Received June 16, 2021, accepted July 8, 2021, date of publication July 16, 2021, date of current version July 26, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3097899

DCD-Based Joint Sparse Channel Estimation for OFDM in Virtual Angular Domain

MINGDUO LIAO, (Member, IEEE), AND YURIY ZAKHAROV, (Senior Member, IEEE)
Department of Electronic Engineering, University of York, York YO10 5DD, 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 under Grant EP/V009591/1 and Grant EP/R003297/1.

ABSTRACT Massive Multiple Input Multiple Output (MIMO) is a promising technique for communications due to the high data transmission rate. To harvest the benefit from the massive MIMO, it is necessary to have accurate channel estimates. Such channels often exhibit sparsity in the virtual angular domain. This paper proposes a dichotomous coordinate descent (DCD) based algorithm for joint sparse channel estimation in the virtual angular domain for the orthogonal-frequency-division-multiplexing massive MIMO. We show that compared to the distributed sparsity adaptive matching pursuit algorithm previously proposed for this purpose, the DCD-based algorithm has significantly lower complexity and better channel estimation performance.

INDEX TERMS Channel estimation, common sparsity, compressive sensing, dichotomous coordinate descent, distributed sparsity adaptive matching pursuit, joint sparse recovery, massive MIMO, virtual angular domain.

I. INTRODUCTION

Massive MIMO has been proposed for next generations of communication systems, since it provides higher spectral efficiency [1], [2]. It can enhance the spectral efficiency by orders of magnitude by equipping the wireless transmitter with a large number of antennas and exploiting the increased degree of freedom in the spatial domain.

Pilot aided channel estimation is widely used in MIMO systems [3]. For channel estimation in a MIMO system with a small number of antennas, orthogonal pilots are often used [4], [5]. However, the pilot overhead increases with the number of antennas [6]. Employing orthogonal pilots for channel estimation would cause unacceptable pilot overhead because of the massive number of antennas at the base station (BS) [7]. In [7], a compressive sensing based channel feedback scheme was proposed, which can reduce the pilot overhead and achieve good channel state information (CSI) acquisition. In this paper, we focus on the channel estimation in the feedback scheme.

Experiments and research have shown that due to the small angle spread seen from a BS between a user and BS, massive MIMO channels exhibit sparsity in the virtual angular domain [8]. Furthermore, according to [6], [7], [9], when applying the orthogonal frequency division multiplexing (OFDM), because of the spatial propagation property of the wireless channel, such as the number of scatterers is nearly unchanged over the system bandwidth, the common sparsity is shared by different subcarriers, which is referred to as the spatially common sparsity over multiple subcarriers. Often, massive MIMO channels can be considered as quasi-static over a coherence time interval [9]. Furthermore, since the angle variation from the user to the BS is relatively slow, and can be often neglected, the support set of the channel in the virtual angular domain can be regarded as unchanged over several OFDM symbols, which is referred to as spatially common sparsity over multiple OFDM symbols [7], [9]. By exploiting the common sparsity in the virtual angular domain, we can jointly estimate the channel for multiple subcarriers.

Sparse recovery techniques are attractive for channel estimation [10]–[12]. There are two ways to find sparse representation, convex optimization and greedy methods [13]. Greedy methods typically have lower complexity [14], such as the orthogonal matching pursuit (OMP) [15], matching pursuit (MP) [14], compressive sampling matching pursuit (CoSAMP) [16]. However, they may provide limited performance when the signal is not very sparse or the noise

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

is too high [17]. Convex optimization algorithms such as Your ALgorithms for ℓ_1 (YALL1) [18], which employs the alternating direction method, provide high accuracy, but the complexity is high [13], [19], [20]. For channel estimation, we usually deal with complex-valued problems [13]. The sparse recovery algorithm used in this paper is for solving complex-valued problems.

The low-complexity coordinate descent (CD) search can be implemented to estimate the channel [21], [22]. In [13], algorithms applying dichotomous CD (DCD) iterations for solving $\ell_2\ell_0$ and $\ell_2\ell_1$ optimization problems have been proposed. By exploiting the DCD, the use of multiplications have been minimized, which significantly reduces the algorithm complexity and makes it well suited for real-time implementation [13]. Here we are interested in the DCD algorithm for the $\ell_2\ell_0$ optimization since it outperforms such greedy algorithms as MP and OMP [13].

The DCD algorithm for $\ell_2\ell_0$ optimization is a greedy algorithm [13], different from the CD algorithm [22], [23]. It does not optimize the step size for each iteration, but employs a set of step sizes defined by the fixed-point representation of the solution [13]. It has been indicated in [13] and [21], 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, which makes it suitable for implementation on real-time design platforms, such as digital signal processors and field-programmable gate arrays [24].

Since the DCD algorithm in [13] can only deal with single sparse channel at one time, by exploiting the spatially common sparsity in the virtual angular domain of the massive MIMO channels, a DCD-Joint-Sparse-Recovery (DCD-JSR) algorithm is proposed here. The DCD-JSR algorithm can jointly estimate multiple sparse channels and provide accurate CSI acquisition with a low computational complexity. Simulation results show that the proposed algorithm has better mean square error (MSE) performance than the Distributed-Sparsity-Adaptive-Matching-Pursuit (DSAMP) algorithm proposed in [7] for solving the same problem.

The paper is organized as follows. Section II describes the system model. Section III presents the proposed DCD-JSR algorithm. 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, and $j = \sqrt{-1}$, $(\mathbf{x})_n$ denotes the nth element of the vector $\mathbf{x}$, $\mathbf{R}^q$ denotes the qth column of the matrix $\mathbf{R}$, and $\mathbf{R}_n$ denotes the nth row of the matrix $\mathbf{R}$, $\mathbf{R}_{m,n}$ denotes an element of the matrix $\mathbf{R}$. The transpose operator is given by $(.)^T$, $(.)^*$ denotes the conjugate operator, $(.)^\dagger$ denotes the Moore-Penrose inversion, and $(.)^H$ 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, $|I|$ is the cardinality of the support I, I^c is the complement of I, $\mathbf{R}_I$ is a matrix obtained from $\mathbf{R}$, and which only contains columns corresponding to support I. $\mathbf{R}_{I,I}$ is an $|I| \times |I|$ matrix obtained from $\mathbf{R}$ by collecting elements from columns and rows corresponding to I, and $\mathbf{x}_I$ 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.

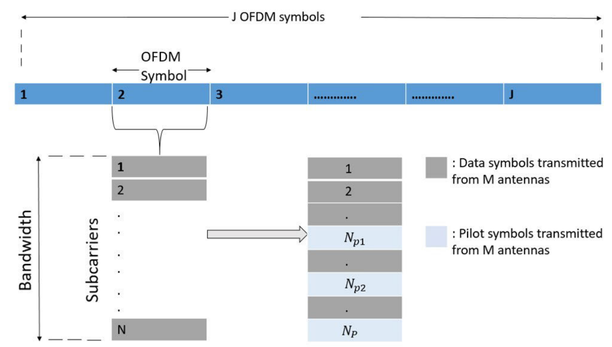

FIGURE 1. Each OFDM symbol contains N subcarriers, while P subcarriers are used to transmit pilot symbols.

II. SYSTEM MODEL AND PROBLEM FORMULATION

A. CHANNEL ESTIMATION SCHEME

The conventional method to acquire the 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 [25]. If we are employing conventional CSI estimation techniques (such as the minimum mean square error (MMSE) estimator), since the number of pilots required at the BS has to scale linearly with the number of transmit antennas at the BS [26], it would cause prohibitively large overhead for both pilot training (downlink) and CSI feedback (uplink). Hence, to solve the overhead issues, 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 $\mathbf{y}_k$ and feeds it back to the BS. The BS recovers the CSI for each user based on the feedback signals $\mathbf{y}_k$, $k = 1, \ldots, K$. As shown in Fig.1 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, a channel estimation algorithm is used to jointly estimate multiple sparse virtual angular domain channels, which are assumed to have the same support I. The least squares (LS) algorithm [27] is employed to acquire the CSI based on an estimate of the common support I.

B. CHANNEL MODEL

In a typical FDD massive MIMO system, consider a coherence time interval consisting of J OFDM symbols.

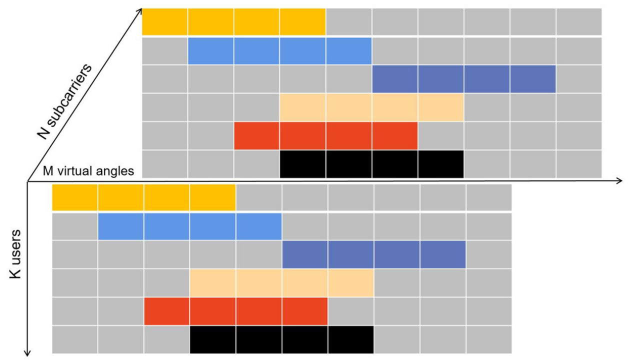

FIGURE 2. The virtual angular domain channel vector exhibits common sparsity within the system bandwidth (adapted from [7]).

M antennas are employed at the BS to serve K single-antenna users simultaneously, where $M \gg K$. At the tth OFDM symbol, $1 \leq t \leq J$, for the nth subcarrier, $1 \leq n \leq N$, the received signal for the kth user, $1 \leq k \leq K$, is given by:

$$y_{k,n}^t = \left(\mathbf{h}_{k,n}^t\right)^T \mathbf{x}_n^t + w_{k,n}^t, \tag{1}$$

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

$$y_n^t = \left(\mathbf{h}_n^t\right)^T \mathbf{x}_n^t + w_n^t. \tag{2}$$

Matrix $\mathbf{A}_B$ is used to modify the channel vector $\mathbf{h}_n^t$ into a vector $\tilde{\mathbf{h}}_n^t$ 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 $d = \lambda/2$, where λ is the wavelength, then $\mathbf{A}_B$ becomes the discrete Fourier transform (DFT) matrix. Thus we obtain:

$$y_n^t = \left(\tilde{\mathbf{h}}_n^t\right)^T \mathbf{A}_B^* \mathbf{x}_n^t + w_n^t, \tag{3}$$

where, $\left(\mathbf{h}_n^t\right)^T = \left(\tilde{\mathbf{h}}_n^t\right)^T \mathbf{A}_B^*$. As illustrated in Fig.2, the channel vector in the angular domain divides the covering area of the BS into angular intervals. The mth element of $\tilde{\mathbf{h}}_n^t$ corresponds to the mth virtual angle, where $1 \leq m \leq M$.

According to experimental study [8] and analysis [26], in practical massive MIMO systems, the BS is usually at a high elevation with a limited number of scatterers (relative to the number of antennas), and the scatterers at the user side are relatively rich. In other words, the BS might only have few active transmit directions for the kth user, which means that the number of multipath arrivals dominating the majority of channel energy is small, and the channel vectors in the virtual angular domain exhibit sparsity. Thus, we have $|I| \ll M$, which means the channel exhibits sparsity in the virtual angular domain. Furthermore, as shown in Fig.2, according to [9] and [7], since the spatial propagation characteristics such as scatterers are almost unchanged over the system bandwidth, the subchannels associated with different subcarriers in the same OFDM symbol share common sparsity. Moreover, in [28], it has been indicated that even in time-varying scenarios, the variation of the arrival angles is usually much slower than that of channel gains. This means, as shown in Fig.2, the channel associated with J successive OFDM symbols shares common sparsity. Since the channel during J OFDM symbols is time invariant, the channel gain can be considered as unchanged during J OFDM symbols, which can be written as:

$$\tilde{\mathbf{h}}_n^1 = \tilde{\mathbf{h}}_n^2 = \ldots = \tilde{\mathbf{h}}_n^J = \tilde{\mathbf{h}}_n. \tag{4}$$

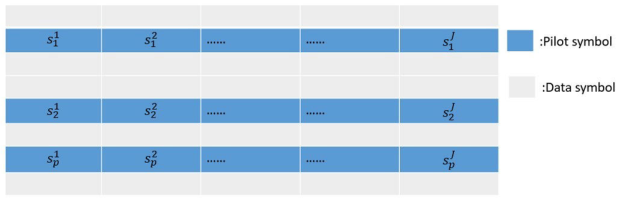

FIGURE 3. Structure of the transmitted *JP* pilot symbols. Each pilot symbol corresponds to the pilot sequence transmitted from *M* antennas.

In this paper, we consider the pilot-aided channel estimation. The structure of the transmitted pilot symbols is shown in Fig.3. To provide accurate channel estimation with multiple pilot subcarriers, for the tth OFDM symbol, a part of subcarriers is used for transmitting pilot symbols $\mathbf{s}_p^t \in C^{M \times 1}$, and the received signal at the pilot subcarrier $n(p)$ is given by:

$$y_{n(p)}^t = \left(\tilde{\mathbf{h}}_{n(p)}\right)^T \mathbf{A}_B^* \mathbf{s}_p^t + w_{n(p)}^t, \tag{5}$$

$$\left[\mathbf{s}_p^t\right]_m = e^{j\theta_{t,m,p}}, \quad 1 \leq p \leq P, \quad 1 \leq m \leq M, \ 1 \leq t \leq J \tag{6}$$

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

C. PROBLEM FORMULATION

As described in Section II-A, after receiving the signal from BS, the user will send the received signal back to the BS without performing the downlink channel estimation, where the feedback channel can be considered as an AWGN channel, and the variance can be neglected [26], [29], [30]. Hence, for the tth OFDM symbol, at the pth pilot subcarrier, the signal received at the BS is given by:

$$r_p^t = \boldsymbol{\phi}_p^t \tilde{\mathbf{h}}_{n(p)} + v_p^t, \quad 1 \leq p \leq P. \tag{7}$$

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

To provide an accurate channel estimation for the pth pilot subcarrier, the BS should jointly utilize the feedback signal over J successive OFDM symbols [7]. We collect the feedback signals $r_p^t, 1 \leq t \leq J$, 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

$$\mathbf{r}_p = \mathbf{\Phi}_p \tilde{\mathbf{h}}_{n(p)} + \mathbf{v}_p, \quad 1 \leq p \leq P, \tag{8}$$

where, $\mathbf{\Phi}_p = \left[\mathbf{S}_p^J \left(\mathbf{A}_B^*\right)^T\right]^T \in C^{J\times M}$, $\mathbf{S}_p = \left[\mathbf{s}_p^1, \mathbf{s}_p^2, \ldots, \mathbf{s}_p^J\right]^T \in C^{J\times M}$, 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, which contains both downlink and uplink noise. Since the channels for all subcarriers exhibit common sparsity, we can jointly estimate the channels associated with multiple pilot subcarriers assuming the common support.

III. DCD-JSR ALGORITHM FOR THE CHANNEL ESTIMATION IN VIRTUAL ANGULAR DOMAIN

In [7], the distributed sparsity adaptive matching pursuit (DSAMP) algorithm was proposed to jointly estimate multiple sparse channels by estimating the common support shared by different subcarriers in OFDM. However, simulation results show that it provides a limited performance when the number of OFDM symbols J used for the channel estimation is not high. In [13], the homotopy $\ell_2\ell_0$ DCD algorithm was proposed, which can be used to estimate the sparse channel, and it can provide accurate sparse estimation with low complexity. However, it was focused on a single sparse problem, and cannot jointly estimate multiple sparse channels. Therefore, based on [7] and [13], we propose the DCD-JSR algorithm, which can jointly estimate multiple sparse channels with a common support.

To simplify notation, we replace $\tilde{\mathbf{h}}_{n(p)}$ with $\mathbf{h}_p \in \mathbf{C}^{M\times 1}$, which is the channel vector to be estimated. We denote $\tilde{\mathbf{h}}_p$ as the final vector estimate. The DCD-JSR algorithm is summarized as follows.

1. For each pilot subcarrier, the $\ell_2\ell_0$ homotopy DCD algorithm is employed to acquire an estimate of $\mathbf{h}_p$.
2. Based on the $\mathbf{h}_p$ estimate, a common support $\tilde{I}$ is found by analysing the distribution of the estimates.
3. Based on the common support $\tilde{I}$, the final channel vector estimate $\tilde{\mathbf{h}}_p$ is acquired by using the LS algorithm [27] on the support.

A. CHANNEL ESTIMATION USING THE $\ell_2\ell_0$ HOMOTOPY DCD ALGORITHM

To estimate the channel at the pth pilot subcarrier using the $\ell_2\ell_0$ homotopy DCD algorithm, we consider the signal model

$$\mathbf{r}_p = \mathbf{\Phi}_p \mathbf{h}_p + \mathbf{v}_p. \tag{9}$$

It is worth to mention that since $\mathbf{h}_p$ is sparse in the virtual angular domain, only $|I|$ elements of the channel vector $\mathbf{h}_p$ are non-zero. We consider that the observation matrix $\mathbf{\Phi}_p$ is available and the support I is unknown.

Based on [13], we can find an estimate of $\mathbf{h}_p$ by applying the homotopy DCD algorithm to the $\ell_2\ell_0$ optimization, considering the minimization of the cost function

$$\mathbf{J}_\tau(\mathbf{h}_p) = \frac{1}{2}\left\|\mathbf{r}_p - \mathbf{\Phi}_p\mathbf{h}_p\right\|_2^2 + \tau\left\|\mathbf{h}_p\right\|_0. \tag{10}$$

Here, $\tau \in [0, 1)$ is a regularization parameter. The second term in (10) makes it 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 adding new elements into the support or removing elements from the support in several iterations following the proposition in [13], we can find an estimate of $\mathbf{h}_p$. Therefore we need to assign initially a high value to the regularization parameter $\tau = \tau_{\max}$ which can dominate 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 [13]. The algorithm stops when $\tau < \tau_{\min}$, where $\tau_{\min} = \mu_\tau\tau_{\max}$ and $\mu_\tau \in [0, 1)$ is a predefined parameter, and $\left(\mathbf{h}_p\right)_g$ is the gth element of the pth estimated channel vector $\mathbf{h}_p$. The structure of the employed $\ell_2\ell_0$ DCD homotopy algorithm is shown in Algorithm 1.

Algorithm 1 $\ell_2\ell_0$ Homotopy DCD Algorithm

Initialization:vector $\mathbf{h}_p = \mathbf{0}$, $I_p = \emptyset$, $\mathbf{b}_p = \mathbf{\Phi}_p{}^H\mathbf{r}_p$, $\mathbf{R}_p = \mathbf{\Phi}_p{}^H\mathbf{\Phi}_p$.

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

As shown in Algorithm 1, by solving the LS problem $\left(\mathbf{R}_p\right)_{I_p,I_p}\left(\mathbf{h}_p\right)_{I_p} = \mathbf{f}_p$ at step 3, $\mathbf{h}_p$ is estimated. According to [13], instead of using the matrix inversion to solve the LS problem, the DCD iterations [13], as shown in Algorithm 2, are employed at step 3 in Algorithm 1. When the DCD

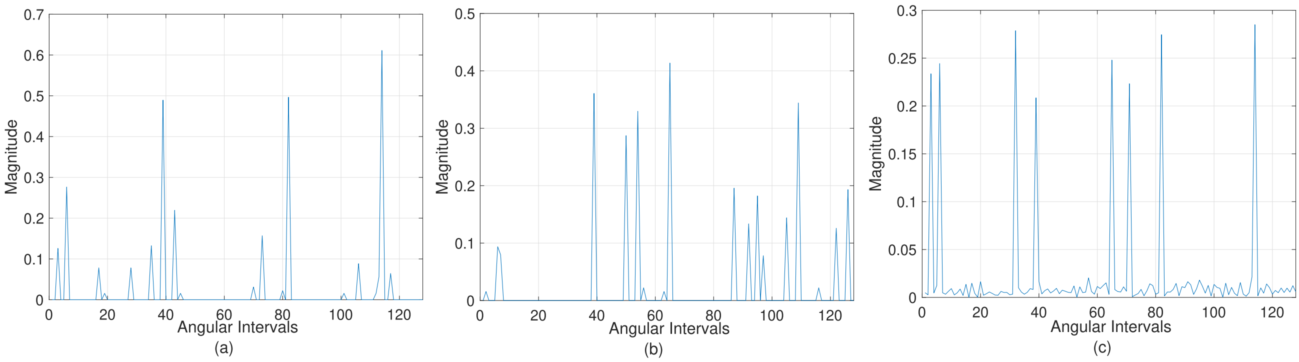

FIGURE 4. **Magnitudes of elements of vectors: (a) $\tilde{\mathbf{h}}_1$, (b) $\tilde{\mathbf{h}}_{64}$, (c) $\mathbf{q}$.**

iterations start, an LS solution for the vector $\mathbf{h}_p$ and the vector $\mathbf{c}$ found at the previous iteration are used as the initialization of the DCD algorithm, which results in the reduction of the computational complexity. In the DCD iterations, N_u is the maximum number of successful iterations and a successful iteration means that the solution is updated in the iteration, M_b and H are predefined parameters.

Algorithm 2 DCD Iterations for LS Minimization

Input: $\mathbf{h}_p$, $\mathbf{c}$, I_p, $\mathbf{R}_p$

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

B. COMMON SUPPORT ACQUISITION AND JOINT CHANNEL ESTIMATION

In this section, the process of estimating the common support I is presented. For example, we consider a scenario with $P = 64$ pilot subcarriers, $M = 128$ transmit antennas, signal to noise ratio SNR $= 20$ dB, $J = 20$ OFDM symbols and $|I| = 8$.

According to [7], among M coordinates of the channel vector $\mathbf{h}_p$, the vast majority of the channel energy will concentrate on $|I|$ coordinates, which are the non-zero elements in $\mathbf{h}_p$. Since we can estimate the channel at the pth pilot subcarrier using the $\ell_2\ell_0$ homotopy DCD algorithm, we can find an estimate of the common support $\tilde{I}$ by jointly analysing estimates $\tilde{\mathbf{h}}_p$ of vectors $\mathbf{h}_p$ for all pilot subcarriers.

In Fig.4(a) and Fig.4(b), magnitudes of elements of vectors $\tilde{\mathbf{h}}_1$ and $\tilde{\mathbf{h}}_{64}$ are shown. For estimation of the joint support, we compute

$$\mathbf{q} = \left(\sum_{p=1}^{P} \left| \tilde{\mathbf{h}}_p \right| \right) / P. \tag{11}$$

An estimate $\tilde{I}$ of the common support I is obtained using thresholding, as a set of elements in the vector $\mathbf{q}$, satisfying the condition

$$\tilde{I} = \{k : (\mathbf{q})_k > \xi\}, \tag{12}$$

where ξ is a predefined threshold parameter.

Based on the estimate $\tilde{I}$, the LS algorithm [27] is employed as follows:

$$\left(\mathbf{R}_p\right)_{\tilde{I},\tilde{I}} \left(\tilde{\mathbf{h}}_p\right)_{\tilde{I}} = \mathbf{f}_{\tilde{I}}, \tag{13}$$

$$\mathbf{f}_{\tilde{I}} = \left(\boldsymbol{\Phi}_p\right)_{\tilde{I}}^H \mathbf{r}_p. \tag{14}$$

Here, $\left(\tilde{\mathbf{h}}_p\right)_{\tilde{I}}$ is the final estimate of the channel vector $\mathbf{h}_p$ on the support $\tilde{I}$.

IV. DSAMP ALGORITHM

The DSAMP algorithm [7], which was developed from the sparsity adaptive matching pursuit algorithm [31], can acquire multiple sparse channel vectors for different pilot subcarriers simultaneously. The DSAMP algorithm has been shown to provide a better channel estimation performance than the orthogonal matching pursuit, sparsity adaptive matching pursuit and subspace pursuit algorithms [7]. We use the DSAMP performance as a benchmark to assess the performance of the proposed DCD-JSR algorithm.

V. SIMULATION RESULTS

A. MSE OF THE CHANNEL ESTIMATION

We will be assessing the algorithm performance using the mean square error (MSE) of the channel estimation. The MSE is given by

$$\text{MSE} = \frac{\left|\left|\mathbf{h}_p - \tilde{\mathbf{h}}_p\right|\right|_2^2}{\left|\left|\mathbf{h}_p\right|\right|_2^2}, \tag{15}$$

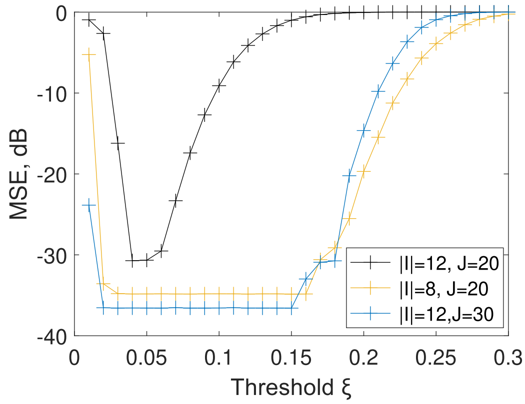

FIGURE 5. MSE performance of the DCD-JSR algorithm against the threshold ξ, SNR = 20 dB, the number of pilot subcarriers $P = 64$, $M = 128$.

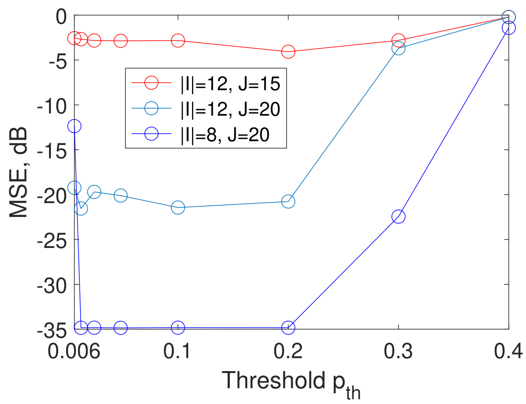

FIGURE 6. MSE performance of the DSAMP algorithm against the threshold p_{th}, SNR = 20 dB, the number of pilot subcarriers $P = 64$, $M = 128$.

$$\left\|\tilde{\mathbf{h}}_p\right\|_2 = \sqrt{\sum_{m=1}^{M}\left[\left(\tilde{\mathbf{h}}_p\right)_m\right]^2}. \qquad (16)$$

where $\tilde{\mathbf{h}}_p$ is the estimated channel vector and $\mathbf{h}_p$ is the true channel vector. When analysing the performance of the estimators, we will also calculate the probability of the estimated support $\tilde{I}$ to be exactly the same as the support I to be estimated.

B. NUMERICAL RESULTS

In this section, we consider simulation scenarios corresponding to a MIMO system with a uniform linear array. We compare the channel estimation performance of the DCD-JSR and DSAMP algorithms. 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.

To provide the best MSE performance, the threshold p_{th} for the DSAMP algorithm and ξ for the DCD-JSR algorithm need to be adjusted. As shown in Fig.5, when SNR = 20 dB, the DCD-JSR algorithm has the best MSE performance when $\xi = 0.055$. In Fig.6, it can be seen that when SNR = 20 dB and $p_{th} = 0.1$, the DSAMP algorithm achieves the best MSE performance. Similarly, appropriate values of ξ and p_{th} for different SNR can be obtained. In this paper, for the DCD-JSR algorithm, $\xi = 0.05$ is considered for both SNR = 20 dB and SNR = 10 dB; for the DSAMP algorithm, p_{th} is set to be 0.1 and 0.17 for SNR = 20 dB and SNR = 10 dB, respectively.

In Fig.7(a) and Fig.7(b), we consider scenarios with different number of pilot subcarriers. The number of pilot subcarriers varies from 48 to 64, and we set $M = 128$, $|I| = 12$, the number of simulation trials is $N_s = 10000$. It can be seen that both the DSAMP and DCD-JSR algorithms benefit from the increasing number of pilot subcarriers, but a larger number of subcarriers results in lower spectral efficiency, since a smaller number of subcarriers are used for data transmission. However, the DCD-JSR algorithm shows significantly better MSE performance.

Fig.8(a) and Fig.8(b), for different number of pilot subcarriers and different SNR, show the probability of the perfect support estimation by the DSAMP and DCD-JSR algorithms, where the perfect support estimation means that the estimated support is exactly the same as the true support. In Fig.8, it can be seen that, compared to the DSAMP algorithm, the DCD-JSR algorithm provides a better probability of correct support estimation. This explains the better MSE performance of the DCD-JSR algorithm, as seen in Fig.7. Compared to the DSAMP algorithm, the DCD-JSR algorithm requires less pilot subcarriers to provide a specified probability of correct support estimation under same scenario.

In Fig.9(a) and Fig.9(b), we show the MSE performance for scenarios with $J = 10$ and $J = 20$ at different SNR. We set $M = 128$, $P = 64$, and the number of simulation trials $N_s = 10000$. In Fig.9(a), for $J = 10$, at SNR = 10 dB, and $|I| \leq 6$, the DCD-JSR algorithm approaches the performance of the oracle LS algorithm [27], while the DSAMP does it only for $|I| \leq 4$. In Fig.9(b), for $J = 20$, when SNR = 10 dB, the DCD-JSR algorithm approaches the performance of the oracle LS algorithm [27] for $|I| \leq 13$, whereas the DSAMP algorithm does not show the LS performance even for $|I| = 10$. When SNR = 20 dB, the DCD-JSR algorithm could approach the oracle performance until $|I| = 13$, while the DSAMP does not. Hence, in these scenarios, the DCD-JSR algorithm outperforms the DSAMP algorithm.

Fig.10(a) and Fig.10(b) present results for different number of employed OFDM symbols J. The number of simulation trials is $N_s = 10000$, $M = 128$, $P = 64$. It can be seen that the DCD-JSR algorithm outperforms the DSAMP algorithm for both SNR = 20 dB and SNR = 10 dB, and requires less OFDM symbols to approach the performance of the oracle LS channel estimator.

Fig.11(a) and Fig.11(b) compare the probability of perfect support estimation by the DSAMP and DCD-JSR channel estimators. It can be seen that the DCD-JSR channel estimator outperforms the DSAMP channel estimator: at SNR = 20 dB, the DCD-JSR channel estimator needs $J = 28$ to provide the perfect support estimation, while the DSAMP algorithm needs $J = 34$, i.e., a lower number

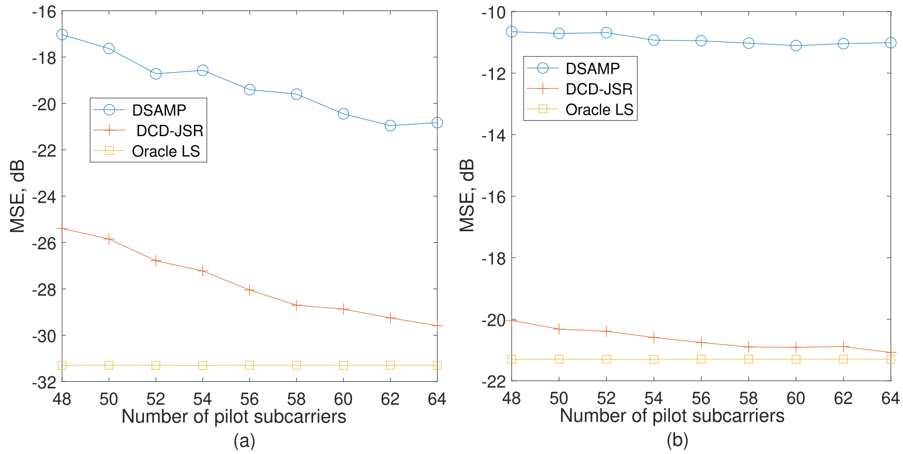

FIGURE 7. MSE performance of Oracle LS, DSAMP, and DCD-JSR algorithms against the number of pilot subcarriers, $M = 128$, $J = 20$: (a) SNR = 20 dB, (b) SNR = 10 dB.

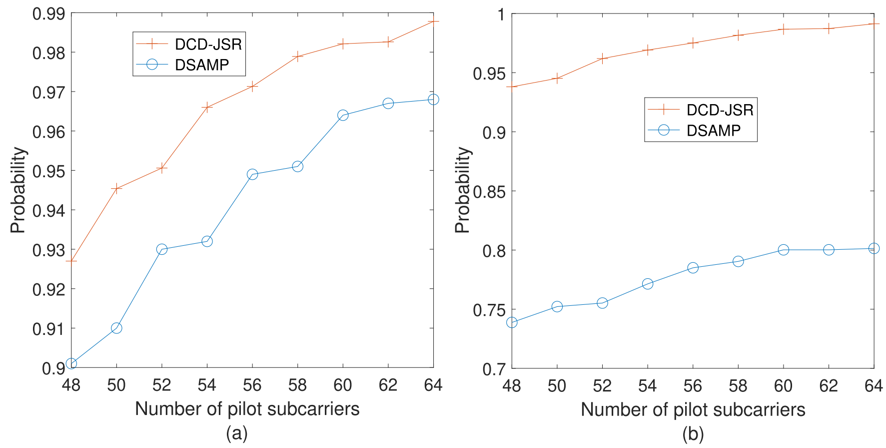

FIGURE 8. Probability of perfect support estimation for DSAMP and DCD-JSR algorithms against the number of pilot subcarriers, $M = 128$ $J = 20$: (a) SNR = 20 dB, (b) SNR = 10 dB.

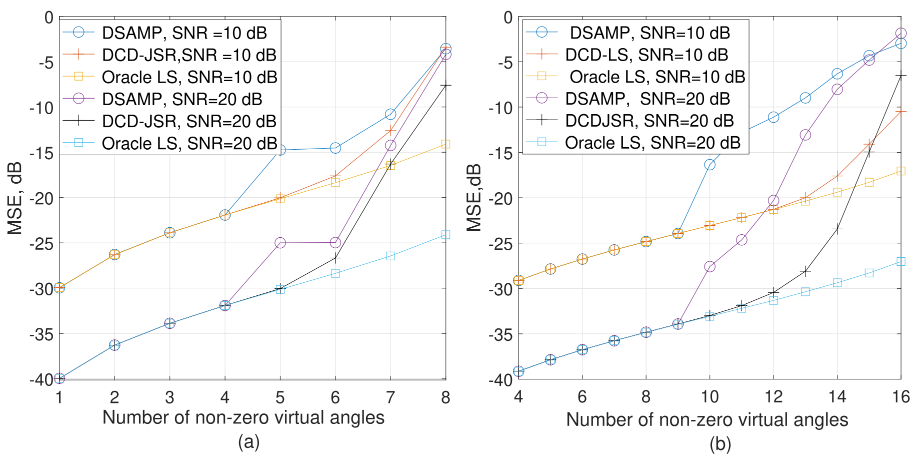

FIGURE 9. MSE performance of Oracle LS, DSAMP, DCD-JSR algorithms against the number of non-zero virtual angles $M = 128$, $P = 64$: (a) $J = 10$, (b) $J = 20$.

of OFDM symbols is required by the DCD-JSR algorithm. Thus, it is easy to see that, compared to the DSAMP channel estimator, the DCD-JSR channel estimator requires less OFDM symbols for an accurate support estimation.

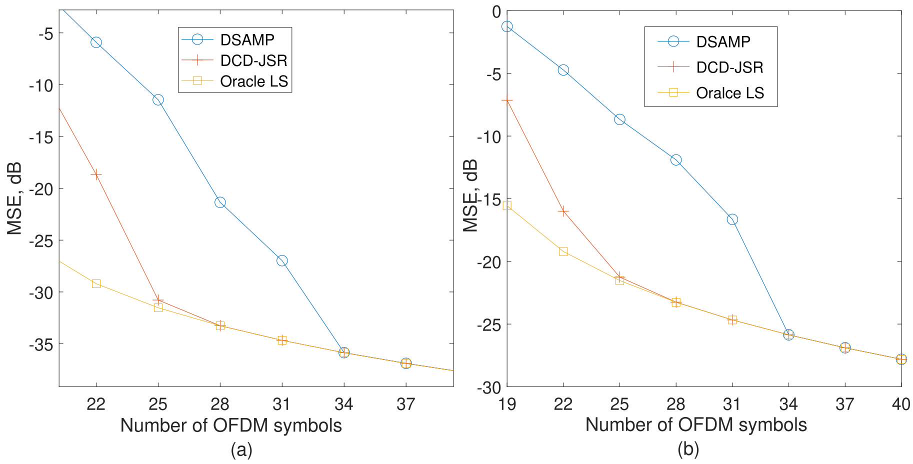

FIGURE 10. MSE performance of Oracle LS, DSAMP, and DCD-JSR algorithms against the number of OFDM symbols $M = 128$, $P = 64$, $|I| = 16$: (a) SNR = 20 dB, (b) SNR = 10 dB.

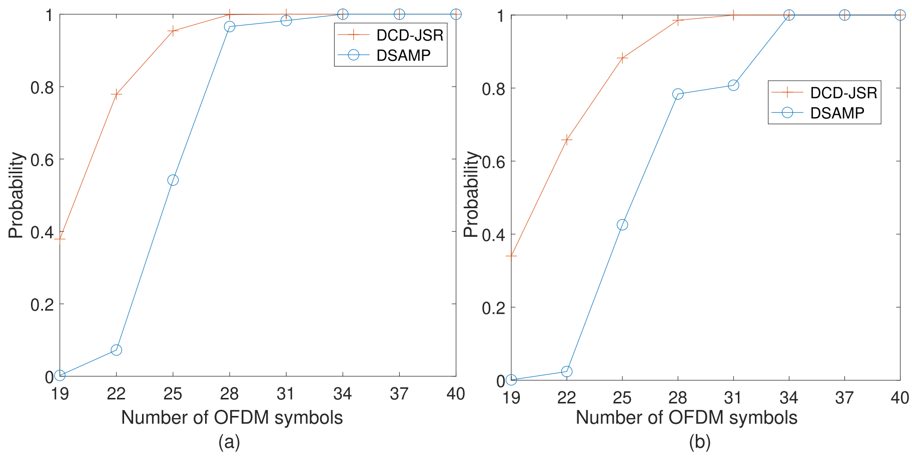

FIGURE 11. Probability of perfect support estimation for DSAMP and DCD-JSR algorithms against the number of OFDM symbols, $M = 128$, $P = 64$, $|I| = 16$: (a) SNR = 20dB, (b) SNR = 10 dB.

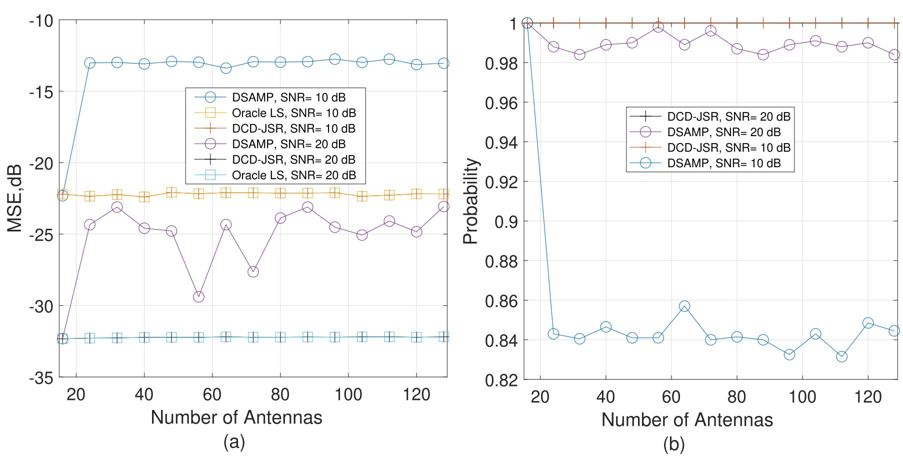

FIGURE 12. Performance of Oracle LS, DSAMP, and DCD-JSR algorithms against the number of antennas, $J = 20$, $P = 64$ (a) MSE. (b) Probability of perfect support estimation.

In Fig.12, we consider the case where the massive MIMO system employs different number of antennas. The number of antenna varies from 16 to 128, the number of simulation trials is $N_s = 10000$. We set the number of OFDM symbols $J = 20$

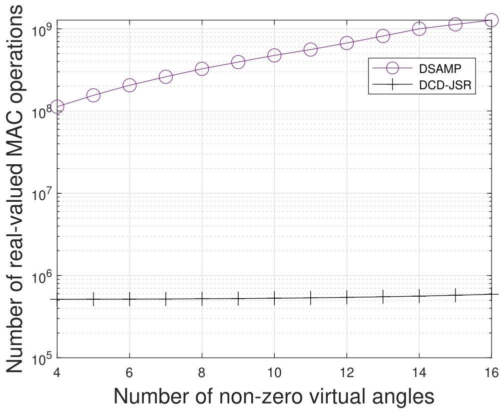

FIGURE 13. Computational complexity of the DSAMP algorithm and the DCD-JSR algorithm, $M = 128$, $J = 20$, $P = 64$, SNR = 20 dB.

and number of non-zero virtual angles $|I| = 11$. In Fig.12(a), it can be seen that when SNR = 10 dB, there exists a significant performance gap between the DSAMP algorithm and oracle LS algorithm [27], while the DCD-JSR algorithm approaches the oracle performance for any number of antennas. When we increase the SNR = 20 dB, the DCD-JSR channel estimator approaches the oracle performance for any number of antennas, while the DSAMP algorithm does not.

Fig.12(b) shows the probability of perfect support estimation in these scenarios. It can be seen that the DCD-JSR algorithm always provides perfect support estimation, while the DSAMP algorithm does not. Thus, we can see that with a large number of antennas, the DCD-JSR channel estimator provides a better MSE performance and more accurate support estimation than the DSAMP algorithm.

To estimate the computational complexity of the algorithms, we decided to update the computational complexity after each line of the algorithm code (both the algorithms have been implemented in Matlab) where an operation occurs. In the DCD-JSR algorithm, most of the operations are additions [13]; to simplify the comparison, we also count the pure additions as multiply-accumulate (MAC) operations.

Fig.13 shows the computational complexity against the number of non-zero virtual angles. We consider the SNR = 20 dB, $J = 20$ and average the results over $N_s = 10000$ simulation trials. It can be seen that the DCD-JSR algorithm has significantly lower complexity. Thus we can say that, compared to the DSAMP algorithm [7], the DCD-JSR algorithm exhibits lower computational complexity.

VI. CONCLUSION

In this paper, based on the original $\ell_2\ell_0$ DCD algorithm, a DCD-JSR algorithm has been proposed to jointly estimate the channel for multiple pilot subcarriers in the virtual angular domain in an FDD massive MIMO system. The DSAMP algorithm is used to compare the channel estimation performance with the DCD-JSR algorithm in different simulation scenario. Simulation results have shown that the proposed DCD-JSR algorithm outperforms the DSAMP algorithm, and requires less OFDM symbols and employed pilot subcarriers for accurate channel estimation, whereas it also exhibits a significantly lower computational complexity.

REFERENCES

[1] 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, Oct. 2014.

[2] Y. Xu, G. Yue, and S. Mao, "User grouping for massive MIMO in FDD systems: New design methods and analysis," *IEEE Access*, vol. 2, no. 1, pp. 947–959, Sep. 2014.

[3] Y. Liu, Z. Tan, H. Hu, L. J. Cimini, and G. Y. Li, "Channel estimation for OFDM," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 4, pp. 1891–1908, 4th Quart., 2014.

[4] D. Angelosante, E. Biglieri, and M. Lops, "Sequential estimation of multipath MIMO-OFDM channels," *IEEE Trans. Signal Process.*, vol. 57, no. 8, pp. 3167–3181, Aug. 2009.

[5] M. Simko, P. S. R. Diniz, Q. Wang, and M. Rupp, "Adaptive pilot-symbol patterns for MIMO OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 12, no. 9, pp. 4705–4715, Sep. 2013.

[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. 2016.

[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] Y. Zhou, M. Herdin, A. M. 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., Feb. 2007.

[9] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[10] 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.

[11] 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.

[12] 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.

[13] Y. V. Zakharov, V. H. Nascimento, R. C. De Lamare, and F. G. De Almeida Neto, "Low-complexity DCD-based sparse recovery algorithms," *IEEE Access*, vol. 5, pp. 12737–12750, Jun. 2017.

[14] P. Maechler, P. Greisen, N. Felber, and A. Burg, "Matching pursuit: Evaluation and implementatio for LTE channel estimation," in *Proc. IEEE Int. Symp. Circuits Syst.*, Paris, France, May 2010, pp. 589–592.

[15] F. Ren, R. Dorrace, W. Xu, and D. Marković, "A single-precision compressive sensing signal reconstruction engine on FPGAs," in *Proc. 23rd Int. Conf. Field Program. Log. Appl.*, Porto, Portugal, Sep. 2013, pp. 1–4.

[16] J. Lu, H. Zhang, and H. Meng, "Novel hardware architecture of sparse recovery based on FPGAs," in *Proc. 2nd Int. Conf. Signal Process. Syst.*, Dalian, China, vol. 1, 2010, p. V1-302.

[17] Y. C. Eldar, *Sampling Theory: Beyond Bandlimited Systems*. Cambridge, U.K.: Cambridge Univ. Press, 2015.

[18] Y. Zhang. (2009). *User's Guide for YALL1: Your a Lgorithms for ℓ_1 Optimization*. [Online]. Available: http://www.caam.rice.edu/optimization/

[19] J. Huang, C. R. Berger, S. Zhou, and J. Huang, "Comparison of basis pursuit algorithms for sparse channel estimation in underwater acoustic OFDM," in *Proc. IEEE OCEANS*, May 2010, pp. 1–6.

[20] G. F. Edelmann and C. F. Gaumond, "Beamforming using compressive sensing," *J. Acoust. Soc. Amer.*, vol. 130, no. 4, pp. EL232–EL237, Oct. 2011.

[21] Y. V. Zakharov and T. C. Tozer, "Multiplication-free iterative algorithm for LS problem," *Electron. Lett.*, vol. 40, no. 9, p. 567, 2004.

[22] J. Friedman, T. Hastie, H. Höfling, and R. Tibshirani, "Pathwise coordinate optimization," *Ann. Appl. Statist.*, vol. 1, no. 2, pp. 302–332, Dec. 2007.

[23] T. T. Wu and K. Lange, "Coordinate descent algorithms for lasso penalized regression," *Ann. Appl. Statist.*, vol. 2, no. 1, pp. 224–244, Mar. 2008.
[24] 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.
[25] 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. Rec. (WCNC)*, vol. 1, Mar. 2002, pp. 189–193.
[26] 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.
[27] 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.*, Chicago, IL, USA, vol. 2, Jul. 1995, pp. 815–819.
[28] 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.
[29] P. Cheng and Z. Chen, "Multidimensional compressive sensing based analog CSI feedback for massive MIMO-OFDM systems," in *Proc. IEEE 80th Veh. Technol. Conf. (VTC-Fall)*, Vancouver, BC, Canada, Sep. 2014, pp. 1–6.
[30] P.-H. Kuo, H. Kung, and P.-A. Ting, "Compressive sensing based channel feedback protocols for spatially-correlated massive antenna arrays," in *Proc. IEEE WCNC*, Paris, France, Apr. 2012, pp. 492–497.
[31] T. T. Do, L. Gan, N. Nguyen, and T. D. Tran, "Sparsity adaptive matching pursuit algorithm for practical compressed sensing," in *Proc. 42nd Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, USA, Oct. 2008, pp. 581–587.

MINGDUO LIAO (Member, IEEE) received the B.Sc. degree from the University of Bristol, Bristol, U.K., in 2015, and the M.Sc. degree from The University of Manchester, Manchester, U.K., in 2016. 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.