IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 9 December 2022, accepted 25 December 2022, date of publication 3 January 2023, date of current version 6 January 2023.

Digital Object Identifier 10.1109/ACCESS.2022.3233823

METHODS

Estimation of Time-Varying Channels in Virtual Angular Domain for Massive MIMO Systems

MINGDUO LIAO, (Member, IEEE), AND YURIY ZAKHAROV, (Senior Member, IEEE)
School of Physics, Engineering and Technology, University of York, YO10 5DD York, U.K.

Corresponding author: Mingduo Liao (ml1806@york.ac.uk)

The work of Yuriy Zakharov was supported in part by the U.K. Engineering and Physical Sciences Research Council (EPSRC) under Grant EP/R003297/1 and Grant EP/V009591/1.

ABSTRACT This paper is concerned with the problem of channel estimation in virtual angular domain in massive MIMO orthogonal frequency division multiplexing (OFDM) systems, in the presence of time varying channels. The proposed channel estimator is based on dichotomous coordinate decent joint sparse recovery (DCD-JSR) algorithm, which was well suited for massive MIMO systems with channels static over several OFDM symbols, but it is not accurate for time-varying channels. The estimator proposed in this paper exploits the basis expansion model (BEM) for approximation of the time-variation of the channel and the sparsity of the channel in the virtual angular domain with a common support over OFDM subcarriers and OFDM symbols for joint estimation of the BEM expansion coefficients. It is shown that, when using the Legendre polynomials as the BEM and depending on the normalized Doppler frequency, only a small number of expansion coefficients is required to provide accurate channel estimation.

INDEX TERMS Basis expansion model (BEM), common sparsity, DCD-JSR algorithm, joint channel estimation, massive MIMO, orthogonal frequency-division multiplexing (OFDM), time varying channels.

I. INTRODUCTION

Massive multiple-input multiple-output (MIMO) has been proposed for the next generation communication systems, since it can provide high data rates by employing the wireless transmitter with a high number of antennas and provide high spectral efficiency by utilising the additional degree of freedom in the spatial domain [1], [2]. The massive MIMO faces challenges such as practical system modeling, channel estimation and so on [3]. Especially, accurate channel state information (CSI) serves as the key information for techniques such as coherent detection, therefore, an accurate channel estimation is necessary for practical applications of massive MIMO systems [4].

In [5], experimental results have shown that massive MIMO channels exhibit sparsity, due to the relatively small angle spread perceived from a base station (BS) between user and BS. According to [6], [7], and [8], the common sparsity is shared by different subcarriers due to the spatial propagation property of the wireless channel, such as the number of scatterers is remaining nearly unchanged over the system bandwidth, which is referred to as the spatially common sparsity over multiple subcarriers. Besides, since the angle variation from the user to the BS is relatively slow, we can consider the channel support is static over a sequence of OFDM symbols [7], [8]. Thus, in this paper, we consider that the channel exhibits sparsity, and shares common support over several OFDM symbols.

Sparse recovery techniques are attractive for channel estimation [9], [10], [11]. The channel estimation deals with estimation of complex-valued parameters. In [12], the dichotomous coordinate descent (DCD) algorithm is proposed for solving complex-valued sparse problems. It has been indicated in [12] and [13] that the computational complexity of the algorithm is dominated by the computational complexity of a small number of successful iterations, while most of the operations of the DCD algorithm are additions and bit-shifts. The algorithm provides accurate channel estimation and is computationally efficient [14], however, it does not suit massive MIMO channel estimation in the

The associate editor coordinating the review of this manuscript and approving it for publication was Wei Feng.

virtual angular domain. In [15], by employing the DCD algorithm and considering that the channel is statistic over several OFDM symbols, the DCD-JSR algorithm is proposed for the channel estimation in massive MIMO in the virtual angular domain; the simulation results have shown that the DCD-JSR algorithm outperformes the distributed sparsity adaptive matching pursuit (DSAMP) algorithm [7], however, it cannot provide accurate channel estimation when the channel is time-varying.

In [16] and [17] and many other publications, it has been shown that, the time-varying channel can be approximated accurately by employing the basis expansion model (BEM). Consequently, estimation of a realization of random process describing the time-varying channel is transformed into estimation of time-invariant expansion coefficients [18]. In [19], the Karhunen-Loeve BEM has been proposed for estimating the time-varying channel, however, it is very sensitive to the variations of the channel statistics. In [20] and [21], algebraic polynomial BEMs have been employed to estimate the time-varying channel, where the channel vectors can be approximated as a linear combination of a set of polynomials. In [17], the experimental results have indicated that, by employing the Legendre polynomials as the BEM, the channel variation could be accurately approximated. Thus, in this paper, we consider employing the Legendre polynomials to approximate the time-variation of the channel in the virtual angular domain.

In this paper, by combining the DCD-JSR algorithm [15] and the BEM, we show that the modified DCD-JSR algorithm can estimate the channel in OFDM system operating over frequency selective and highly mobile wireless time-varying channels. Simulation results show that, compared to the original DCD-JSR algorithm, the modified DCD-JSR algorithm could provide better MSE performance when estimating time-varying channels.

The paper is organized as follows. Section II describes the system model. In Section III, channel model for time varying channel is introduced. In Section IV, the processes of employing the basis functions, and estimating the time-varying channel are described. In Section V, numerical examples are analysed and, finally, Section VI presents the conclusion.

In this paper, capital and small bold fonts are used to denote matrices and vectors, respectively, $(\mathbf{x})_n$ denotes the nth element of the vector $\mathbf{x}$, $\mathbf{R}^{\mathrm{q}}$ denotes the qth column of the matrix $\mathbf{R}$, and $\mathbf{R}_{\mathrm{n}}$ denotes the nth row of the matrix $\mathbf{R}$, $\mathbf{R}_{\mathrm{m,n}}$ denotes an element of the matrix $\mathbf{R}$. The transpose operator is given by $(.)^{\mathrm{T}}$, $(.)^*$ denotes the conjugate operator, $(.)^{\dagger}$ denotes the Moore-Penrose inversion, and $(.)^{\mathrm{H}}$ denotes the Hermitian transpose operator. The ℓ_0-norm and ℓ_2-norm are represented by $||.||_0$ and $||.||_2$, respectively. We use I to denote a support, $|\mathrm{I}|$ is the cardinality of the support I, I^{c} is the complement of I, $\mathbf{R}_{\mathrm{I}}$ is a matrix obtained from $\mathbf{R}$, and which only contains rows corresponding to support I. $\mathbf{R}_{\mathrm{I,I}}$ is an $|\mathrm{I}| \times |\mathrm{I}|$ matrix obtained from $\mathbf{R}$ by collecting elements from columns and rows corresponding to I, and $\mathbf{x}_{\mathrm{I}}$ is the subset of $\mathbf{x}$ that includes non-zero elements from $\mathbf{x}$ corresponding to I. We use $\mathbf{h}$ to denote a channel vector and $\tilde{\mathbf{h}}$ to denote the channel vector in the virtual angular domain, $\tilde{\mathbf{h}_n}$ denotes the channel vector corresponding to the nth subcarrier. $\Re$ denotes the real part of a complex number, and $\mathrm{j} = \sqrt{-1}$.

II. SYSTEM MODEL

A. CHANNEL MODEL

In a massive MIMO system, consider a time interval consisting of J OFDM symbols. M antennas are employed at the BS to serve K single-antenna users simultaneously, where $\mathrm{M} \gg \mathrm{k}$. At the tth OFDM symbol, $1 \leq \mathrm{t} \leq \mathrm{J}$, for the nth subcarrier, $1 \leq \mathrm{n} \leq \mathrm{N}$, the received signal model for the kth user, $1 \leq \mathrm{k} \leq \mathrm{K}$, in the frequency domain is given by

$$\mathrm{y}^{\mathrm{t}}_{\mathrm{k,n}} = \left(\mathbf{h}^{\mathrm{t}}_{\mathrm{k,n}}\right)^{\mathrm{T}} \mathbf{x}^{\mathrm{t}}_{\mathrm{n}} + \mathrm{w}^{\mathrm{t}}_{\mathrm{k,n}}, \tag{1}$$

where $\mathbf{h}^{\mathrm{t}}_{\mathrm{k,n}} \in \mathrm{C}^{\mathrm{M}\times 1}$ represents the downlink channel between the M BS antennas and the kth user, $\mathbf{x}^{\mathrm{t}}_{\mathrm{n}} \in \mathrm{C}^{\mathrm{M}\times 1}$ is the vector of transmitted symbols (data or pilot symbols) and $\mathrm{w}^{\mathrm{t}}_{\mathrm{k,n}}$ is the corresponding additive white Gaussian noise (AWGN). For a single user, we can drop the index k, thus we can write

$$\mathrm{y}^{\mathrm{t}}_{\mathrm{n}} = \left(\mathbf{h}^{\mathrm{t}}_{\mathrm{n}}\right)^{\mathrm{T}} \mathbf{x}^{\mathrm{t}}_{\mathrm{n}} + \mathrm{w}^{\mathrm{t}}_{\mathrm{n}}. \tag{2}$$

Matrix $\mathbf{A}_{\mathrm{B}}$ is used to modify the channel vector $\mathbf{h}^{\mathrm{t}}_{\mathrm{n}}$ into a channel vector $\tilde{\mathbf{h}}^{\mathrm{t}}_{\mathrm{n}}$ in the virtual angular domain, and it is determined by the geometric structure of the antenna array. We consider a uniform linear array with the antenna spacing $\lambda/2$, where λ is the wavelength, then $\mathbf{A}_{\mathrm{B}}$ becomes the discrete Fourier transform (DFT) matrix. Thus we obtain

$$\mathrm{y}^{\mathrm{t}}_{\mathrm{n}} = \left(\tilde{\mathbf{h}}^{\mathrm{t}}_{\mathrm{n}}\right)^{\mathrm{T}} \mathbf{A}^{*}_{\mathrm{B}} \mathbf{x}^{\mathrm{t}}_{\mathrm{n}} + \mathrm{w}^{\mathrm{t}}_{\mathrm{n}}, \tag{3}$$

where $\left(\mathbf{h}^{t}_{n}\right)^{T} = \left(\tilde{\mathbf{h}}^{t}_{n}\right)^{T} \mathbf{A}^{*}_{\mathrm{B}}$. The channel vector in the angular domain divides the covering area of the BS into angular intervals, and the mth element of $\tilde{\mathbf{h}}^{t}_{n}$ corresponds to the mth virtual angle, where $1 \leq \mathrm{m} \leq \mathrm{M}$.

In massive MIMO systems, the BS is often positioned at a high elevation with a small number of scatterers (in contrast to the number of antennas) [5], [22], but the scatterers on the user side are relatively rich. Thus, we can consider for the kth user, in the virtual angular domain, only a small number $|\mathrm{I}|$ of multipath arrivals contains the majority of the channel energy. Hence, we have $|\mathrm{I}| \leq \mathrm{M}$, and the channel vector exhibits sparsity in the virtual angular domain.

We consider that the channel is static over one OFDM symbol, furthermore, as indicated in [7], [8], and [15], since the spatial propagation characteristics such as scatterers are almost unchanged over the system bandwidth, as shown in Fig.1, we can consider that in the same OFDM symbol, for different subcarriers, the subchannel exhibits common sparsity.

According to [23], in time-varying scenarios, the variation of the arrival angles is usually much slower than that of channel gain, thus, we can consider that during J successive OFDM symbols, the channel exhibits common sparsity. Thus,

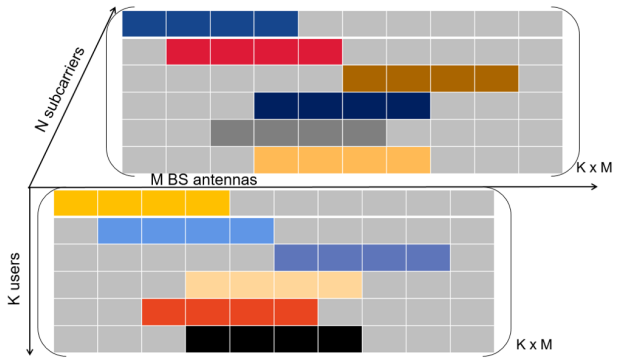

FIGURE 1. The channel vector exhibits common sparsity within the system bandwidth in the virtual angular domain.

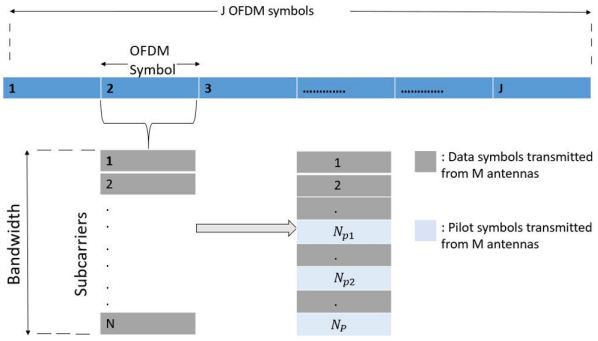

FIGURE 2. Each OFDM symbol contains N subcarriers of which P subcarriers are used to transmit pilot symbols.

we can obtain

$$\mathrm{I}_1^{\mathrm{l}} = \mathrm{I}_{\mathrm{n}}^{\mathrm{j}} = \mathrm{I}_{\mathrm{N}}^{\mathrm{J}} = \mathrm{I},\ 1 \le \mathrm{j} \le \mathrm{J},\ 1 \le \mathrm{n} \le \mathrm{N}, \tag{4}$$

where $\mathrm{I}_{\mathrm{n}}^{\mathrm{j}}$ is the support set for the Jth OFDM symbol at the nth subcarrier.

In this paper, we consider the pilot-aided channel estimation. For the tth OFDM symbol, a part of subcarriers is used for transmitting pilot symbols $\mathbf{s}_{\mathrm{p}}^{\mathrm{t}} \in \mathrm{C}^{\mathrm{M}\times 1}$, and the received signal at the pilot subcarrier n(p) is given by:

$$\mathrm{y}_{\mathrm{n(p)}}^{\mathrm{t}} = \left(\tilde{\mathbf{h}}_{\mathrm{n(p)}}^{\mathrm{t}}\right)^{\mathrm{T}} \mathbf{A}_{\mathrm{B}}^{*} \mathbf{s}_{\mathrm{p}}^{\mathrm{t}} + \mathrm{w}_{\mathrm{n(p)}}^{\mathrm{t}}, \tag{5}$$

$$\left[\mathbf{s}_{\mathrm{p}}^{\mathrm{t}}\right]_{\mathrm{m}} = e^{j\theta_{t,m,p}}, \tag{6}$$

$$1 \le \mathrm{p} \le \mathrm{P},\ 1 \le \mathrm{m} \le \mathrm{M},\ 1 \le \mathrm{t} \le \mathrm{J}$$

where $\theta_{\mathrm{t,m,p}}$ are independent random numbers uniformly distributed in $(0, 2\pi]$.

B. CHANNEL ESTIMATION APPROACH

The conventional method to acquire the channel state information (CSI) in frequency-division-duplexing (FDD) systems is as follows: the BS transmits downlink pilot symbols to a user, so the user can estimate the downlink CSI locally and then feed it back to the BS via an uplink channel [24]. If we are employing conventional CSI estimation techniques (such as the minimum mean square error (MMSE) estimator), since the number of pilot symbols required at the BS has to scale linearly with the number of transmit antennas [22], it would cause prohibitively large overhead for both pilot training (downlink) and CSI feedback (uplink). Hence, to solve the overhead issue, as suggested in [7], the channel estimation is performed at the BS. The channel estimation scheme is summarized as follows.

1 . In each OFDM symbol, every BS antenna broadcasts pilot symbols to users, the kth user receives the signal $\mathbf{y}_{\mathrm{k}} \in \mathrm{C}^{\mathrm{M}\times 1}$ and feeds it back to the BS. The BS recovers the CSI for each user based on the feedback signals $\mathbf{y}_{\mathrm{k}}$, $\mathrm{k} = 1, \ldots, \mathrm{K}$. As shown in Fig. 2, each OFDM symbol contains N subcarriers, while P subcarriers are used to transmit pilot symbols. The user feeds back the received signal to the BS without performing downlink channel estimation.

2 . At the BS, by employing the BEM to approximate the time variation of the channel, the DCD-JSR algorithm can jointly estimate the common support I for multiple sparse virtual angular domain channels. The least squares (LS) algorithm [25] is then employed to acquire the CSI based on an estimate of the common support I.

C. RECEIVED SIGNAL

For the tth OFDM symbol, at the pth pilot subcarrier, the signal received at the BS is given by:

$$\mathrm{r}_{\mathrm{p}}^{\mathrm{t}} = \boldsymbol{\phi}_{\mathrm{p}}^{\mathrm{t}} \tilde{\mathbf{h}}_{\mathrm{n(p)}}^{\mathrm{t}} + \mathrm{v}_{\mathrm{p}}^{\mathrm{t}},\ 1 \le \mathrm{p} \le \mathrm{P},\ 1 \le \mathrm{t} \le \mathrm{J}. \tag{7}$$

Here, $\boldsymbol{\phi}_{\mathrm{p}}^{\mathrm{t}} = \left(\mathbf{s}_{\mathrm{p}}^{\mathrm{t}}\right)^{\mathrm{T}} \left(\mathbf{A}_{\mathrm{B}}^{*}\right)^{\mathrm{T}} \in \mathrm{C}^{1\times \mathrm{M}}$ is the sensing vector defined by the DFT matrix and the pilot symbols. $\tilde{\mathbf{h}_{n(p)}^{t}} \in \mathrm{C}^{\mathrm{M}\times 1}$ is the sparse channel vector for the n(p)th subcarrier, and $\mathrm{v}_{\mathrm{p}}^{\mathrm{t}}$ is the corresponding noise, which contains both downlink and uplink channel noise.

III. TIME VARYING CHANNEL

In practice, due to the user mobility, the propagation of wireless signals would face the time-varying environment [1]. Due to the simple implementation and the orthogonality between columns of the basis-expansion matrix, using the Legendre polynomials as basis functions on the investigated time interval has been considered in the literature [17]. In this paper, we consider that for the pth pilot subcarrier, the time-varying channel vector can be approximated by $\mathrm{N_b}$ basis functions:

$$\hat{\mathbf{h}}_{\mathrm{n(p)}}^{\mathrm{t}} = \sum_{i=1}^{\mathrm{N_b}} \mathrm{b_i(t)} \mathbf{c}_{\mathrm{i,p}},\quad 1 \le \mathrm{t} \le \mathrm{J}, \tag{8}$$

where $\mathrm{b_i(t)}$ is the tth element of a vector $\mathbf{b}_{\mathrm{i}} \in \mathrm{C}^{\mathrm{J}\times 1}$ representing samples of the basis function $\mathrm{b_i(t)}$, $\mathbf{c}_{\mathrm{i,p}} \in \mathrm{C}^{\mathrm{M}\times 1}$ are expansion coefficients for the ith basis function at the pth pilot

subcarrier. This approximation is valid if the channel variation are slow enough over one OFDM symbol, so that we can ignore the intercarrier interference between the OFDM subcarriers. Thus, we mostly deal with variation of the channel over a sequence of J OFDM symbols. In such scenarios, the channel approximation in (8) can be made arbitrary accurate by choosing a large enough number of basis functions N_b. In our numerical investigation below, we consider scenarios with a limited product f_dT of the Doppler frequency f_d by the (orthogonality) duration of the OFDM symbol T, specifically, $f_dT \leq 0.05$.

In this paper, we consider employing the Legendre polynomials as the basis functions [26], defined as:

$$b_i(t) = \frac{1}{2^{i-1}i!}\frac{d^{i-1}}{dt^{i-1}}[(t^2-1)^{i-1}],\ i \geq 1. \tag{9}$$

For the tth OFDM symbol at the pth pilot subcarrier, by substituting (8) into (7), we can obtain:

$$r_p^t = \boldsymbol{\phi}_p^t \sum_{i=1}^{N_b} b_i(t)\mathbf{c}_{i,p} + v_p^t,\ 1 \leq p \leq P,\ 1 \leq t \leq J. \tag{10}$$

Since $\tilde{\mathbf{h}^{t}_{n(p)}} \in C^{M\times 1}$ exhibit common sparsity, the expansion coefficient vectors $\mathbf{c}_{i,p} \in C^{M\times 1}$ also exhibit common sparsity. Thus, the task of estimating JM channel coefficients is transformed into estimating only $N_b|I|$ expansion coefficients with usually $N_b \ll J$ and $|I| \leq M$.

We collect the received signal samples r_p^t, $1 \leq t \leq J$, in a vector $\mathbf{r}_p = \left[r_p^1, r_p^2, \ldots, r_p^J\right]^T \in C^{J\times 1}$, then we have:

$$\mathbf{r}_p = \sum_{i=0}^{N_b} \mathbf{F}_{i,p}\mathbf{c}_{i,p} + \mathbf{v}_p,\quad 1 \leq p \leq P, \tag{11}$$

where $\mathbf{F}_{i,p}^t = \boldsymbol{\phi}_p^t b_i(t) \in C^{1\times M}$, $\mathbf{F}_{i,p} = [\mathbf{F}_{i,p}^1, \mathbf{F}_{i,p}^2, \ldots, \mathbf{F}_{i,p}^t] \in C^{J\times M}$ is a matrix whose tth row is $\mathbf{F}_{i,p}^t$, and $\mathbf{v}_p = \left[v_p^1, v_p^2, \ldots, v_p^J\right]^T \in C^{J\times 1}$ is the noise vector. Since the expansion coefficients exhibit common sparsity, we can firstly estimate the common support and then find the expansion coefficients.

IV. DCD-JSR ALGORITHM FOR TIME-VARYING CHANNEL

Here, the homotopy DCD algorithm [12] is used to estimate the support of the expansion coefficients, as shown in Table 1. First, we apply the homotopy DCD algorithm to the $\ell_2\ell_0$ optimization problem of minimizing:

$$\mathbf{J}_\tau(\tilde{\mathbf{c}}_{i,p}) = \frac{1}{2}\left\|\mathbf{r}_p - \mathbf{F}_{i,p}\tilde{\mathbf{c}}_{i,p}\right\|_2^2 + \tau\left\|\tilde{\mathbf{c}}_{i,p}\right\|_0. \tag{12}$$

Here, we solve the optimization problem for the pth pilot subcarrier of the ith expansion coefficient, and $\tau \in [0, 1)$ is a regularization parameter. The second term in (12) makes it a non-convex problem and the solution of it is NP-hard. To solve the problem, we initially assign the support set $I_p = \emptyset$, and by following the proposition in [12] we can add new elements into the support or remove elements from the support in several iterations, thus, the estimated expansion coefficients $\tilde{\mathbf{c}_{i,p}}$ can be obtained.

Therefore we need to assign initially a high value to the regularization parameter $\tau = \tau_{max}$, so that the second term in (12) dominates the cost function to provide an empty support $I_p = \emptyset$. In the homotopy iterations, by gradually reducing value of τ as $\tau \leftarrow \gamma\tau$, where $\gamma \in [0, 1)$, new elements can be added to the support or removed from the support [12]. The algorithm stops when $\tau < \tau_{min}$, where $\tau_{min} = \mu_\tau\tau_{max}$ and $\mu_\tau \in [0, 1)$ is a predefined parameter.

To reduce the computational complexity [12], instead of solving the LS problem in Table 1, step 3, we employ the DCD iterations to solve the LS problem, as shown in Table 2, where N_u is the number of successful DCD iterations, and a successful DCD iteration means that the solution is updated.

Following is an example of how we estimate the common support for the expansion coefficients. For the

TABLE 1. $\ell_2\ell_0$ homotopy DCD algorithm.

Initialization: For the ith expansion coefficient at the pth pilot subcarrier, vector $\mathbf{d} = \mathbf{0}$, $I_p = \emptyset$, $\mathbf{b}_p = \mathbf{F}_{i,p}{}^H\mathbf{r}_p$, $\mathbf{R}_p = \mathbf{F}_{i,p}{}^H\mathbf{F}_{i,p}$.

1: $g = \arg\max_{k\in I_p^c} |(\mathbf{b}_p)_k|^2/(\mathbf{r}_p)_{k,k}$,
$\tau_{max} = (1/2)\max_k |(\mathbf{b}_p)_k|^2/(\mathbf{R}_p)_{k,k}$,
$\tau = 0.5\left|(\mathbf{b}_p)_g\right|^2/(\mathbf{R}_p)_{g,g}$, $I_p = \{g\}$.
2: **Repeat** until the termination condition is met:
3: **If** the support I_p has been updated **then**
Solve $(\mathbf{R}_p)_{I_p,I_p}\left[\tilde{\mathbf{c}}_{i,p}\right]_{I_p} = \mathbf{f}_p$,
where $\mathbf{f}_p = \left[(\mathbf{F}_{i,p}\right]_{I_p}^H\mathbf{R}_p$
$\mathbf{d} \leftarrow \mathbf{b} - (\mathbf{R}_p)_{I_p,I_p}\left[\tilde{\mathbf{c}}_{i,p}\right]_{I_p}$
4: Update the regularization parameter : $\tau \leftarrow \gamma\tau$
5: Add the g-th element into the support I_p,
where $g \in I_p{}^c$,
and $g = \arg\max_{k\in I_p^c} \frac{|(\mathbf{c})_k|^2}{(\mathbf{R}_p)_{k,k}}$ s.t $\left|(\mathbf{c})_g\right|^2 > 2\tau\,(\mathbf{R}_p)_{g,g}$,
then assign to $\left[\tilde{\mathbf{c}}_{i,p}\right]_g$ the value $(\mathbf{d})_g/(\mathbf{R}_p)_{g,g}$,
update $\mathbf{d} \leftarrow \mathbf{d} - \left[\tilde{\mathbf{c}}_{i,p}\right]_g\mathbf{R}_p^g$.
6: Remove the gth element from the support I_p,
where $g \in I_p$, and
$g = \arg\min_{k\in I_p}\left[\frac{1}{2}|\left[\tilde{\mathbf{c}}_{i,p}\right]_k|^2(\mathbf{R}_p)_{k,k} + \Re\left\{\left[\tilde{\mathbf{c}}_{i,p}\right]_k^*(\mathbf{c})_k\right\}\right]$,
s.t. $\frac{1}{2}\left|(\mathbf{c}_{i,p})_g\right|^2(\mathbf{R}_p)_{g,g} + \Re\left\{\left[(\tilde{\mathbf{c}}_{i,p}\right]_g^*(\mathbf{d})_g\right\} < \tau$
for every removed element,
update $\mathbf{d} \leftarrow \mathbf{d} + \left(\tilde{\mathbf{c}}_{i,p}\right)_g\mathbf{R}_p^g$ and set $\left[(\tilde{\mathbf{c}}_{i,p}\right]_g = 0$.
7: If $\tau < \tau_{min}$, Stop.

TABLE 2. DCD iterations for LS minimization.

Input: $\tilde{c}_{i,p}$, $\mathbf{d}$, I_p, $\mathbf{R}_p$
Initialization: $s = 0$, $\delta = H$
1: **for** $m = 1, ..., M_b$ **do** until $s = N_u$
2: $\delta = \delta/2$, $\boldsymbol{\alpha} = [\delta, -\delta, j\delta, -j\delta]$, State =0
3: **for** $n = 1, ..., |I_p|$ **do**: $e = I_p(n)$
4: **for** $k = 1, ..., 4$ **do**
5: **if** $\Re\left\{(\alpha)_k(d)_e^*\right\} > \left[(R_p)_{e,e}\right]\delta^2/2$ **then**
6: $\left[\tilde{c}_{i,p}\right]_e \leftarrow \left[\tilde{c}_{i,p}\right]_e + (\alpha)_k$, $\mathbf{d} \leftarrow \mathbf{d} - (\alpha)_k\mathbf{R}_p{}^{(e)}$
7: State=1, $s \leftarrow s + 1$
8: **if** State=1, **go to step 3**

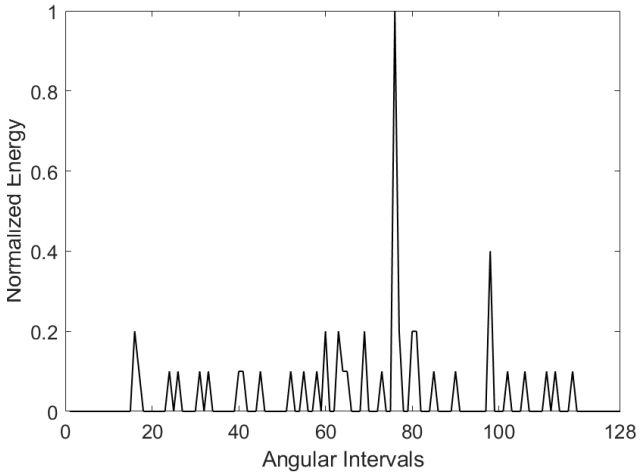

FIGURE 3. Average of normalized energy of elements of the vector $\tilde{\mathbf{c}}_{1,1}$, for the expansion coefficient of the first basis function (zero-order Legendre polynomial) for the first pilot subcarrier against the angular intervals. SNR = 20 dB, J = 100, |I| = 8.

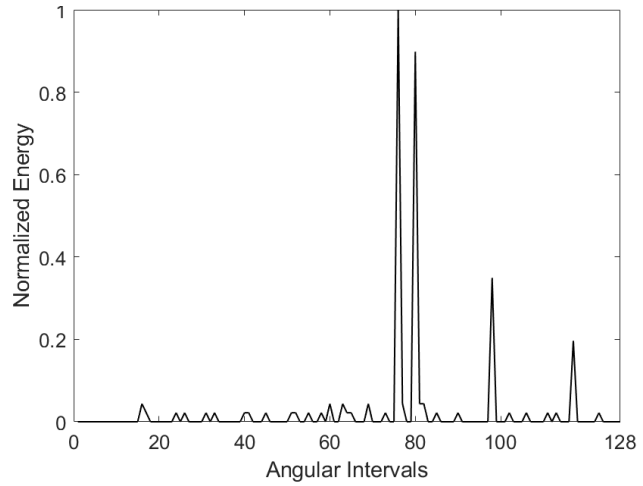

FIGURE 4. Average of the normalized energy of elements of the vector $\mathbf{q}_1$, for the expansion coefficients of two basis functions (zero-order and first order Legendre polynomials) and all pilot subcarriers against the angular intervals. SNR = 20 dB, J = 100, |I| = 8.

simulation scenario, we consider a massive MIMO system, SNR=20 dB, M = 128, P = 64, J = 100, the normalized Doppler frequency $f_dT = 0.05$ and $N_b = 3$, $|I| = 8$.

Fig. 3 shows the normalized energy of the elements of the estimated expansion coefficient $\mathbf{c}_{1,1}$, for the first basis function (zero-order Legendre polynomial) at the first pilot subcarrier in the angular intervals. It is easy to see that, due to the large variance of the energy of the elements in the angular intervals, we can not clearly identify the support.

Since all expansion coefficients share a common support, for the pth pilot subcarrier, we can compute another vector with contribution from all the expansion coefficients:

$$\mathbf{q}_p = \sum_{i=1}^{N_b} \left|\tilde{\mathbf{c}}_{i,p}\right|^2, \tag{13}$$

$$\tilde{\mathbf{q}_p} = \mathbf{q}_p/(\max(\mathbf{q}_p)N_b). \tag{14}$$

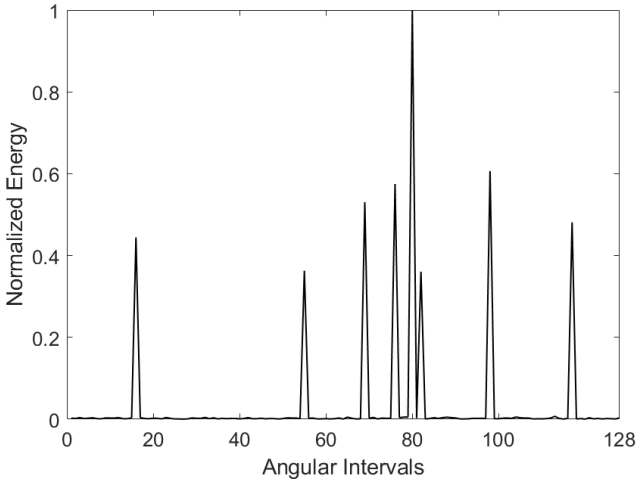

FIGURE 5. Average of the normalized energy of elements of the vector q, for all basis functions and all pilot subcarriers against the angular intervals. SNR = 20dB, J = 100, |I| = 8.

Here, as shown in Fig. 4, $\tilde{\mathbf{q}_p}$ is a vector that contains normalized energy of elements for all expansion coefficients at the pth pilot subcarrier in the angular intervals. The new plot shows clearly 4 of 8 non-zero directions. However, the variance is still large and we can not estimate reliably the support at this step.

As indicated in the previous section, since the expansion coefficients $\mathbf{c}_{i,p}$ share the common support among P subcarriers, we can compute a new vector, which takes this into account:

$$\mathbf{q} = \sum_{i=1}^{N_b}\sum_{p=1}^{P} \left|\tilde{\mathbf{c}}_{i,p}\right|^2, \tag{15}$$

$$\tilde{\mathbf{q}} = \mathbf{q}/(\max(\mathbf{q})N_bP). \tag{16}$$

As shown in Fig.5, here, $\tilde{\mathbf{q}} \in C^{M\times 1}$ is a sparse vector with elements averaging contribution from all pilot subcarriers and all expansion coefficients. We can acquire now the common support $\tilde{I}$ by using the hard thresholding

$$I = \{k : [\mathbf{q}]_k \tilde{} > \xi \max[\mathbf{q}]\}, \tag{17}$$

where ξ is a predefined thresholding parameter.

Based on the support estimate $\tilde{I}$, the MMSE approach [25] is employed to estimate the expansion coefficients $\tilde{\mathbf{c}_p}$:

$$\tilde{\mathbf{c}_p} = \mathbf{R}_{pp}(\mathbf{R}_{gg} + \sigma^2\mathbf{w})^{-1} \tag{18}$$

$$\mathbf{R}_{pp} = \mathbf{F}_p^H\mathbf{r}_p \tag{19}$$

Here, $\mathbf{w}$ is the identity matrix with size of $|\tilde{I}| \times |\tilde{I}|$, $\tilde{\mathbf{c}_p} \in C^{1\times MN_b}$ is a vector containing $\tilde{\mathbf{c}}_{i,p}$ for N_b basis functions, and $\mathbf{F}_p \in C^{J\times MN_b}$ is the matrix containing N_b vectors $\mathbf{F}_{i,p}$, and σ^2 is the noise variance, which is assumed to be known. The estimation of noise variance is a well known problem (e.g., see [27], [28], [29]) and it is not considered here.

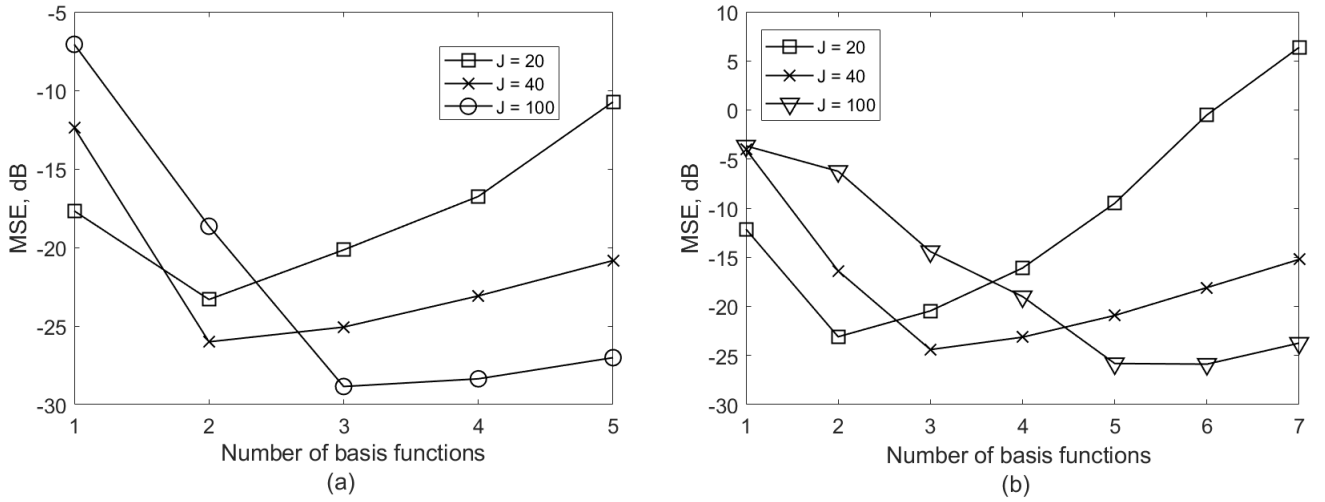

FIGURE 6. MSE performance of the DCD-JSR algorithm against the number of employed basis functions; SNR = 20 dB, |I| = 3. (a) $f_dT = 0.02$, (b) $f_dT = 0.05$.

V. SIMULATION RESULTS

The mean squared error (MSE) of the channel estimation will be used to asses the algorithm performance. First, we compute the MSE for the tth OFDM symbol at the pth subcarrier:

$$\mathrm{MSE}_p^t = \frac{||\hat{\mathbf{h}}_p^t - \tilde{\mathbf{h}}_{n(p)}^t||_2^2}{||\tilde{\mathbf{h}}_{n(p)}^t||_2^2}, \tag{20}$$

$$||\tilde{\mathbf{h}}_{n(p)}^t||_2^2 = (\tilde{\mathbf{h}}_{n(p)}^t)^H(\tilde{\mathbf{h}}_{n(p)}^t), \tag{21}$$

where $\hat{\mathbf{h}}_p^t$ is the estimated channel vector obtained from (8), and $\tilde{\mathbf{h}}_p^t$ is the true channel vector. Then the overall MSE for the channel estimation is computed by:

$$\mathrm{MSE} = \frac{1}{JP}\sum_{p=1}^{P}\sum_{t=1}^{J}\mathrm{MSE}_p^t. \tag{22}$$

The MSE in (22) is further averaged over the simulation trials.

We consider simulation scenarios corresponding to a MIMO system with a uniform linear array. For the massive MIMO system, in most simulation scenarios, we consider the number of antennas $M = 128$, the number of pilot subcarriers $P = 64$, the sampling frequency $f_s = 15.36$ MHz, the time interval for one OFDM symbol $T = 66.7\ \mu$s, the carrier frequency $f_c = 2.5$ GHz and the number of simulation trials $N_s = 500$. The performance of the oracle LS algorithm [27] with known support is adopted as the performance bound. In most scenarios, we consider two cases, SNR = 10 dB and SNR = 20 dB, and the user mobility with $v = 120\frac{km}{h}$ and $v = 300\frac{km}{h}$. The Doppler frequency f_d can be obtained by using:

$$f_d = f_c\frac{v}{v_c}, \tag{23}$$

where $v_c = 3 \times 10^8$ m/s is the light speed.

The channel estimation performance is investigated in several ways. First, we investigate the MSE performance of the proposed algorithm with different number of employed OFDM symbols, then we compare the MSE performance against the number of basis functions, with normalized Doppler frequencies $f_dT = 0.02$, $f_dT = 0.05$, where the Doppler frequency is approximately $f_d = 300$ Hz, $f_d = 700$ Hz, respectively. After that, we compare the MSE performance for different SNR, considering the normalized Doppler frequency $f_dT = 0.05$. The MSE performance against the number of DCD-iterations is investigated to show the convergence of the proposed algorithm, and the MSE performance against the number of employed antennas is also investigated. Furthermore, we compare the MSE performance of the DCD-JSR algorithm and the distributed sparsity adaptive matching pursuit (DSAMP) algorithm from [7] against the number of non-zero virtual angles, the oracle LS algorithm with known support is adopted as the performance bound [25]. At last, the computational complexity of the DCD-JSR algorithm is analyzed.

In Fig. 6, we show the MSE performance of the DCD-JSR algorithm for different number of employed OFDM symbols, for SNR = 20 dB, |I| = 3, $f_dT = 0.02$ and $f_dT = 0.05$. It can be seen that, as the number of OFDM symbols increases, the better MSE performance is provided by employing more basis functions. Thus we can conclude that, with longer data packets, more basis functions are required for the DCD-JSR algorithm to provide the best MSE performance, and the minimum MSE is also reduced.

In Fig. 7, we compare the MSE performance for different normalized Doppler frequencies, $f_dT = 0.02$ and $f_dT = 0.05$. The number of employed OFDM symbols is set to $J = 100$, and the number of non-zero virtual angles $|I| = 3$. It can be seen that, for the higher normalized Doppler frequency, we need to employ more basis functions to obtain the best

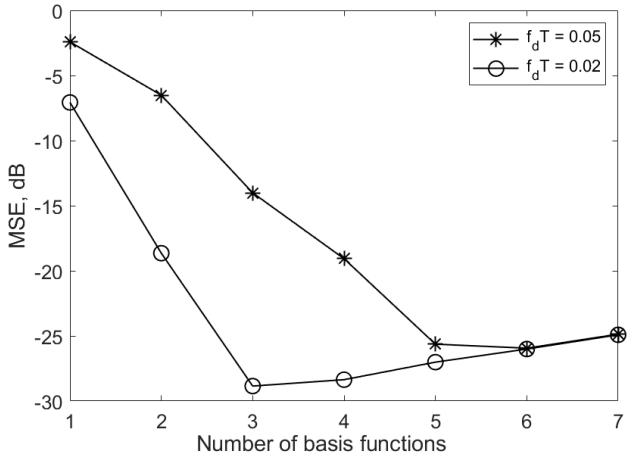

FIGURE 7. MSE performance of the DCD-JSR algorithm against the number of employed basis functions; SNR = 20 dB, J = 100, |I| = 3.

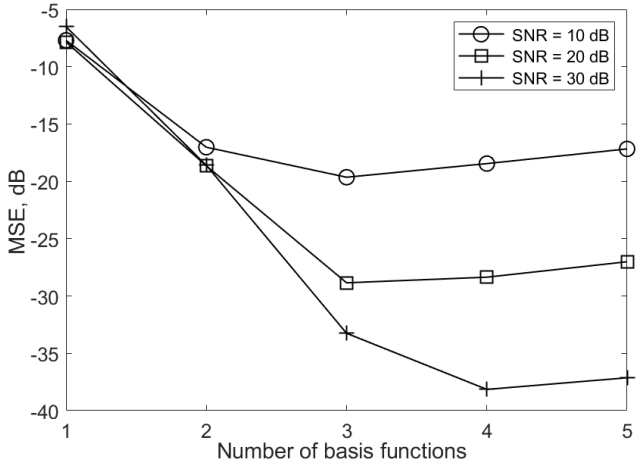

FIGURE 8. MSE performance of the DCD-JSR algorithm against the number of employed basis functions; $f_dT = 0.02$, J = 100, |I| = 3.

MSE performance. In other words, to provide the best MSE performance, the number of basis functions to be employed increases with the normalized Doppler frequency, which means that with higher user mobility, the more basis functions is required to provide accurate channel estimation.

In Fig. 8, we investigate the number of basis functions required to provide the best MSE performance under different SNR scenarios, for the case $f_dT = 0.02$, J = 100, |I| = 3. It can be seen that, as the SNR increases, the number of basis functions required to provide the best MSE performance is increased. This is because when the SNR is low, the main issue for channel estimation is the noise, a small number of basis functions is required to approximate the channel. When the SNR is high, the algorithm can focus more on the time variation of the channel, thus the number of basis functions required will be larger. Hence, we can say that, the number of basis functions required to approximate the time-varying channel should be higher for higher SNR.

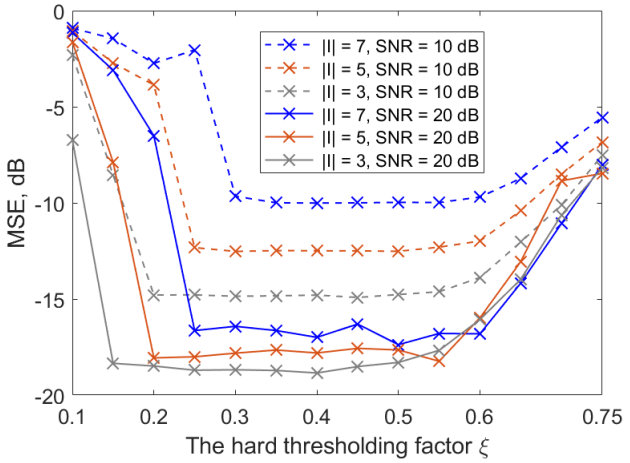

FIGURE 9. MSE performance of the DCD-JSR algorithm against the hard thresholding factor ξ; $f_dT = 0.05$, $N_b = 3$, J = 40.

To provide the best MSE performance, the thresholding factor ξ needs to be properly adjusted. In Fig. 9, we investigate the MSE performance of the DCD-JSR algorithm against the hard thresholding factor ξ, for the case $N_b = 3$, $f_dT = 0.05$, J = 40. It is clear that, for both cases SNR = 10 dB and SNR = 20 dB, as the number of non-zero virtual angles |I| increases, the range for the thresholding factor ξ which can provide the best MSE performance decreases. However, in all the cases, the thresholding factor can be chosen in the interval $[0.30, 0.55]$ to provide the minimum MSE.

Fig. 10 shows the MSE performance of the DCD-JSR algorithm in scenarios with different number of non-zero virtual angles against the number of DCD iterations, for the case $N_b = 2$, $f_dT = 0.02$, J = 40. It can be seen that, in all these scenarios, after a few DCD iterations, the algorithm converges to the best MSE. However, the smaller number of non-zero angles, the faster is the convergence. For $|I| \leq 9$, a single DCD iteration is enough for the convergence.

In Fig. 11(a), we show the MSE performance for different number of employed antennas, for the case $N_b = 3$, $f_dT = 0.05$, J = 20, |I| = 4. It can be seen that, with a small number of antennas, the MSE performance of the DCD-JSR algorithm is poor. For SNR = 10 dB, it requires M = 56 to approach the oracle performance, for SNR = 20dB, it requires at least M = 32. In Fig. 11 (b), we show the probability of perfect support estimation against the number of employed antennas, where a perfect support estimation means the estimated support is exactly the same as the true support, for the case |I| = 4, $f_dT = 0.05$, J = 20. It can be seen that, at SNR = 10 dB, with a small number of employed antennas, we cannot estimate the support correctly until M = 64; this explains why the MSE performance is poor with small number of antennas. We have run our simulations up to M = 512 and observed that the MSE performance does not change.

In Fig. 12(a) and Fig. 12(b), for the DCD-JSR algorithm with $N_b = 1$ and $N_b = 2$, and the distributed sparsity

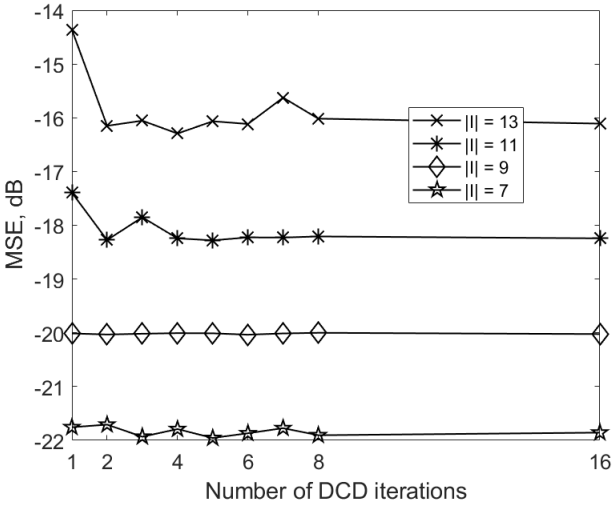

FIGURE 10. MSE performance of the DCD-JSR algorithm against the number of DCD iterations; $f_dT = 0.02$, $N_b = 2$, $J = 40$.

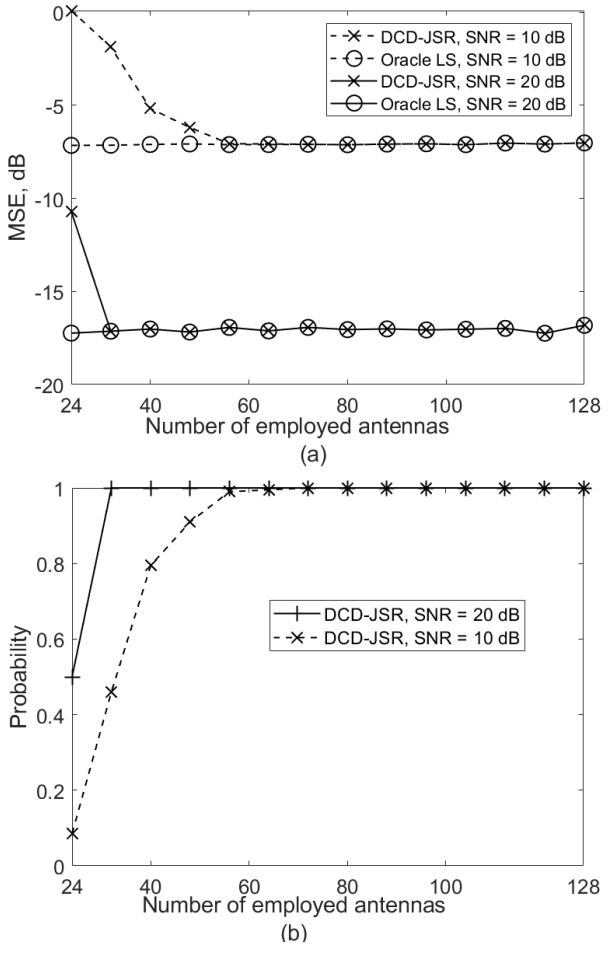

FIGURE 11. (a) MSE performance of the DCD-JSR algorithm against the number of employed antennas; (b) Probability of perfect support estimation against the number of employed antennas. $f_dT = 0.05$, $N_b = 3$, $J = 20$, $|I| = 4$.

adaptive matching pursuit algorithm (DSAMP) [7], we show the MSE performance for different number of non-zero

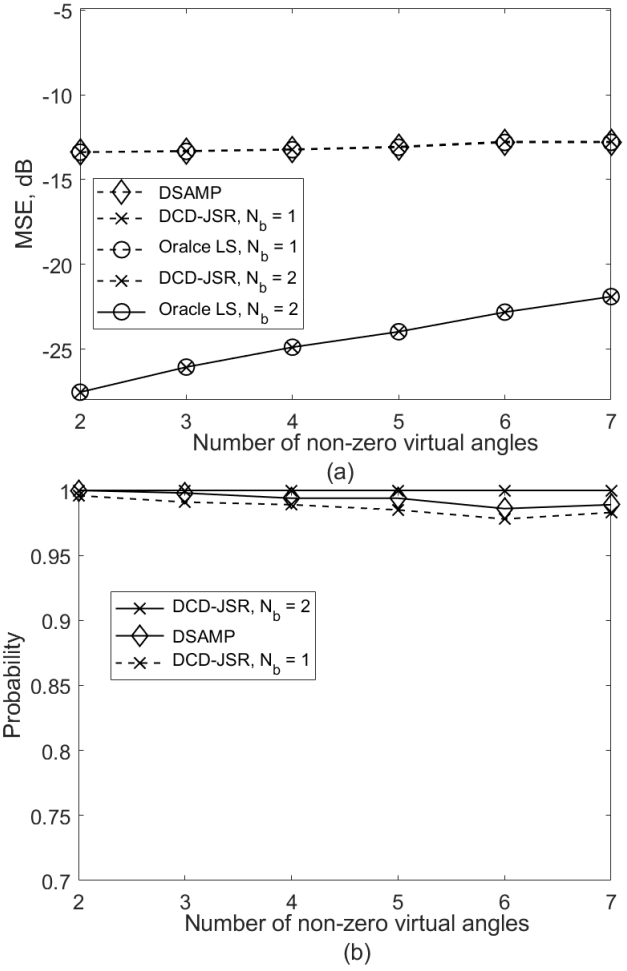

FIGURE 12. (a) MSE performance against the number of non-zero virtual angles; (b) Probability of perfect support estimation against the number of non-zero virtual angles. $f_dT = 0.02$, $J = 40$, SNR = 20 dB.

virtual angles, and the probability of perfect support estimation, respectively; here, the DCD-JSR algorithm with $N_b = 1$ corresponds to the version of the DCD-JSR algorithm previously proposed in [15] for time-invariant channels. For simulation scenario, we consider the normalized Doppler frequency $f_dT = 0.02$, $J = 40$, SNR $= 20$ dB. It can be seen that, in Fig. 12(a), when $f_dT = 0.02$, since the time variation of the channel is slow, we can estimate the channel quite well using only one basis function in the DCD-JSR algorithm or using the DSAMP algorithm, while both of them shows the MSE performance close to the oracle performance. The DCD-JSR algorithm with $N_b = 2$ also shows close to the oralce MSE performance. In Fig. 12(b), it is seen that for the DCD-JSR algorithm with $N_b = 2$, the support estimation is slightly better than that for the DCD-JSR algorithm with $N_b = 1$ and DSAMP algorithm, which explains why the DCD-JSR algorithm with $N_b = 2$ can provide a better MSE performance in this case.

In Fig. 13, a lower SNR is considered compared to Fig. 12, SNR $= 10$ dB. It can be seen that, when the noise level is higher, the MSE performance provided by the DCD-JSR

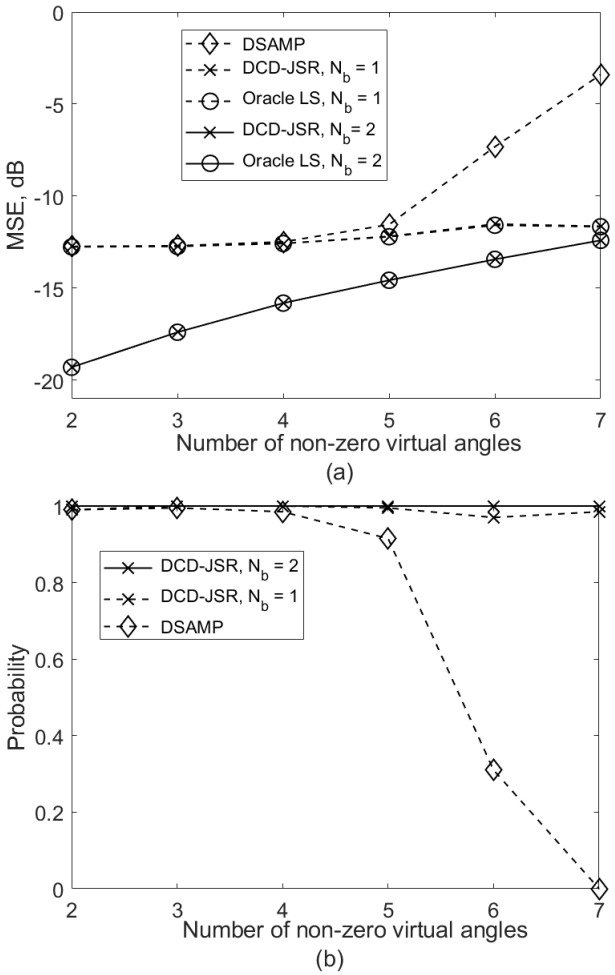

FIGURE 13. (a) MSE performance against the number of non-zero virtual angles; (b) Probability of perfect support estimation against the number of non-zero virtual angles. $f_dT = 0.02$, $J = 40$, SNR = 10 dB.

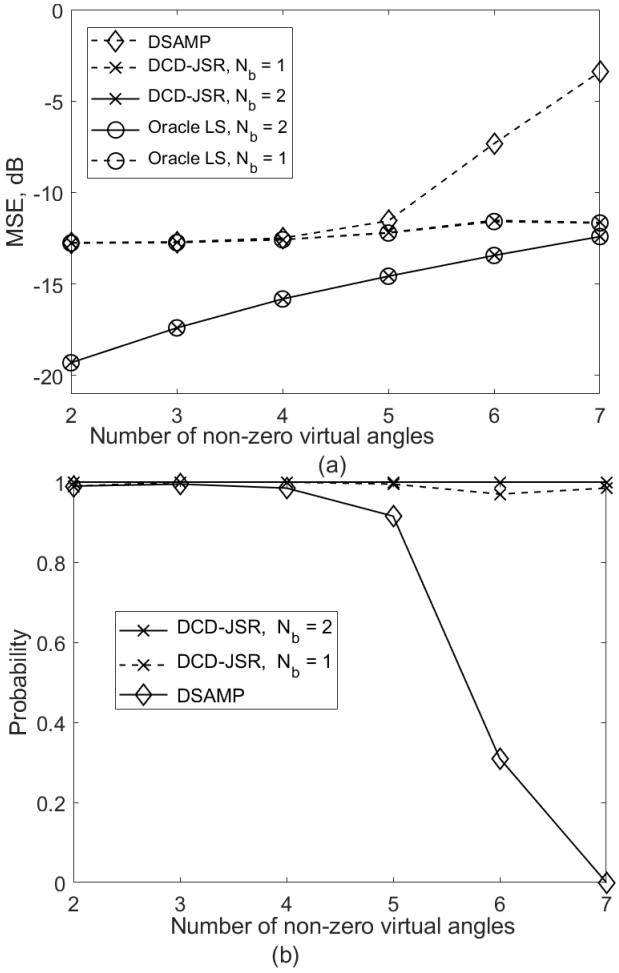

FIGURE 14. (a) MSE performance against the number of non-zero virtual angles; (b) Probability of perfect support estimation against the number of non-zero virtual angles. $f_dT = 0.05$, $J = 40$, SNR = 20 dB.

algorithm with $N_b = 2$ still shows close to the oracle performance and provides the perfect support estimation, whereas the DCD-JSR algorithm with a single basis function ($N_b = 1$) and the DSAMP algorithm, both developed for time-invariant channels, show inferior MSE performance and support estimation, although the DCD-JSR algorithm with $N_b = 1$ still outperforms the DSAMP algorithm.

In Fig. 14(a) and Fig. 14(b), for the DCD-JSR algorithm, and the DSAMP algorithm, we show the MSE performance for different number of non-zero virtual angles, and the probability of perfect support estimation, respectively. For simulation scenario, we consider the normalized Doppler frequency $f_dT = 0.05$, $J = 40$, and SNR = 20 dB. It can be seen in Fig. 14(a) that the DCD-JSR algorithm shows close to the oracle MSE performance, while the DSAMP algorithm shows a poor performance. In Fig. 14(b), it is seen that the DCD-JSR algorithm with $N_b = 3$ always provides the perfect support estimation, while the DCD-JSR algorithm with only one basis function $N_b = 1$ shows inferior performance, and the DSAMP algorithm cannot estimate the support accurately. This is because the DSAMP algorithm is developed for static channel, and when $f_dT = 0.05$, i.e. the time variation of the channel is fast, the algorithm is incapable of providing a high estimation performance.

In Fig. 15, results are shown for a higher noise level compared to Fig. 14, we set SNR = 10 dB. It can be seen in Fig. 15(a), that the DSAMP algorithm has again a poor MSE performance, while the DCD-JSR algorithm with $N_b = 1$ and $N_b = 3$ shows close to the oracle performance. In Fig. 15(b), it is clear that the DSAMP algorithm cannot provide an accurate support estimation in this case. The probability of perfect support estimation provided by the DCD-JSR algorithm with $N_b = 1$ decreses as the number of non-zero virtual angels increases, while the DCD-JSR algorithm with $N_b = 3$ can always provides the perfect support estimation. This is because as the time variation of the channel becomes faster, more basis functions is required to accurately approximate the channel.

Hence, from Fig. 12 to Fig. 15, we can conclude that the DCD-JSR algorithm outperforms the DSAMP algorithm. The improvement in the performance provided by the DCD-JSR algorithm against the DSAMP algorithm is more

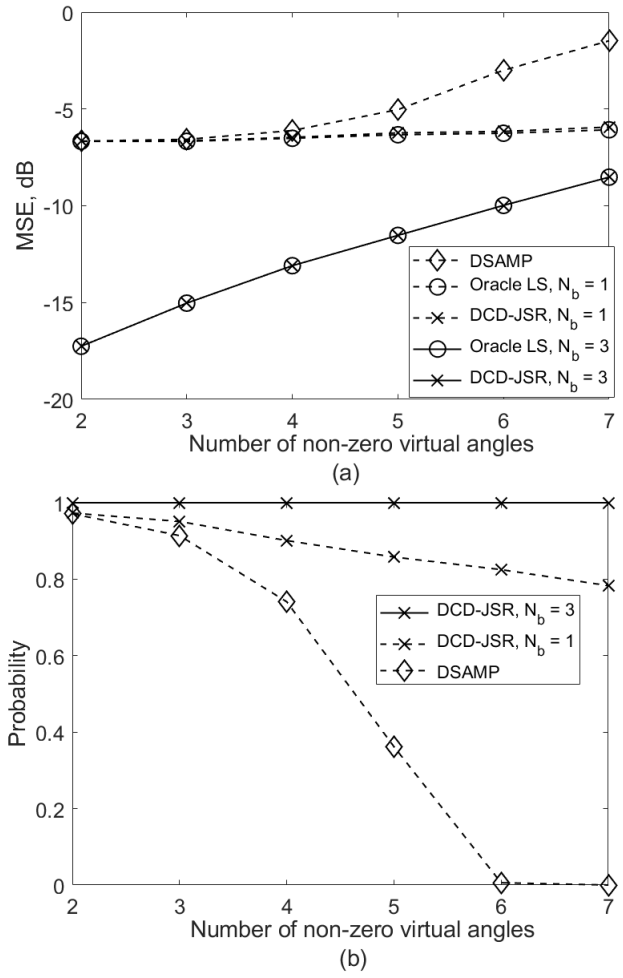

FIGURE 15. (a) MSE performance against the number of non-zero virtual angles; (b) Probability of perfect support estimation against the number of non-zero virtual angles. $f_dT = 0.05$, $J = 40$, SNR = 10 dB.

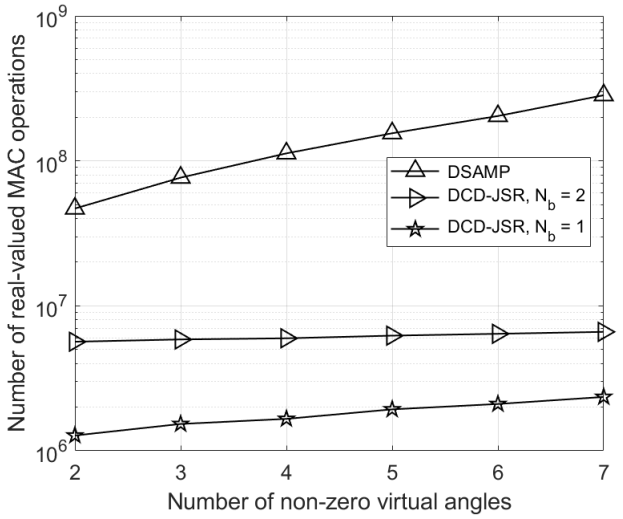

FIGURE 16. Computational complexity of the DCD-JSR algorithm and the DSAMP algorithm; $f_dT = 0.02$, $J = 20$, SNR = 20 dB.

significant in time-varying channels. For faster time varying channels, by employing more basis functions,we can significantly improve the performance of the DCD-JSR algorithm compared to the case $N_b = 1$ developed in [15] for static channels.

Fig. 16 shows the computational complexity of the DCD-JSR algorithm and DSAMP algorithm against the number of non-zero virtual angles, obtained for the case $f_dT = 0.02$, $J = 20$, SNR = 20 dB. It can be seen that the DCD-JSR algorithm has significantly lower computational complexity than the DSAMP algorithm. The DCD-JSR algorithm, when $N_b = 2$, has slightly higher computational complexity than the DCD-JSR algorithm with $N_b = 1$, while the increase in the number of basis functions provides a significantly better MSE performance.

VI. CONCLUSION

In this paper, by combining the BEM approach and the DCD-JSR algorithm, an efficient algorithm for estimation of the fast time-varying channels in virtual angular domain for massive MIMO systems is proposed. Simulation results have shown that compared to the previously proposed algorithms for channel estimation in virtual angular domain designed for time-invariant channels, the proposed DCD-JSR algorithm could provide significantly better MSE performance in time-varying channels.

REFERENCES

[1] J. Zhao, H. Xie, F. Gao, W. Jia, S. Jin, and H. Lin, "Time varying channel tracking with spatial and temporal BEM for massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 8, pp. 5653–5666, Aug. 2018.
[2] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.
[3] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, Jun. 2014.
[4] 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.
[5] Y. Zhou, M. Herdin, A. Sayeed, and E. Bonek, "Experimental study of MIMO channel statistics and capacity via the virtual channel representation," Univ. Wisconsin-Madison, Madison, WI, USA, Tech. Rep. 5, 2007, pp. 10–15.
[6] 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 Trans. Commun.*, vol. 64, no. 2, pp. 601–617, Feb. 2015.
[7] Z. Gao, L. Dai, Z. Wang, and S. Chen, "Spatially common sparsity based adaptive channel estimation and feedback for FDD massive MIMO," *IEEE Trans. Signal Process.*, vol. 63, no. 23, pp. 6169–6183, Dec. 2015.
[8] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
[9] S. F. Cotter and B. D. Rao, "Sparse channel estimation via matching pursuit with application to equalization," *IEEE Trans. Commun.*, vol. 50, no. 3, pp. 374–377, Mar. 2002.
[10] W. Li and J. C. Preisig, "Estimation of rapidly time-varying sparse channels," *IEEE J. Ocean. Eng.*, vol. 32, no. 4, pp. 927–939, Oct. 2007.
[11] G. Z. Karabulut and A. Yongacoglu, "Sparse channel estimation using orthogonal matching pursuit algorithm," in *Proc. IEEE 60th Veh. Technol. Conf., (VTC-Fall)*, vol. 6, Sep. 2004, pp. 3880–3884.
[12] Y. V. Zakharov, V. H. Nascimento, R. C. De Lamare, and F. G. D. A. Neto, "Low-complexity DCD-based sparse recovery algorithms," *IEEE Access*, vol. 5, pp. 12737–12750, 2017.
[13] Y. V. Zakharov and T. C. Tozer, "Multiplication-free iterative algorithm for LS problem," *Electron. Lett.*, vol. 40, no. 9, p. 567, 2004.
[14] J. Liu, Y. V. Zakharov, and B. Weaver, "Architecture and FPGA design of dichotomous coordinate descent algorithms," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 56, no. 11, pp. 2425–2438, Nov. 2009.

[15] M. Liao and Y. Zakharov, "DCD-based joint sparse channel estimation for OFDM in virtual angular domain," *IEEE Access*, vol. 9, pp. 102081–102090, 2021.
[16] Z. Tang, R. C. Cannizzaro, G. Leus, and P. Banelli, "Pilot-assisted time-varying channel estimation for OFDM systems," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 2226–2238, May 2007.
[17] H. Senol, E. Panayirci, and H. V. Poor, "Nondata-aided joint channel estimation and equalization for OFDM systems in very rapidly varying mobile channels," *IEEE Trans. Signal Process.*, vol. 60, no. 8, pp. 4236–4253, Aug. 2012.
[18] Z. Tang and G. Leus, "Time-multiplexed training for time-selective channels," *IEEE Signal Process. Lett.*, vol. 14, no. 9, pp. 585–588, Sep. 2007.
[19] K. A. D. Teo and S. Ohno, "Optimal MMSE finite parameter model for doubly-selective channels," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, vol. 6, Jan. 2005, p. 5.
[20] S. Tomasin, A. Gorokhov, H. Yang, and J. P. Linnartz, "Iterative interference cancellation and channel estimation for mobile OFDM," *IEEE Trans. Wireless Commun.*, vol. 4, no. 1, pp. 238–245, Jan. 2005.
[21] D. K. Borah and B. T. Hart, "Frequency-selective fading channel estimation with a polynomial time-varying channel model," *IEEE Trans. Commun.*, vol. 47, no. 6, pp. 862–873, Jun. 1999.
[22] X. Rao and V. K. N. Lau, "Distributed compressive CSIT estimation and feedback for FDD multi-user massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3261–3271, Jun. 2014.
[23] I. E. Telatar and D. N. C. Tse, "Capacity and mutual information of wideband multipath fading channels," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1384–1400, Jul. 2000.
[24] Q. Sun, D. C. Cox, H. C. Huang, and A. Lozano, "Estimation of continuous flat fading MIMO channels," in *Proc. IEEE Wireless Commun. Netw. Conf. Record. (WCNC)*, vol. 1, Mar. 2002, pp. 189–193.
[25] J.-J. van de Beek, O. Edfors, M. Sandell, S. K. Wilson, and P. O. Borjesson, "On channel estimation in OFDM systems," in *Proc. IEEE 45th Veh. Technol. Conf.*, vol. 2, Jul. 1995, pp. 815–819.
[26] L. Shen, Y. Zakharov, L. Shi, and B. Henson, "Performance of adaptive filtering based on legendre polynomials," in *Proc. IEEE Stat. Signal Process. Workshop (SSP)*, Jul. 2021, pp. 6–10.
[27] V. Savaux, F. Bader, and Y. Louët, "A joint MMSE channel and noise variance estimation for OFDM/OQAM modulation," *IEEE Trans. Commun.*, vol. 63, no. 11, pp. 4254–4266, Nov. 2015.
[28] T. Cui and C. Tellambura, "Power delay profile and noise variance estimation for OFDM," *IEEE Commun. Lett.*, vol. 10, no. 1, pp. 25–27, Jan. 2006.
[29] A. Das and B. D. Rao, "SNR and noise variance estimation for MIMO systems," *IEEE Trans. Signal Process.*, vol. 60, no. 8, pp. 3929–3941, Aug. 2012.

MINGDUO LIAO (Member, IEEE) received the B.Sc. degree from the University of Bristol, Bristol, U.K., in 2015. He is currently pursuing the Ph.D. degree in electronic engineering with the Communication Research Group, Department of Electronic Engineering, University of York. His research interest includes signal processing for massive MIMO communication systems.

YURIY ZAKHAROV (Senior Member, IEEE) received the M.Sc. and Ph.D. degrees in electrical engineering from the Power Engineering Institute, Moscow, Russia, in 1977 and 1983, respectively. From 1977 to 1983, he was with the Special Design Agency, Moscow Power Engineering Institute. From 1983 to 1999, he was with the N. N. Andreev Acoustics Institute, Moscow. From 1994 to 1999, he was a DSP Group Leader with Nortel. Since 1999, he has been with the Communications Research Group, University of York, York, U.K., where he is currently a Reader with the Department of Electronic Engineering. His research interests include signal processing, communications, and underwater acoustics.