennas are greatly reduced from N_tNL to N_tQL, and the mathematical model of the frequency domain can be considered using the CS algorithm.

The combination of CE-BEM and CS algorithm can further reduce the pilot overhead of DS channel estimation and achieve good estimation effect. The single antenna channel estimation scheme based on time-frequency training OFDM is implemented in [17]. The model is further extended to MIMO systems and a novel compression recovery algorithm is proposed in [18]. The fast time-varying of the channel causes the error of symbol reconstruction in the time-frequency joint channel estimation scheme. Therefore, some studies directly use the pilot sequence to complete the frequency domain channel estimation. The distributed compressive sensing-simultaneous orthogonal matching pursuit (DCS-SOMP) algorithm is proposed for channel estimation of single antenna and symbol in [19], and the block simultaneous orthogonal matching pursuit (BSOMP) algorithm is proposed for channel estimation of multi symbols in [20]. The system model is further extended to large-scale MIMO systems and the optimal pilot placement scheme is proposed in [21].

The nonorthogonal pilot scheme in which the pilots of different antennas are placed at the same position is proposed in [11], which can greatly reduce the pilot overhead of the massive MIMO system compared to the conventional orthogonal pilot scheme. In addition, the scheme can also perform superimposed block processing of the same position in combination with the spatial common sparsity of the channel. In view of this, the nonorthogonal pilot group structure was improved, and the optimal placement scheme based on discrete stochastic optimization (DSO) is used to determine the pilot position in [21]. The DSO is the process of performing numerous random selections within a sample to find the optimal result. The goal of the optimal placement scheme is to obtain the position where the correlation coefficient of each column of the measurement matrix as small as possible, so as to make the measurement matrix approximately satisfy the restricted isometry property (RIP) criterion and improve the recovery accuracy.

In this paper, we estimate multiple channel state information corresponding to multiple OFDM symbols of all antennas at the same time. Firstly, the CE-BEM model is used to convert the channel tap gain into CE-BEM coefficients. Then the CE-BEM model is reconstructed by using the sparsity of the channel. The proposed block sparsity adaptive matching pursuit (B-SAMP) algorithm is used to estimate the BEM coefficients and obtain all channel coefficients. The following improvements are mainly implemented:

(i) The equally placed pilot group scheme is proposed, and a constant amplitude zero autocorrelation (CAZAC) sequence is used to determine the pilot amplitude.

(ii) The CE-BEM model is chunked by the spatiotemporal common sparsity of the channel, and the B-SAMP algorithm is used to obtain the accurately sparse support set.

The equally placed pilot group scheme is equally spaced to insert multiple pilot clusters on subcarriers, and each cluster has only one nonzero central pilot in the middle and surrounded by zero pilots on the left and right sides. In addition, the RIP criterion of measurement matrix can be satisfied by using the CAZAC sequence when the number of antennas is large [11]. The equally placed scheme is more convenient to implement in actual communication systems, and the computational complexity is lower.

The BSOMP algorithm is only used in a single antenna system in [20], and the algorithm requires that the sparsity is known. However, the proposed B-SAMP algorithm can be used in massive MIMO system and can achieve more accurate estimation by exploiting block sparse features. In addition, the B-SAMP algorithm does not require sparsity S to be known and can be more widely used in practice.

To further reduce the estimation error caused by CE-BEM, a piecewise linear approximation model is used to smooth the estimated channel taps. It has been proved in [22] that the piecewise linear approximation model has a good estimate of DS channel even for normalized Doppler of up to 0.2, and this model has been improved for smoothing treatment for single antenna system in [20]. The simulation part shows that smoothing treatment can greatly improve the estimation effect.

The rest of this paper is organized as follows. The system model is presented in Section 2. The proposed estimation scheme and algorithm are in Section 3. Section 4 holds the smoothing treatment and complexity analysis of algorithm. In Section 5, simulation results are presented to demonstrate the performance of the proposed scheme. Finally, conclusion is drawn in Section 6.

2. System Model

In this section, we first describe the basic model of OFDM system under the DS channel in massive MIMO. Then, the CE-BEM form of the channel and the basic criteria of the CS algorithm are introduced.

2.1. OFDM System Model over DS Channel. In massive MIMO systems, base stations usually employ a large number of antennas, which exchange data with terminals of single antennas to achieve complex communication. It is assumed that there are N_t antennas at the sender and $\mathbf{s}_r^{(k)}$ is the r-th OFDM symbol transmitted by the k-th antenna. The subcarrier sequence $\mathbf{X}_r^{(k)} = [X_r^{(k)}[0], \cdots X_r^{(k)}[\mathrm{N}-1]]^T$ of length N is composed of the pilot subcarrier group $[\mathbf{X}_r^{(k)}]_\Omega$ and the data subcarrier $[\mathbf{X}_r^{(k)}]_\sigma$, in which $\Omega(|\Omega| = P)$ is the set of pilot subcarrier indices and $\sigma(|\sigma| = N - P)$ represents the set of data carrier indices. After inverse discrete Fourier transform (IDFT), the time-domain symbol sequence $\mathbf{x}_r^{(k)} = \mathbf{F}_N^H \mathbf{X}_r^{(k)}$ is obtained ($\mathbf{F}_N^H$ is an N-point IDFT matrix). Then the OFDM symbol $\mathbf{s}_r^{(k)} \in \mathbb{C}^{(M+N)\times 1}$ is precoded by a cyclic prefix (CP) sequence with length M.

The OFDM symbols transmitted by N_t antennas reach the receiver through a complex DS channel. The discrete mathematical model is expressed as

$$y_r[\mathrm{n}] = \sum_{k=1}^{N_t} \sum_{l=1}^{L} h_r^{(k)}[n,l]\, s_r^{(k)}[n-l] + w_r[n] \tag{1}$$

where l is discrete delay samples and L is the equivalent channel length, $h_r^{(k)}[n,l] = 0$ when $n \notin [0, M+N-1]$ or $l \notin [1, L]$. Due to the large attenuation in the transmission process, the path with significant gain becomes less. Therefore, channels exhibit the sparsity in the delay domain with several relatively large tap coefficients, $\ell_{k,r} = \mathrm{supp}\{\mathbf{h}_r^{(k)}\} = \{l : |\mathbf{h}_r^{(k)}[l]| > p_{th}\}$ is the support set of $\mathbf{h}_r^{(k)}$ and p_{th} is the noise floor according to [23]. After removing the CP and DFT transform, the frequency domain subcarrier sequences are obtained:

$$\mathbf{y}_r = \sum_{k=1}^{N_t} \mathbf{H}_{\mathrm{F},r}^{(k)} \mathbf{X}_r^{(k)} + \mathbf{w}_r \tag{2}$$

where the frequency domain channel matrix $\mathbf{H}_{\mathrm{F},r}^{(k)} = \mathbf{F}_N \mathbf{H}_{\mathrm{T},r}^{(k)} \mathbf{F}_N^H$, $\mathbf{H}_{\mathrm{T},r}^{(k)} \in \mathbb{C}^{N\times N}$ is the time domain channel matrix, and $\mathbf{w}_r$ is the corresponding additive white Gaussian noise (AWGN) vector. The elements for $\mathbf{H}_{\mathrm{T},r}^{(k)}$ can be expressed as

$$\left[\mathbf{H}_{\mathrm{T},r}^{(k)}\right]_{p,q} = h_r^{(k)}\left[M + p, \mathrm{mod}\left(p-q, N\right)\right], \quad p,q \in [0, N-1] \tag{3}$$

It is worth noting that $\mathbf{H}_{\mathrm{T},r}^{(k)}$ is a cyclic matrix and $\mathbf{H}_{\mathrm{F},r}^{(k)}$ is a diagonal matrix in the slow fading channel. However, $\mathbf{H}_{\mathrm{T},r}^{(k)}$ is a pseudocyclic matrix and $\mathbf{H}_{\mathrm{F},r}^{(k)}$ becomes a complete matrix due to the Doppler shift inducing ICI in DS channel. In the channel estimation process, there are N_tNL channel coefficients that need to be estimated due to the rapid change of channel status. In this case, a large number of pilots need to be placed in order to ensure that ICI can be observed, which will greatly reduce the spectral efficiency. Therefore, even using compressive sensing it is challenging to estimate the channel with sufficient accuracy under reasonable model [4].

2.2. CE-BEM Based System Model. The variation of $h_r^{(k)}[n,l]$ with time n is usually very smooth because the Doppler spread of the channel is limited, so it can be approximated by a small part smooth function. Using the BEM to fit the channel coefficients can greatly reduce the computational complexity of the estimation. Specifically

$$\mathbf{h}_{r,l}^{(k)} = \sum_{q=0}^{Q-1} c_{r,q}^{(k)}[l] \bullet \mathbf{b}_q + \boldsymbol{\xi}_{r,l}^{(k)} \tag{4}$$

in which $\boldsymbol{\xi}_{r,l}^{(k)} \in \mathbb{C}^{N\times 1}$ is the BEM error, $Q(Q \ll N)$ is the BEM order, $\mathbf{b}_q \in \mathbb{C}^{N\times 1}$ is q-th BEM function, and $c_{r,q}^{(k)}[l]$ is the corresponding BEM coefficient with the tap l. Besides, $\mathbf{h}_{r,l}^{(k)} = (h_r^{(k)}[M,l], \cdots, h_r^{(k)}[M+N-1,l])^T$ is the impulse response of the N points of the l-th channel tap. Using the BEM can significantly reduce the number of unknown coefficients from N_tNL to N_tQL.

Although the gain of channel changes rapidly in DS channel, the delay of channel changes slowly [24]. Therefore, it can be ensured that S significant channel taps keep the corresponding position $\ell_k = \{l_1, \cdots l_S\}$ unchanged during R successive OFDM symbols [20]. The R consecutive OFDM symbols transmitted by N_t antennas can be considered to pass through the same multipath channel and share common sparsity under the condition of $d_{\max}/c \le 1/10BW$ [21], in which $d_{\max}$ denotes the maximum distance between any two transmit antennas, c is the speed of light, and BW is the signal bandwidth. It should be pointed out that the path gains of different transmit receive antenna pairs can be different due to the nonisotropic antennas [25].

We can get that $\mathbf{c}_{r,q}^{(k)} = [c_{r,q}^{(k)}[1], \cdots, c_{r,q}^{(k)}[\mathrm{L}]]^T$ has the same set of support ℓ and sparsity S for any $q \in [0, Q-1]$, $r \in [1,R]$, and $k \in [1, N_t]$ by exploiting the spatial and temporal channel correlations. Combining the formula (3) for analysis and derivation, we can obtain

$$\mathbf{H}_{T,r}^{(k)} = \sum_{q=0}^{Q-1} \mathrm{diag}\left(\mathbf{b}_q\right) \widetilde{\mathbf{C}}_{r,q}^{(k)} + \mathbf{E} \tag{5}$$

where $\mathbf{E}$ is the error matrix of the BEM, $\mathrm{diag}(\mathbf{b}_q)$ is the diagonal matrix with the diagonal element $\mathbf{b}_q$, and $\widetilde{\mathbf{C}}_{r,q}^{(k)}$ is the cyclic matrix with the first column $\mathbf{c}_{r,q}^{(k)} \in \mathbb{C}^L$. The BEM error will be omitted in the following calculation, and the estimated $\overline{\mathbf{H}}_T$ will be smoothed to further reduce the impact of errors in Section 5. Equation (2) can be simplified as

$$\mathbf{y}_r = \sum_{k=1}^{N_t} \sum_{q=0}^{Q-1} \mathbf{B}_q \mathbf{C}_{r,q}^{(k)} \mathbf{X}_r^{(k)} + \mathbf{w}_r \tag{6}$$

in which $\mathbf{B}_q = \mathbf{F_N}\text{diag}(\mathbf{b}_q)\mathbf{F}_N^H$ and $\mathbf{C}_{r,q}^{(k)} = \text{diag}(\mathbf{F_N}((\mathbf{c}_{r,q}^{(k)})^T, 0_{1\times(N-L)})^T)$. The simplification of the formula takes advantage of the diagonalization of the circulant matrix.

The basis functions of CE-BEM are Fourier vector with a period of N, which is able to make the frequency-domain channel matrix $\mathbf{H}_{F,r}^{(k)}$ strictly banded [19]. The q-th basis function $\mathbf{b}_q$ can be expressed as

$$\mathbf{b}_q = \Big(1, \cdots, e^{j(2\pi/N)n(q-(Q-1)/2)}, \cdots e^{j(2\pi/N)(N-1)(q-(Q-1)/2)}\Big)^T \tag{7}$$

Using the expression of $\mathbf{b}_q$, the formula (6) can be further simplified as

$$\mathbf{y}_r = \sum_{k=1}^{N_t}\sum_{q=0}^{Q-1}\mathbf{I}_N^{\langle q-(Q-1)/2\rangle}\text{diag}\left(\mathbf{X}^{(k)}\right)\mathbf{F}'_N\mathbf{c}_{r,q}^{(k)} + \mathbf{w}_r \tag{8}$$

where $\mathbf{I}_N^{\langle q\rangle}$ denotes a permutation matrix obtained from an identity matrix $\mathbf{I}_N$ and $\mathbf{F}'_N \in \mathbb{C}^{N\times L}$ is the first L column of the DFT matrix $\mathbf{F}_N$. In the next, the specific BEM coefficient expression will be obtained by the corresponding pilot group structure.

2.3. The Basic Criteria of CS Algorithm. The high-dimensional signal can be successfully recovered when two important criteria in CS are satisfied: the high dimensional signal $\mathbf{x}$ is a sparse vector (i.e., $\|\mathbf{x}\|_0 = S$ and $S \ll L$) and the measurement matrix $\mathbf{\Phi}$ needs to meet the RIP criterion. However, it is difficult to verify RIP condition due to the prohibitive complexity and the tremendous computation. The mutual coherence property (MCP) is widely adopted in the practical schemes because it is simple to calculate [26]. It can be expressed as

$$\mu(\mathbf{\Phi}) = \max_{1\le i\ne j\le L}\frac{\left|\left\langle\phi_i, \phi_j\right\rangle\right|}{\|\phi_i\|_2\|\phi_j\|_2} \tag{9}$$

where ϕ_i, ϕ_j denote the i-th and the j-th columns of $\mathbf{\Phi}$. $\mu(\mathbf{\Phi})$ is used to represent the self-correlation of the measurement matrix, and a smaller value of the $\mu(\mathbf{\Phi})$ means the orthogonality of the columns of the measurement matrix gets better, so that the more accurately support set can be selected in the loop iteration of the CS algorithm.

3. The Proposed Channel Estimation Scheme

Our goal is to obtain channel state information under the DS channel, and the compressed sensing algorithm used due to the pilot overhead can be reduced. The equally spaced pilot group scheme will be mentioned firstly, which can obtain the CE-BEM vector model without ICI and can satisfy the MCP criterion of the measurement matrix. Then, the B-SAMP algorithm based on spatiotemporal common sparsity will be proposed, which has good sparse adaptability and estimation stability.

3.1. The Equally Placed Pilot Group Scheme. The nonorthogonal pilot scheme is used in the paper, which is that the pilots of different transmit antennas are at the same subcarrier position. The nonorthogonal pilot scheme can reduce the pilot overhead compared with the traditional orthogonal pilot scheme and can perform the joint channel estimation of multiple antennas by using the sparsity of the channel and the compressed sensing algorithm.

The total number of pilot subcarriers and the corresponding pilot indices within one OFDM symbol in the k-th antenna are denoted as P and Ω, respectively. The pilot subcarriers are grouped in G clusters, and each cluster has only one non-zero central pilot in the middle and surrounded by $Q-1$ zero pilots on the left and right sides. The central pilot index set $\Omega_{cent}(|\Omega_{cent}| = G)$ is expressed as

$$\Omega_{cent} = \{p_0, \cdots, p_{G-1}\} \tag{10}$$

where $0 \le p_0 < \cdots < p_{G-1} \le N-1$ and must have $|p_i - p_j| \ge 2Q-1$, $i \ne j$ to prevent the overlapping of the effective pilot subcarriers and the guard pilot subcarriers. In the next, we take Q subsets from Ω with

$$\begin{aligned}\Omega_0 &= \Omega_{cent} - \frac{Q-1}{2}\\ &\vdots\\ \Omega_{(Q-1)/2} &= \Omega_{cent}\\ &\vdots\\ \Omega_{Q-1} &= \Omega_{cent} + \frac{Q-1}{2}\end{aligned} \tag{11}$$

in which $\Omega_{cent} - (Q-1)/2$ denotes a set where each element in Ω_{cent} subtracts $(Q-1)/2$. The pilot pattern with $Q = 3$ is shown in Figure 1, and the value of Q is related to the Doppler shift. The Doppler shift characteristic can be expressed by using the pilot pattern and the CE-BEM, and the estimation of Q sparse CE-BEM coefficient vectors could be decoupled from (8) by Q separate equations without ICI as

$$\begin{aligned}[\mathbf{y}_r]_{\Omega_0} &= \sum_{k=1}^{N_t}\text{diag}\left(\mathbf{P}_{\Omega_{cent}}^{(k)}\right)\left[\mathbf{F}'_N\right]_{\Omega_{cent}}\mathbf{c}_{r,0}^{(k)} + \mathbf{w}_{r,0}\\ &\vdots\\ [\mathbf{y}_r]_{\Omega_{(Q-1)/2}} &= \sum_{k=1}^{N_t}\text{diag}\left(\mathbf{P}_{\Omega_{cent}}^{(k)}\right)\left[\mathbf{F}'_N\right]_{\Omega_{cent}}\mathbf{c}_{r,(Q-1)/2}^{(k)} + \mathbf{w}_{r,(Q-1)/2}\\ &\vdots\\ [\mathbf{y}_r]_{\Omega_{Q-1}} &= \sum_{k=1}^{N_t}\text{diag}\left(\mathbf{P}_{\Omega_{cent}}^{(k)}\right)\left[\mathbf{F}'_N\right]_{\Omega_{cent}}\mathbf{c}_{r,Q-1}^{(k)} + \mathbf{w}_{r,Q-1}\end{aligned} \tag{12}$$

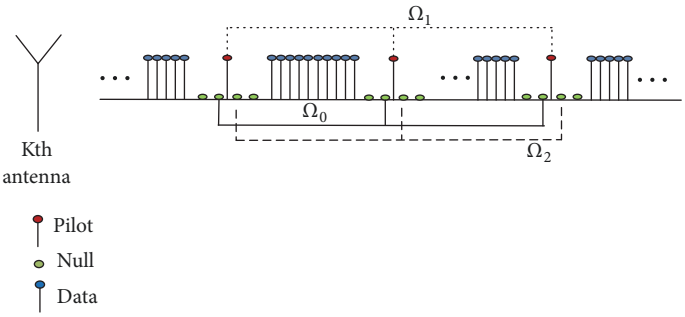

Figure 1: The pilot group pattern within single symbol of k-th antenna.

The above formula is an undetermined condition due to $G \ll N_t L$, which is impossible to estimate the signal using the traditional estimation algorithm. The high-dimensional signal can be successfully recovered when two important criteria in CS are satisfied. We have obtained that the BEM coefficient $\mathbf{c}_{r,q}^{(k)}$ is sparse. Then we will construct the measurement matrix $\mathbf{\Phi}$, and channel estimation will be performed by using an improved CS based on the spatiotemporal common sparsity of channel.

Considering the correlation of channels over R consecutive OFDM symbols, (12) can be expressed as

$$\begin{aligned}
[\mathbf{Y}]_{\Omega_0} &= \left[\mathbf{\Phi}^{(1)} \cdots \mathbf{\Phi}^{(N_t)}\right] \mathbf{C}_0 + \mathbf{W}_0 \\
&\vdots \\
[\mathbf{Y}]_{\Omega_{(Q-1)/2}} &= \left[\mathbf{\Phi}^{(1)} \cdots \mathbf{\Phi}^{(N_t)}\right] \mathbf{C}_{(Q-1)/2} + \mathbf{W}_{(Q-1)/2} \\
&\vdots \\
[\mathbf{Y}]_{\Omega_{Q-1}} &= \left[\mathbf{\Phi}^{(1)} \cdots \mathbf{\Phi}^{(N_t)}\right] \mathbf{C}_{Q-1} + \mathbf{W}_{Q-1}
\end{aligned} \tag{13}$$

where $\mathbf{Y} = [\mathbf{y}_1 \cdots \mathbf{y}_R] \in \mathbb{C}^{N \times R}$ is the received subcarriers during multiple OFDM symbols, $\mathbf{C}_q = [\mathbf{c}_{1,q} \cdots \mathbf{c}_{\mathrm{R},q}] \in \mathbb{C}^{N_t L \times R}$ represents the BEM coefficients that needs to be estimated, and $\mathbf{c}_{r,q} = [(\mathbf{c}_{r,\mathrm{q}}^{(1)})^T \cdots (\mathbf{c}_{r,\mathrm{q}}^{(N_t)})^T]^T$. The measurement matrix is $\mathbf{\Phi} = [\mathbf{\Phi}^{(1)} \cdots \mathbf{\Phi}^{(N_t)}] \in \mathbb{C}^{R \times N_t L}$, and $\mathbf{\Phi}^{(k)} = \mathrm{diag}(\mathbf{P}_{\Omega_{cent}}^{(k)})[\mathbf{F}'_N]_{\Omega_{cent}}$. Since Q equations have the same measurement matrix and sparsity, (12) can be further expressed as

$$\left[[\mathbf{Y}]_{\Omega_0} \cdots [\mathbf{Y}]_{\Omega_{Q-1}}\right] = \mathbf{\Phi}\left[\mathbf{C}_0 \cdots \mathbf{C}_{Q-1}\right] + \mathbf{W} \tag{14}$$

The above formula is the multiobservation vector representation form of DCS theory. Estimation accuracy can be improved by using DCS with multidimensional observed vectors [19]. It can be found that the size $\mathbf{P}_{\Omega_{cent}}^{(k)}$ and position Ω_{cent} of the pilot determine the value of $\mu(\mathbf{\Phi})$ because of $\mathbf{\Phi}^{(k)} = \mathrm{diag}(\mathbf{P}_{\Omega_{cent}}^{(k)})[\mathbf{F}'_N]_{\Omega_{cent}}$.

The pilots of different antennas occupy the identical subcarriers position, and the pilot groups are equally spaced on each antenna. The pilot subcarriers of each antenna are grouped in G clusters, and each cluster has only one central pilot p_κ in the middle and surrounded by $Q-1$ zero pilots on the left and right sides. Specifically, $\{p_\kappa\}_{\kappa=0}^{G-1}$ is selected from the set of $\{0+(\mathrm{Q}-1), \cdots, N-1-(\mathrm{Q}-1)\}$ with the equal interval $\lfloor N/G \rfloor$ for all antennas, and $p_\kappa = p_0 + \kappa\lfloor N/G \rfloor$ for $0 \le \kappa \le G-1$, where p_0 is the subcarrier index of the first pilot with $Q-1 \le p_0 < \lfloor N/G \rfloor - Q$. In addition, we consider $\{\theta_{\kappa,m}\}_{\kappa=0,k=1}^{G-1,N_t}$ to follow the independent and identically distributed uniform distribution in $[0, 2\pi]$, where $e^{j\theta_{\kappa,k}}$ denotes the $\kappa+1$ th element of $\mathbf{P}_{\Omega_{cent}}^{(k)} \in \mathbb{C}^{G \times 1}$ for the k-th antenna.

It has been proved in [11] that smaller value of $\mu(\mathbf{\Phi})$ can be obtained using this pilot scheme with the increase of the number of antennas. Compared to the optimal pilot placement scheme [21], the equally interval placement scheme has lower computational complexity and can be applied more widely in practical communication.

3.2. The B-SAMP Algorithm. In previous section, we have obtained the DCS form by exploiting the common sparsity of the time and BEM functions. However, the spatial common sparsity can be used to further improve the accuracy of the estimation. Specifically, the measurement matrix $\mathbf{\Phi}$ can be rearranged as

$$\mathbf{\Psi} = \left[\mathbf{\Psi}_1, \cdots, \mathbf{\Psi}_L\right] \tag{15}$$

where $\mathbf{\Psi}_l = [\mathbf{\Phi}_l^{(1)}, \cdots, \mathbf{\Phi}_l^{(N_t)}] \in \mathbb{C}^{G \times N_t}$. So (13) can be reformulated as

$$\mathbf{Y} = \mathbf{\Psi}\mathbf{D} + \mathbf{W} \tag{16}$$

where $\mathbf{D} = [\mathbf{D}_0^T \cdots \mathbf{D}_{L-1}^T]^T$, $\mathbf{D}_l = [\mathbf{d}_{1,0}[l] \cdots \mathbf{d}_{\mathrm{R},0}[l] \cdots \mathbf{d}_{1,\mathrm{Q}-1}[l] \cdots \mathbf{d}_{\mathrm{R},\mathrm{Q}-1}[l]] \in \mathbb{C}^{N_t \times RQ}$, and $\mathbf{d}_{r,q}[l] = [\mathbf{c}_{r,\mathrm{q}}^{(1)}[l], \cdots \mathbf{c}_{r,\mathrm{q}}^{(N_t)}[l]]^T \in \mathbb{C}^{N_t \times 1}$. The model becomes block-sparse form after reorganization. Our goal is to find the support set ℓ with sparsity S in $\mathbf{D}_l$ $(0 \le l \le L-1)$.

We propose the B-SAMP algorithm by improving the original SAMP algorithm in order to obtain the sparse position of the channel more accurately. Block sparsity can greatly improve the accuracy of recovery by using Q BEM equations and spatiotemporal common sparsity. In the B-SAMP algorithm as shown in Algorithm 1, the initialization parameters are set at first and then iterate through the loop to update the support set. When the stop criterion is satisfied,

Input: sensing matrix $\mathbf{\Psi}$, observed matrix $\mathbf{Y}$, step size s;
Output: S-sparse block approximation $\widehat{\mathbf{D}}$
Initialization: the iterative index $i = 1$, the support set $F_0 = \phi$, the residual matrix $\mathbf{R}_{i-1} = \mathbf{Y}$, the initial channel sparsity level $T = s$ and the stage index $j = 1$
Repeat
(1) (correlation) $\mathbf{Z} = \mathbf{\Psi}^H \mathbf{R}_{i-1}$
(2) (support obtain) $C_i = F_{i-1} \cup \prod^T (\{\|\mathbf{Z}_l\|_F\}_{l=1}^L)$
(3) (support pruning) $\check{\mathbf{D}}_{C_i} = \mathbf{\Psi}_{C_i}^{\dagger} \mathbf{Y}, \check{\mathbf{D}}_{(C_i)^c} = 0$

$$F = \prod^T \left(\left\{ \left\| \check{\mathbf{D}}_l \right\|_F \right\}_{l=1}^L \right)$$

(4) (block estimate) $\check{\mathbf{D}}_F = \mathbf{\Psi}_F^{\dagger} \mathbf{Y}, \check{\mathbf{D}}_{(F)^c} = 0$
(5) (residue update) $\mathbf{R}_i = \mathbf{Y} - \mathbf{\Psi} \check{\mathbf{D}}$
If $\|\mathbf{R}_i\|_F < \|\mathbf{R}_{i-1}\|_F$
(6) (Continue loop iteration)
$F_i = F$
$i = i + 1$
Else
(7) (stage index update) $j = j + 1$
(8) (sparsity level update) $T = j \times s$
Until halting condition ture
Output: $\widehat{\mathbf{D}} = \check{\mathbf{D}}$ and obtain the estimation of channel $\mathbf{h}_{r,l}^{(k)}$

Algorithm 1: The B-SAMP Algorithm.

the estimated $\check{\mathbf{D}}$ is obtained and $\mathbf{h}_{r,l}^{(k)} \in \mathbb{C}^{N \times 1}$ can be obtained according to (4)-(14) and (16).

The purpose of step (1)-(6) of the loop iteration is to obtain a support set with sparsity level S. The step (1) is to obtain the correlation between each column of the measurement matrix and the residual. Then select the most relevant T-block, get its position, and update the support set in step (2). $\prod^T (\{\|\mathbf{Z}_l\|_F\}_{l=1}^L)$ denotes a set of size T, which finds the position of the T largest number by calculating the F-norm of the L block. A more accurate support set is obtained through inversion and support pruning.

The residual is updated at last, and then proceed to the next iteration of loop. Steps (7)-(8) are performed when $\|\mathbf{R}_i\|_F \geq \|\mathbf{R}_{i-1}\|_F$, indicating that the current js-sparse solution has been obtained, and the sparsity level needs to be updated to find $(j+1)s$-sparse solution. The step size s determines the number of supports selected for each cycle.

In the algorithm, stop iterating when $\|\mathbf{R}_i\|_F > \|\mathbf{R}_{i-2}\|_F$ $(i \geq 2)$ or $\|\check{\mathbf{D}}_{l_{\min}}\|_F \leq p_{th}$ is satisfied. The sparsity level T is increased to reduce residuals when $\|\mathbf{R}_i\|_F > \|\mathbf{R}_{i-1}\|_F$. However, if the residuals continue to increase, it must stop iterating immediately to ensure the accuracy of the estimate. $\|\check{\mathbf{D}}_{l_{\min}}\|_F$ denotes the smallest $\|\check{\mathbf{D}}_l\|_F$ for $l \in [0, L-1]$, and its value will become smaller and smaller with repeated iterations. When a complete support set is obtained, $\|\check{\mathbf{D}}_{l_{\min}}\|_F$ implies that the l-th path is dominated by noise. Therefore, we can set the appropriate threshold p_{th} based on the size of the noise to achieve a good estimation effect.

The MCP criterion is the basic criterion of the CS algorithm and is of great significance for measuring the performance of the algorithm [26]. According to the previous analysis, the equal-interval pilot group scheme can obtain lower cross-correlation values, thus ensuring the validity of the loop iteration of the B-SAMP algorithm [11]. As the loop iterates continually, the correct sparse support set is gradually selected and the residual $\mathbf{R}_i$ is reduced. It is worth noting that the size of the threshold p_{th} is mainly determined by the noise and the fitting error caused by CE-BEM, so it should be continuously tested and optimized during application [11, 16]. The construction of the measurement matrix and the setting of the stopping criterion ensure the validity and convergence of the algorithm, so that accurate estimation results can be obtained.

Simulation experiments show that the B-SAMP algorithm achieves similar results to the DCS-SOMP algorithm by setting appropriate stopping criteria. It should be pointed out that the DCS-SOMP algorithm is block-processed as the B-SAMP algorithm, instead of being originally applied to a single-symbol and single-antenna system in [19]. The DCS-SOMP algorithm is not available in the real communication process due to the sparsity S needs to be known. Therefore, the B-SAMP algorithm has a wider application significance. Comparison of specific algorithms will be compared in Section 5.

4. Smoothing Treatment and Complexity Analysis

4.1. Smoothing Treatment. The CE-BEM can greatly reduce the estimated parameters under DS channels. However, this model has a serious Gibbs effect at the edge of the detection window and causes a large error due to spectral leakage.

It has been proved that the piecewise linear approximation model can be used in DS channel estimation, which can achieve good estimation effect even if the normalized Doppler shift (NDS) reaches 0.2 [20, 22]. We define $\mathrm{h}_{r,l}^{(k)ave} \triangleq E[\mathbf{h}_{r,l}^{(k)}[M], \cdots, \mathbf{h}_{r,l}^{(k)}[M+N-1]]$ to denote the time average of the already obtained the l-th CIR during the r-th OFDM system of the k-th antenna. It has been found that $|\mathrm{h}_{r,l}^{(k)ave} - \mathrm{h}_{r,l}^{(k)}[M+n]|$ is minimized with $n = N/2 - 1$ in [22]. So we can use the estimated $\mathrm{h}_{r,l}^{(k)ave}$ to get $\mathrm{h}_{r,l}^{(k)}[M + N/2 - 1]$ with a small error for $r \in [1, R]$ and then use a piecewise linear approximation model for smoothing treatment.

Specifically, the estimate of the slope between the $(r-1)$-th and the r-th OFDM symbol can be obtained

$$\alpha_{r-1,l}^{(k)} = \frac{\mathrm{h}_{r,l}^{(k)ave} - \mathrm{h}_{r-1,l}^{(k)ave}}{N+M}, \quad l \in \widehat{\ell},\ k \in [1, N_t] \tag{17}$$

Because of the existence of the CP, the discrete interval is $N + M$. Similarly, the slope between the r-th and the $(r+1)$-th OFDM symbol can be obtained

$$\alpha_{r,l}^{(k)} = \frac{\mathrm{h}_{r+1,l}^{(k)ave} - \mathrm{h}_{r,l}^{(k)ave}}{N+M}, \quad l \in \widehat{\ell},\ k \in [1, N_t] \tag{18}$$

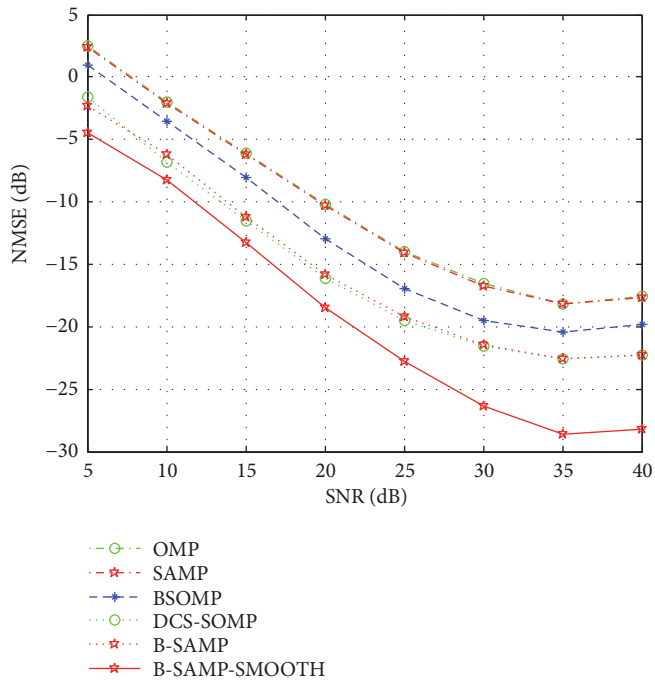

FIGURE 2: Comparison of NMSE performance between the proposed estimation algorithm and other algorithms at speed $350km/h$ (NDS $v_d = 0.065$).

We can utilize both $\alpha_{r-1,l}^{(k)}$ and $\alpha_{r,l}^{(k)}$ to calculate the CIR of the r-th OFDM.

$$h_{r1,l}^{(k)}(n) = \left(M + n + 1 + \frac{N}{2}\right)\alpha_{r-1,l}^{(k)} + h_{r-1,l}^{(k)ave}$$

$$h_{r2,l}^{(k)}(n) = \left(n + 1 - \frac{N}{2}\right)\alpha_{r,l}^{(k)} + h_{r,l}^{(k)ave}, \quad (19)$$

$$0 \le n \le N - 1$$

Then by calculating the average of $\mathbf{h}_{r1,l}^{(k)}$ and $\mathbf{h}_{r2,l}^{(k)}$, we obtain more accurate CIR via smoothing treatment, represented as

$$\mathbf{h}_{r,l}^{(k)} = \frac{1}{2}\left(\mathbf{h}_{r1,l}^{(k)} + \mathbf{h}_{r1,l}^{(k)}\right) \quad l \in [0, \mathrm{L} - 1], \; k \in [1, N_t] \quad (20)$$

After smoothing treatment, the estimation error will be greatly reduced, and the effect of smoothing will be verified in Section 5.

4.2. Complexity Analysis. The iteration of steps (1)-(5) constitutes the main part of each cycle in the B-SAMP algorithm. For step (1), the complexity of finding the correlation of the measurement matrix is $o(GN_tLQR)$. For step (2), the complexity of the F-norm of the matrix block $\mathbf{Z}_l$ and the select T largest correlations $\prod^T$ is $o(N_tLQR)$ and $o(L)$, respectively. The complexity of Moore-Penrose matrix inversion is $o(2G(N_tT)^2 + (N_tT)^3)$ in step (3), and the complexity of getting F is $o(N_tLQR)$ and $o(L)$. Similarly, steps (4) and (5) have the complexity of $o(2G(N_tT)^2 + (N_tT)^3)$ and $o(GN_tLQR)$, respectively. After simple analysis of complexity, it can be found that the main computational load lies in the inverse of the matrix.

5. Simulation Results and Discussion

In this section, the proposed channel estimation scheme will be validated by MATLAB simulation. We mainly compare the B-SAMP algorithm with the DCS-SOMP algorithm to verify the accuracy of the estimation and highlight the self-adaptiveness of the algorithm. In addition, the effects of equally spaced pilot placement scheme and the smoothing method will also be validated.

The important parameters in the simulation system are as follows: the length of subcarriers was $N = 4096$, the lengths of channel and CP are $L = 165$ and $M = 170$, system carrier was $f_c = 3GHz$, subcarrier spacing $\Delta f = 15KHz$, and system bandwidth $f_s = 61.44MHz$. The DS channel is generated by Jakes simulation model and the power spectrum obeys U-shaped distribution. The International Telecommunications Union Vehicular-B channel model with $S = 6$ was adopted in simulation. Assume that the uniform array of antennas is arranged at the pitch of $\lambda/2$. We set $N_t = 15$ to ensure that all antennas have common sparsity according to [21].

5.1. NMSE Comparison of Different Algorithms. The normalized mean square error (NMSE) comparisons are performed in scenarios where the speeds are $350km/h$ and $500km/h$, respectively. It is known by [20] that the size of Q and R is affected by the speed, so we can set $Q = 3$ and $R = 5$ to meet the requirements.

The B-SAMP algorithm is compared with the original SAMP algorithm and OMP algorithm. Meanwhile, the algorithm compares with the BSOMP algorithm and the DCS-SOMP algorithm. Figure 2 shows the comparison of NMSE performance for different algorithms at the speed of $350km/h$

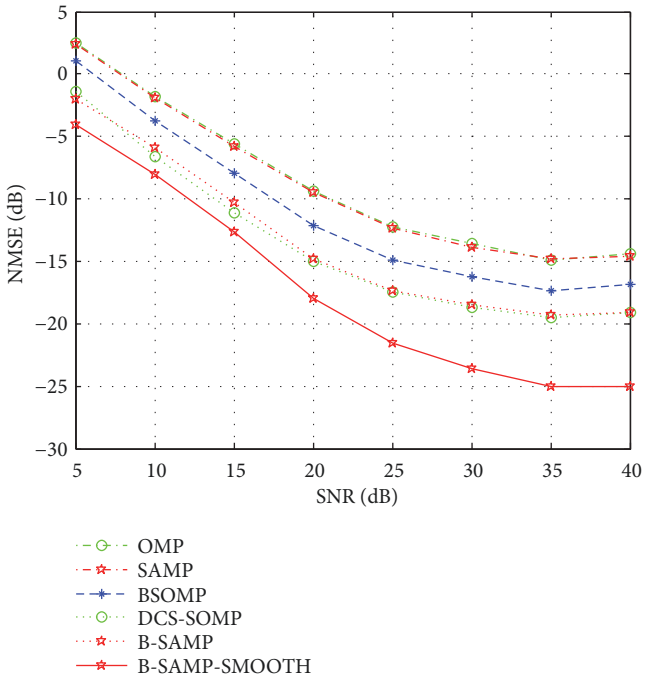

Figure 3: Comparison of NMSE performance between the proposed estimation algorithm and other algorithms at speed $500km/h$ (NDS $v_d = 0.093$).

(NDS $v_d = 0.065$). It can be seen from the figure that the NMSE of different algorithms is gradually reduced with the increase of signal to noise ratio. The OMP algorithm has the same recovery effect as the SAMP algorithm, and the NMSEs are the largest. The BSOMP algorithm utilizes the common sparsity of multiple OFDM symbols and BEM coefficients to change the position selection of single element into vector and uses this superposition form to improve the selection accuracy of the support set. The DCS-SOMP algorithm and the B-SAMP algorithm further reduce the NMSE by exploiting the common sparsity of the antennas of the massive MIMO system. In addition, it can be found that, after smoothing treatment, the NMSE is greatly reduced and a very good estimation effect is achieved.

In order to strengthen the contrast, the comparison at the speed of $500km/h$ (NDS $v_d = 0.093$) is shown in Figure 3. The recovery effect similar to the NDS of 0.065 is shown in the figure, except that the NMSE of all algorithms is improved by 3 dB due to the increase of Doppler shift. After these comparisons, it can be found that the B-SAMP algorithm can reduce the error by about 5 dB compared with the original SAMP algorithm using the block search support set, and the smooth processing can further reduce the NMSE by 5 dB.

After this part, we can conclude that the B-SAMP algorithm using the spatiotemporal common sparsity of the channel to improve the original SAMP algorithm can greatly improve the estimation accuracy, and the algorithm has good adaptive adjustment capability because it achieves the same effect as the DCS-SOMP algorithm with known sparsity. In addition, the smoothing process using the piecewise linear approximation model further improves the estimation accuracy.

5.2. NMSE Comparison of Pilots Placement Schemes. The equally spaced pilot scheme was used in this paper. In order to prove that the equally spaced placement can achieve good estimation results, we compare the equal interval placement scheme with the optimal placement scheme based on DSO.

The optimal placement scheme is to find a location index that minimizes the $\mu(\mathbf{\Phi})$ value, which can makes the measurement matrix have good orthogonality, and the estimated support set is very accurate. Figure 4 compares the NMSE performance of two placement schemes, where the proposed B-SAMP algorithm and the smoothed B-SAMP algorithm are used. We can find that the two pilot schemes have very similar performance.

It is mentioned that the equispaced pattern reaches $\mu(\mathbf{\Phi}) = 0.99$, and the estimation effect is extremely poor in [21], because the number of pilot sequence samples is too small to make the measurement matrix have good orthogonality under single antenna system. In this paper, the distribution of the phase $\theta_{\kappa,m}$ is more uniformly with the increase of antennas number, which can make the measurement matrix achieve better orthogonal effect. The equally spaced pilot scheme can be applied more widely in the communication protocol due to its regular distribution.

5.3. NMSE versus Number of Antennas. The NMSE performance comparison of different algorithms with the number of antennas k is shown in Figure 5. In the comparison, the fixed parameters are used: SNR=30dB, v=350km/h, and G=190. The total pilot overhead reaches $2(Q-1)G/N \approx 23.2\%$, and the important reason for the high pilot overhead is the existence of guard pilots. The comparison is to observe the change of NMSE with the increase of the number of supporting antennas when the number of pilots is fixed.

We can find that the NMSEs of the OMP algorithm and the SAMP algorithm rise rapidly with the increase of the number of antennas, and the NMSE of the BSOMP algorithm also appears similar to them. However, the B-SAMP algorithm and the DCS-SOMP algorithm change relatively slowly. This is because the estimation model becomes a block sparse form in the B-SAMP algorithm by using the common sparsity of the spatial. The sparse block becomes larger as the number of antennas increases, so that the accuracy of the estimation does not decrease rapidly.

We can also find the channel information obtained by the B-SAMP algorithm, and the NMSE will be further reduced after smoothing. In the case of the same pilot overhead, the NMSE of the smoothed B-SAMP algorithm at 20 antennas are consistent with the B-SAMP algorithm at 10 antennas. We can conclude that the improved algorithms and methods can support channel estimation of more antennas under the same conditions, that is, reduce the pilot overhead when the number of antennas is constant.

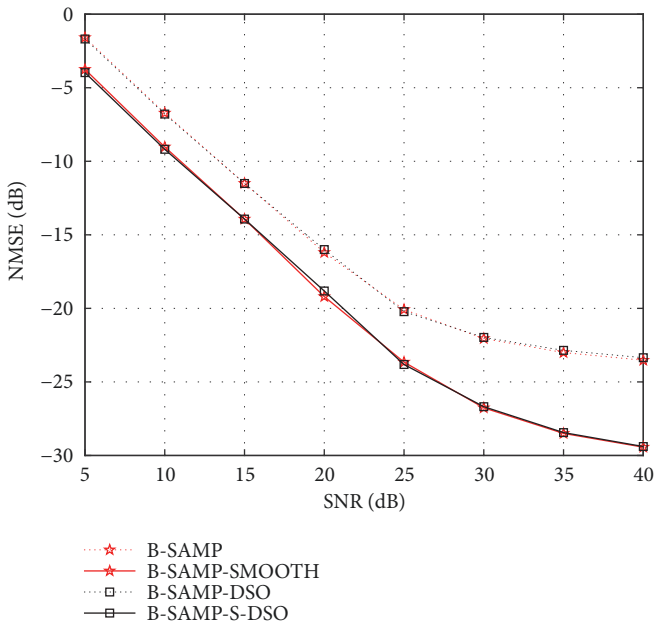

Figure 4: NMSE performance comparison of the equally spaced pilot scheme and the optimal placement scheme.

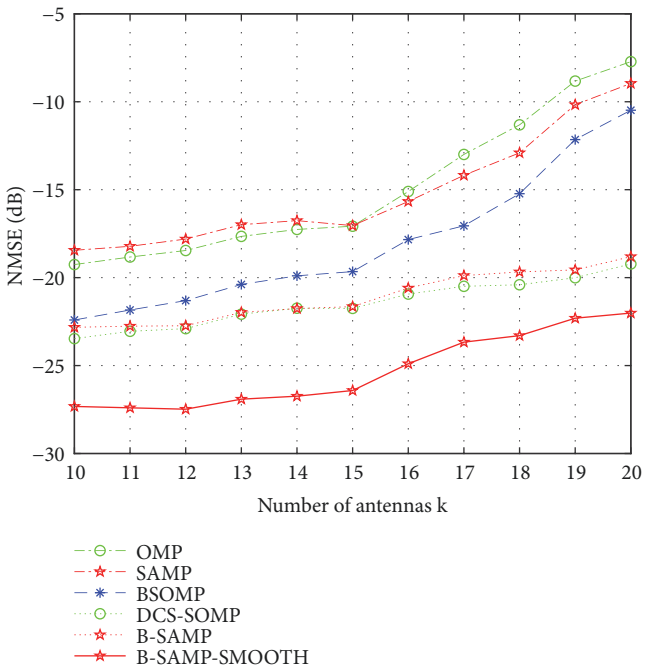

Figure 5: The NMSE performance versus the number of antennas k with SNR=30dB and v=350km/h.

6. Conclusions

In this paper, we use CE-BEM and proposed CS-based channel estimation scheme to complete DS channel estimation for massive MIMO systems. The ICI-free model is obtained by using the proposed equally placed pilot group scheme, and the B-SAMP algorithm is proposed by using the spatiotemporal common sparsity of the channel. Simulation results show that channel estimation using the proposed equally placed pilot group scheme and the B-SAMP algorithm achieves good estimation results. Compared with the existing DS channel estimation scheme, the proposed equally placed pilot group scheme has lower computational complexity and can meet the requirements of the measurement matrix; moreover the B-SAMP algorithm has good adaptive adjustment capability and stable estimation effects. In addition, the proposed channel estimation scheme can be easily implemented in practical systems.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request (email: kongsi0627@163.com).

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The authors would like to thank Shuangshuang Song, Yujie Fan, and Shanlin Wei, for their constructive comments and suggestions that improved the quality of the paper. This work was supported by the National Natural Science Foundation of China (Grant No. 61571364), the Shaanxi Provincial Natural Science Foundation (Grant No. 2017JM6073), the Seed Foundation of Innovation and Creation for Graduate Students in Northwestern Polytechnical University (Grant No.

ZZ2018118), and Innovation Foundation for Doctoral Dissertation of Northwestern Polytechnical University (Grant No. CX201833).

References

[1] M. Agiwal, A. Roy, and N. Saxena, "Next generation 5G wireless networks: a comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 3, pp. 1617–1655, 2016.

[2] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: benefits and challenges," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 742–758, 2014.

[3] H. Xie, F. Gao, and S. Jin, "An overview of low-rank channel estimation for massive MIMO systems," *IEEE Access*, vol. 4, pp. 7313–7321, 2016.

[4] C. R. Berger, Z. Wang, J. Huang, and S. Zhou, "Application of compressive sensing to sparse channel estimation," *IEEE Communications Magazine*, vol. 48, no. 11, pp. 164–174, 2010.

[5] W. U. Bajwa, J. Haupt, A. M. Sayeed, and 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.

[6] W. Ding, Y. Lu, F. Yang et al., "Spectrally efficient csi acquisition for power line communications: a Bayesian compressive sensing perspective," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 7, pp. 2022–2032, 2016.

[7] W. Ding, F. Yang, S. Liu, and J. Song, "Structured compressive sensing-based non-orthogonal time-domain training channel state information acquisition for multiple input multiple output systems," *IET Communications*, vol. 10, no. 6, pp. 685–690, 2016.

[8] X. Ma, F. Yang, S. Liu, J. Song, and Z. Han, "Design and optimization on training sequence for mmWave communications: a new approach for sparse channel estimation in massive MIMO," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 7, pp. 1486–1497, 2017.

[9] L. Dai, Z. Wang, and Z. Yang, "Spectrally efficient time-frequency training OFDM for mobile large-scale MIMO systems," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 251–263, 2013.

[10] W. Ding, F. Yang, S. Liu, X. Wang, and J. Song, "Nonorthogonal time-frequency training-sequence-based CSI acquisition for MIMO systems," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 7, pp. 5714–5719, 2016.

[11] Z. Gao, L. Dai, W. Dai, B. Shim, and Z. Wang, "Structured compressive sensing-based spatio-temporal joint channel estimation for FDD massive MIMO," *IEEE Transactions on Communications*, vol. 64, no. 2, pp. 601–617, 2016.

[12] X. Rao and V. K. Lau, "Distributed compressive CSIT estimation and feedback for FDD multi-user massive MIMO systems," *IEEE Transactions on Signal Processing*, vol. 62, no. 12, pp. 3261–3271, 2014.

[13] Z. Gao, L. Dai, Z. Wang, and S. Chen, "Spatially common sparsity based adaptive channel estimation and feedback for FDD massive MIMO," *IEEE Transactions on Signal Processing*, vol. 63, no. 23, pp. 6169–6183, 2015.

[14] M. Stojanovic and J. Preisig, "Underwater acoustic communication channels: propagation models and statistical characterization," *IEEE Communications Magazine*, vol. 47, no. 1, pp. 84–89, 2009.

[15] G. B. Giannakis and C. Tepedelenlioglu, "Basis expansion models and diversity techniques for blind identification and equalization of time-varying channels," *Proceedings of the IEEE*, vol. 86, no. 10, pp. 1969–1986, 1998.

[16] X. Ma, G. B. Giannakis, and S. Ohno, "Optimal training for block transmissions over doubly selective wireless fading channels," *IEEE Transactions on Signal Processing*, vol. 51, no. 5, pp. 1351–1366, 2003.

[17] X. Ma, F. Yang, S. Liu, W. Ding, and J. Song, "Structured compressive sensing-based channel estimation for time frequency training OFDM systems over doubly selective channel," *IEEE Wireless Communications Letters*, vol. 6, no. 2, pp. 266–269, 2017.

[18] X. Ma, F. Yang, S. Liu, J. Song, and Z. Han, "Sparse channel estimation for MIMO-OFDM Systems in high-mobility situations," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 7, pp. 6113–6124, 2018.

[19] P. Cheng, Z. Chen, Y. Rui et al., "Channel estimation for OFDM systems over doubly selective channels: a distributed compressive sensing based approach," *IEEE Transactions on Communications*, vol. 61, no. 10, pp. 4173–4185, 2013.

[20] Q. Qin, L. Gui, B. Gong, X. Ren, and W. Chen, "Structured distributed compressive channel estimation over doubly selective channels," *IEEE Transactions on Broadcasting*, vol. 62, no. 3, pp. 521–531, 2016.

[21] B. Gong, L. Gui, Q. Qin, X. Ren, and W. Chen, "Block distributed compressive sensing-based doubly selective channel estimation and pilot design for large-scale MIMO systems," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 10, pp. 9149–9161, 2017.

[22] Y. Mostofi and D. C. Cox, "ICI mitigation for pilot-aided OFDM mobile systems," *IEEE Transactions on Wireless Communications*, vol. 4, no. 2, pp. 765–774, 2005.

[23] F. Wan, W.-P. Zhu, and M. N. Swamy, "Semi-blind most significant tap detection for sparse channel estimation of OFDM systems," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 57, no. 3, pp. 703–713, 2010.

[24] I. E. Telatar and D. N. Tse, "Capacity and mutual information of wideband multipath fading channels," *Institute of Electrical and Electronics Engineers Transactions on Information Theory*, vol. 46, no. 4, pp. 1384–1400, 2000.

[25] E. Bjonson, J. Hoydis, M. Kountouris, and M. Debbah, "Massive MIMO systems with non-ideal hardware: energy efficiency, estimation, and capacity limits," *Institute of Electrical and Electronics Engineers Transactions on Information Theory*, vol. 60, no. 11, pp. 7112–7139, 2014.

[26] M. F. Duarte and Y. C. Eldar, "Structured compressed sensing: from theory to applications," *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp. 4053–4085, 2011.