Massive MIMO-OFDM Channel Estimation via Distributed Compressed Sensing

Abbas Akbarpour-Kasgari and Mehrdad Ardebilipour

Abstract—**Massive multiple-input multiple-output orthogonal frequency division multiplexing (mMIMO-OFDM) channel estimation is considered recently utilizing compressed sensing (CS) based methods. Here, we proposed to use the joint sparsity of mMIMO-OFDM channels using stage-wise forward backward pursuit (StFBP) algorithm. In order to increase the speed of convergence and accuracy of estimation, we proposed to gather multiple good atoms in each step and to exploit common sparsity in the system model, respectively. Furthermore, the backward steps improve the accuracy by omitting bad previously gathered atoms. Simulation results represent the superiority of the proposed StFBP approach rather than the conventional CS-based and non-CS-based approaches.**

Index Terms—**Channel estimation, compressed sensing, massive multiple input multiple output-orthogonal frequency division multiplexing (mMIMO-OFDM), stage-wise forward backward pursuit (StFBP).**

I. Introduction

MASSIVE Multiple-Input Multiple-Output (mMIMO) is one the promising approaches for the future 5G telecommunication systems. By employing a large number of antennas in Base Station (BS), it facilitates the implementation of high-throughput wireless systems. In order to increase the data availability, accurate channel estimation is mandatory. Hence, researchers devote lots of their studies on the channel estimation in MIMO-OFDM systems, recently [1]–[6].

The sparse behavior of the mMIMO-Orthogonal Frequency Division Multiplexing (OFDM) channel ensembles is utilized in Compressed Sensing (CS)-based channel estimation methods which are the result of a small number of significant scatterers in the medium. Moreover, the small distance of the antenna pairs on the mMIMO node rather than the considerable conveyed distance by the wave leads to unique and common support for all the channel ensembles between two mMIMO nodes. Hence, using common sparsity channels, the compressed channel estimation is improved significantly [1]–[8]. A channel estimation scheme based on training sequence (TS) design and optimization with high accuracy and spectral efficiency is investigated in the framework of structured compressive sensing in [2]. Channel estimation for indoor mMIMO systems is decomposed in angular domain and a unified channel estimation for TDD and FDD system is proposed in [3]. Moreover, to reduce the required pilot and accurate channel estimation in mMIMO systems, exploiting spatio-temporal sparsity, an adaptive structured subspace pursuit (ASSP) is introduced in [1]. ASSP algorithm estimates the sparsity of the channel and channel coefficients, simultaneously. Choi *et al.* [5] utilize the temporal correlation of the channels to develop the locally common support (LCS) algorithm for channel estimation. The Bayesian estimation was used via marginal based SCS method in order to improve the channel estimation accuracy in [6]. Stage-wise OMP (StOMP) was utilized in block sparsity mode to estimate the channel state information of MIMO-OFDM rapidly and accurately in [7]. In each step of OMP algorithm, the estimation is based on just one of the atoms while in StOMP the estimation works using multiple atoms. Hence, StOMP based algorithms converge more rapid than the former. The OMP and its derivatives the same as StOMP suffers from an inherent drawback which is caused by not eliminating bad atoms which are gathered in previous steps. This drawback is compensated in Forward-Backward Pursuit (FBP) algorithm introduced in [9].

In this letter, we have proposed to utilize the common sparsity together with the fading channel large scale behavior. First of all, we have considered all the channel ensembles common sparsity by using the μ-norm. Then, in order to deploy the channel large scale behavior, the channel is modeled by a weighted cost function based on the distance of the channel coefficients from the origin. Hence, the channel coefficients will be extracted from the random measurements optimally. In order to solve the resultant optimization problem which address the common sparsity and large scale behavior, we proposed a Stage-wise FBP (StFBP) algorithm mode wherein forward steps multiple good atoms are gathered while in the conventional FBP method only one atom is selected in each forward step. Consequently, the proposed method increases the speed of convergence. Moreover, exploiting backward steps makes it possible to omit bad atoms gathered previously, as a consequence it would improve the accuracy of estimation rather than StOMP and OMP. To exploit the proposed StFBP, we design the channel matrix and measurement matrix to exploit the common sparsity using μ-norm which is introduced. The system model is formulated using matrix representation to exploit the common sparsity of the channel. Then, a StFBP algorithm-based channel estimation is developed to estimate the channel coefficients in the time domain.

The remainder of this letter is as follows. In Section II the system model is represented, and the problem is formulated.

Manuscript received September 3, 2018; accepted September 21, 2018. Date of publication October 1, 2018; date of current version April 9, 2019. The associate editor coordinating the review of this paper and approving it for publication was J. Choi. *(Corresponding author: Mehrdad Ardebilipour.)*

The authors are with the Department of Electrical and Computer Engineering, K. N. Toosi University of Technology, Tehran 19697, Iran (e-mail: mehrdad@eetd.kntu.ac.ir).

Digital Object Identifier 10.1109/LWC.2018.2873339

Section III the proposed channel estimation method is introduced. Numerical results and Concluding remarks are represented in Sections IV and V, respectively.

II. System Model and Problem Formulation

A. System Model

Consider a mMIMO BS $\mathcal{A}$ and a user terminal (UT) $\mathcal{B}$ which are equipped by $N_{\mathcal{A}}$ and $N_{\mathcal{B}}$ transmit-receive antennas, respectively. Each terminal employs OFDM signaling on each of *u*-th antenna with $u = 0, 1, \ldots, N_{\mathcal{A}}(N_{\mathcal{B}}) - 1$ to exchange data with the other terminal. The emitted OFDM symbol is constructed using *N* subcarriers which are spaced by Δf in hertz while prefaced using $N_{\mathcal{G}}$ cyclic prefix (CP). The received signal by *v*-th antenna of the receiver with $v = 0, 1, \ldots, N_{\mathcal{B}}(N_{\mathcal{A}}) - 1$, is passed through a frequency selective channel which can be determined by a Finite Impulse Response (FIR) filter and constructed of multiple paths. $T_s = 1/(N\Delta f)$ is used as the sampling period in the receiver. Ignoring the CP fallen chips and transforming other chips into Fourier transform domain, the received signal is ready to be selected on the pilot subcarriers. This selection is made on each of the antennas and respective pilot sequence. As a consequence, the received signal on the *n*-th pilot symbol can be derived as

$$y_{uv}(n) = \sum_{l=0}^{L-1} h_{uv}(l)e(l,u,n) + w_{uv}(n) \tag{1}$$

where $n = 0, 1, \ldots, N_p - 1$ is the pilot index, $l = 0, 1, \ldots, L - 1$ is the paths index and $e(l,u,n) = e^{-j2\pi l p_u(n)/N}$. Moreover, $p_u(n)$ represents the *n*-th pilot of *u*-th transmit antenna with $n = 0, 1, \ldots, N_p$. $w_{uv}(n)$ denotes the received Additive White Gaussian Noise (AWGN) sample between the *u*-th transmit antenna, and *v*-th received antenna in the *n*-th sample of pilot which obeys the $\mathcal{CN}(0, \sigma_w^2)$. Besides, $h_{uv}(l)$ is the *l*-th channel tap between the *u*-th transmitter antenna, and *v*-th received antenna. The vector representation of (1) can be formulated as

$$\mathbf{y}_{uv} = \mathbf{\Phi}_u \mathbf{h}_{uv} + \mathbf{w}_{uv} \tag{2}$$

where $\mathbf{y}_{uv} = [y_{uv}(0), y_{uv}(1), \ldots, y_{uv}(N_p - 1)]^T$ is the received pilot vector, $\mathbf{\Phi}_u = \text{diag}\{\mathbf{x}_u\}\mathbf{F}_u$ is the measurement matrix with $\mathbf{x}_u = [x_u(0), x_u(1), \ldots, x_u(N_p - 1)]^T$ as the transmitted pilot vector and $\mathbf{F}_u$ as the partial Fourier matrix with N_p rows corresponding to the pilot sequence of transmit antenna *u* and *L* first columns, $\mathbf{h}_{uv} = [h_{uv}(0), h_{uv}(1), \ldots, h_{uv}(L-1)]^T$ is the channel coefficient vector and $\mathbf{w}_{uv} = [w_{uv}(0), w_{uv}(1), \ldots, w_{uv}(N_p - 1)]^T$ is the received noise vector. Besides, in this formulation the superscript ^{T} is the sign of transpose and diag$\{.\}$ denotes the diagonal representation of the corresponding vector.

B. Channel Model

As mentioned, the channel between transmit-receive pair which is denoted by $\mathbf{h}_{uv}$ is consist of *L* resolvable paths. These resolvable paths result from *L* scatterers which are encountered by the signal conveying from the transmitter antenna *u* to the receiver antenna *v*. *K* of these *L* scatterers are significant scatterers where $K << L$. Consequently, the channel could be modeled using sparse vectors. Moreover, the signal conveyed distance is large relative to the transmit-receive antenna spacing in each terminal. Hence, the encountered scatterers in each chip period are identical between different antennas. In other words, the delays of different paths are the same in all the channel ensembles between two terminals. Thus, the sparsity pattern of varying channel pairs could be assumed to be the same while the channel attenuation is different. Since each path channel attenuation is composed of multiple distinct scatterers with zero-mean and identically independent distributed (i.i.d.) subpaths, it is assumed to be $\mathcal{CN}(0, \sigma^2)$. Hence, the channel coefficient could be represented as

$$h_{uv}(l) = \sum_{i=0}^{I-1} \alpha_{uv}(i) g(lT - \tau(i)) \tag{3}$$

where $l \in [0, 1, \ldots, L]$ is the channel path index, $\tau(I-1) \geq \cdots \geq \tau(1) \geq \tau(0)$ are the respective paths' delay, α_{uv} is the corresponding paths' gain, and *g*(.) is the shaping pulse in the continuous domain. The shaping pulse is zero outside the interval [0, T_g], where T_g is the integer multiple of chip time *T*. Without loss of generality, we assumed that $\tau(i)$ are integer multiples of *T*. Thus, the number of channel paths, caused by channel itself and shaping filter is derived by $L = \tau(I-1)/T + T_g/T + 1$. Furthermore, we assume that *L* is lower than $T_{\mathcal{G}}/T$. Using the mentioned notations, we can represent the channel impulse response using $\mathbf{h}_{uv} = [h_{uv}(0), h_{uv}(1), \ldots, h_{uv}(L-1)]^T$.

C. Problem Formulation

Up to now, the channel model and the system model are represented. In order to estimate the channel impulse response, we have developed a formulation to utilize the channel common sparsity together with the large scale characteristic. To develop the model we consider Multiple Input and Single Output (MISO) since it is practical according to the mMIMO systems. As mentioned earlier, the channel impulse response could be represented as $\mathbf{h}_{uv}$ wherein MISO channels $u = 0, 1, \ldots, N_{\mathcal{B}} - 1$, and $v = 0$. Using the channel ensembles in MISO system, we can represent the channel matrix as $\mathbf{H} = [\mathbf{h}_{00}, \mathbf{h}_{10}, \ldots, \mathbf{h}_{(N_{\mathcal{A}}-1)0}]$ where the size of $\mathbf{H}$ is $L \times N_{\mathcal{A}}$. By collecting all the received pilot sequences, we can represent the $N_p \times N_{\mathcal{A}}$ received pilot matrix as $\mathbf{Y} = [\mathbf{y}_{00}, \mathbf{y}_{10}, \ldots, \mathbf{y}_{(N_{\mathcal{B}}-1)0}]$. Hence, the extension of (2) could be represented in matrix form for MISO case as

$$\mathbf{Y} = \mathbf{\Phi}\mathbf{H} + \mathbf{W} \tag{4}$$

where $\mathbf{W}$ is the AWGN matrix with corresponding columns according to the $\mathbf{w}_{u0}$. $\mathbf{\Phi}$ is the measurement matrix with the size of $N_p \times L$. For the sake of simplicity, we drop the index 0 for the UT antenna. Defining μ norm for matrices as $\mu(\mathbf{H}) = \text{card}\{\|\mathbf{H}_u\|_2 \neq 0\}$ where $\mathbf{H}_u$ is the *u*-th column of $\mathbf{H}$, and card$\{S\}$ is the number of elements in the set *S*, estimating the channel could be accomplished by following optimization criterion.

$$\begin{aligned} \min \quad & \|\mathbf{Y} - \mathbf{\Phi}\mathbf{H}\|_2^2 \\ s.t. \quad & \mu(\mathbf{H}) \leq K. \end{aligned} \tag{5}$$

Apparently, K is the maximum sparsity of the columns of $\mathbf{H}$. We called the objective function as $\mathbf{F}(\mathbf{H}) = \|\mathbf{Y} - \mathbf{\Phi}\mathbf{H}\|_2^2$. Utilizing $\mu(\mathbf{H})$, we can exploit the joint sparsity of the channel ensembles in the system. The objective function in (5) represents the error of channel estimation method, and the constraint controls the sparsity order of the channel ensembles in $\mathbf{H}$. Moreover, using μ-norm definition, the block sparsity of the channels are exploited.

III. Proposed Stage-Wise Forward-Backward Pursuit

To solve the optimization problem in Eq. (5), pseudo-inverse of the matrix $\mathbf{\Phi}$ is needed. Calculating the channel coefficients without any consideration of channel characteristics, would cause some difficulties in inversion of the matrix. Actually, this inversion didn't consider any characteristics of channels. In order to consider the channel impulse response, we have considered the large scale fading phenomena. Each path of the channel is changed in the power, according to the exponential distribution based on the distance of the origin. In fact, each path could be defined by the following equation

$$Z(H) = \sum_{l=0}^{L-1} \|\mathbf{h}_l\|_2^2 \omega_l \tag{6}$$

where

$$\omega_l = \begin{cases} 1, & \text{if } l = 0 \\ \frac{\sigma}{l^\alpha}, & \text{otherwise} \end{cases} \tag{7}$$

where $0 \le l \le L$-1, σ is the coefficient of the pathloss model and α is the environment factor. Accordingly, by defining the new vector $\mathbf{g}_l = [h_1(l), h_2(l), \dots, h_{N_\mathcal{A}}(l)]$ and consequently the matrix $\mathbf{G} = [\mathbf{g}_0^T, \mathbf{g}_1^T, \dots, \mathbf{g}_{L-1}^T]^T$, the $Z(H)$ could be considered as the

$$\mathbf{Z}(\mathbf{H}) = \mathrm{Tr}(\mathbf{G}\mathbf{\Omega}\mathbf{G}^H) \tag{8}$$

where $\Omega = \mathrm{diag}\{\omega_0, \omega_1, \dots, \omega_{L-1}\}$. Thus, in order to consider the channel ensembles, we should change the optimization problem in Eq. 5 as

$$\begin{aligned} \min \quad & \|\mathbf{Y} - \mathbf{\Phi}\mathbf{H}\|_2^2 + \lambda \mathrm{Tr}(\mathbf{G}\mathbf{\Omega}\mathbf{G}^H) \\ s.t. \quad & \mu(\mathbf{H}) \le K \end{aligned} \tag{9}$$

where λ is the regulation factor. In order to solve the problem, we try to consider the gradient of the optimization problem equal to zero.

$$\mathbf{\Phi}^H \mathbf{\Phi}\mathbf{H} + \lambda/2\mathbf{H}\mathbf{\Omega} = \mathbf{\Phi}^H \mathbf{Y} \tag{10}$$

which is a Lyapunov equation [10]. Hence, by solving a Lyapunov equation, we can calculate the channel coefficients.

In order to handle the optimization in (9), we have proposed a Stage-wise Forward-Backward Pursuit (StFBP) based on [9] where ℓ_0 norm was used. The algorithm is represented in details in Algorithm 1. Eq. (9) could be solved using three different methods as convex relaxation, greedy processes, and message-passing algorithms. StFBP which is based on the Message-Passing algorithm is used here because of its forward selection and backward fixing. Specifically, StOMP which is a greedy algorithm is a particular case of StFBP which the forward selection is present, but the backward fixing is absent, consequently, it cannot fix its own mistakes in previous steps. Moreover, FBP algorithm constructs the new subspace by adding just one atom to the previous subspace, and in the backward steps, it reconstructs the subspace by omitting bad atoms. To increase the speed of convergence, we have proposed to add multiple elite atoms in each forward step and in the backward step it excludes various atoms, too. In this case, the speed of convergence would increase, while the accuracy is guaranteed using backward steps. As a consequence, the proposed StFBP algorithm could be compared with its greedy one called StOMP, where StOMP is the particular case of StFBP without backward steps to increase the estimation accuracy.

Algorithm 1 StFBP-Based Channel Estimation

1: $\mathbf{R}^{(0)} = \mathbf{Y}$, $\lambda^0 = \phi$ and $t = 0$
2: **while** stop criterion not met **do**
3: Proxy of the signal is formed by $\mathbf{P} = \mathbf{\Phi}^H \mathbf{R}^{(t-1)}$
4: $\mathbf{r} = \mathbf{P}\theta$
5: Select good atoms according to $\{i^{(t)} : r_i \ge \tau\}$
6: Merging supports of the previous iteration and the present one $\lambda^{(t)} = \lambda^{(t-1)} \cup i^{(t)}$
7: Calculate $\mathbf{H}^{(t)}$ by solving Eq. (10)
8: $\delta_F^{(t)} = \mathbf{F}(\mathbf{H}^{(t-1)}) - \mathbf{F}(\mathbf{H}^{(t)})$
9: **while** 1 **do**
10: $j^{(t)} = \arg\min_{j \in \lambda^{(t)}} \mathbf{F}(\mathbf{H}^{(t)} - \mathbf{H}_{\not{j}}^{(t)})$
11: $\delta_B^{(t)} = \mathbf{F}(\mathbf{H}_{\not{j}}^{(t)}) - \mathbf{F}(\mathbf{H}^{(t)})$
12: **if** $\delta_B^{(t)} \ge 0.5\delta_F^{(t)}/|i^{(t)}|$ **then**
13: Update residual $\mathbf{R}^{(t)} = \mathbf{Y} - \mathbf{\Phi}_{\lambda^{(t)}} \mathbf{H}^{(t)}$
14: break
15: **end if**
16: Exclude bad atom $\lambda^{(t)} = \lambda^{(t)} - j^{(t)}$
17: Update $\mathbf{H}^{(t)}$ by solving Eq. (10).
18: Update residual $\mathbf{R}^{(t)} = \mathbf{Y} - \mathbf{\Phi}_{\lambda^{(t)}} \mathbf{H}^{(t)}$
19: **end while**
20: **end while**

In Algorithm 1, $|i^{(t)}|$ denotes the number of selected atoms in forwarding step, and $\mathbf{H}_{\not{j}}^{(t)}$ represents the $\mathbf{H}^{(t)}$ while $j^{(t)}$-th column of it is omitted. Furthermore, τ is the threshold of selection in forwarding steps and θ is $N_\mathcal{B} \times 1$ ones vector (i.e., all the elements equal to one).

In algorithm 1 each step is included of different multiplications which will be mentioned in the following. The proxy matrix costs $4N_\mathcal{A}N_p$ real multiplications. Moreover, In 9th step $4N_\mathcal{A}N_p$ real multiplications is encountered. In each step t the cardinality of $\boldsymbol{\lambda}^{(t)}$ could be defined by $|\boldsymbol{\lambda}^{(t)}|$; hence, the overall cost of 10^{th} step is evaluated by $4|\boldsymbol{\lambda}^{(t)}| + 4N_\mathcal{B}$. In backward stages, calculating $j^{(t)}$ and $\delta_B^{(t)}$ cost $4(|\boldsymbol{\lambda}^{(t)}| - 1) + 4N_\mathcal{B}$. Furthermore, defining residual in 15^{th} and 20^{th} step costs $4|\boldsymbol{\lambda}^{(t)}|$ real multiplications. Finally, updating $\mathbf{H}^{(t)}$ in each step consumes $4N_\mathcal{A}N_p$ real multiplications.

IV. Numerical Results

The simulation results are represented in this section to compare the performance of the proposed method with other

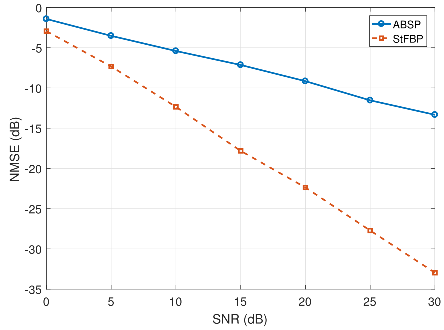

Fig. 1. Comparison of different estimation algorithms in using $N_p = 32$ pilots and $N_{\mathcal{A}} = 20$ antennas from NMSE view point.

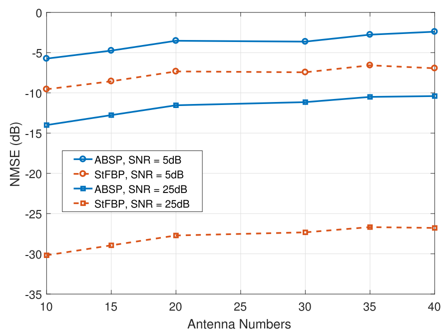

Fig. 2. Comparison of different number of antennas in BS by using $N_p = 32$ pilots from NMSE view point.

known approaches. In the simulations, the number of subcarriers were $N = 2048$ while there were 15 kHz spacing between them. The BS was equipped with large number of antennas. Moreover, in each transmitter, the interleaved bits are modulated using 16 Quadrature-Amplitude-Modulation (16-QAM). Transmit power of both data and pilot subcarriers are the same. Besides, the total available power is divided among transmitting antennas. Channels are distributed according to Extended Pedestrian A (EPA) available in LTE-A standard channel model [11]. Moreover, the channel taps are scaled to support $N_{\mathcal{A}}N_{\mathcal{B}}$ total channel power. In the detection, after channel estimation, the zero-forcing equalizer omits the impact of the channel. Furthermore, the simulation results are averaged on 2000 independent runs.

At first, the proposed channel estimation is compared with the other well-known DCS-based channel estimation methods called ABSP from [4]. The results are shown in Fig. 1. This simulation is carried on utilizing $N_p = 32$ pilot subcarriers for each of the transmitting antenna. As represented in Fig. 1, the proposed StFBP approach outperforms the other method. From NMSE perspective, the proposed approach is superior to ABSP method almost 7 dB. Moreover, increasing SNR more than 30 dB is seemingly useless for ABSP. The main advantage of the proposed StFBP approach to the well-known ABSP is the fixing process which is absent in ABSP. As a consequence, previous steps selected bad atoms are omitted, and the NMSE is getting lower while in ABSP the previous steps selected bad atoms are not omitted and remain in the last estimated channel impulse response. Hence, adjusting a reasonable threshold for selecting atoms in each step is essential in ABSP while in the proposed StFBP this is not as important as ABSP since in fixing process we can omit the previous steps chosen bad atoms.

In Fig. 2, the comparison of different number of antennas is represented. As obvious, NMSE is only 3 dB effected by increasing the number of antennas. This demonstrate the proposed method application in large number of antennas in BS. Moreover, the proposed channel estimation is superior approximately 5 dB in lower SNR and 15 dB in higher SNRs than ABSP method for all the number of antennas.

V. Conclusion

In this letter channel estimation of the mMIMO-OFDM system using DCS approach is considered. To utilize the channel block sparsity together with the MIMO channel large scale behavior, we have proposed to study StFBP algorithm. The proposed estimation approach which is called StFBP utilizes the channel block sparsity and adds multiple atoms in each forward steps while omits badly selected atoms. Skipping bad atoms makes it possible improve the precision of the channel estimation which is entirely appreciable in comparison with ABSP. In order to increase the speed of convergence, we have proposed to select multiple atoms in each forward step.

References

[1] 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.

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

[3] D. Fan, F. Gao, G. Wang, Z. Zhong, and A. Nallanathan, "Angle domain signal processing aided channel estimation for indoor 60GHz TDD/FDD massive MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 1948–1961, Sep. 2017.

[4] Y. Nan, L. Zhang, and X. Sun, "Efficient downlink channel estimation scheme based on block-structured compressive sensing for TDD massive MU-MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 4, no. 4, pp. 345–348, Aug. 2015.

[5] J. W. Choi, B. Shim, and S.-H. Chang, "Downlink pilot reduction for massive MIMO systems via compressed sensing," *IEEE Commun. Lett.*, vol. 19, no. 11, pp. 1889–1892, Nov. 2015.

[6] M. Masood, L. H. Afify, and T. Y. Al-Naffouri, "Efficient coordinated recovery of sparse channels in massive MIMO," *IEEE Trans. Signal Process.*, vol. 63, no. 1, pp. 104–118, Jan. 2015.

[7] D. Lee, "MIMO OFDM channel estimation via block stagewise orthogonal matching pursuit," *IEEE Commun. Lett.*, vol. 20, no. 10, pp. 2115–2118, Oct. 2016.

[8] A. Akbarpour-Kasgari and M. Ardebilipour, "Probability-based pilot allocation for MIMO relay distributed compressed sensing based channel estimation," *EURASIP J. Adv. Signal Process.*, vol. 2018, no. 1, pp. 1–18, Mar. 2018.

[9] T. Zhang, "Adaptive forward-backward greedy algorithm for learning sparse representations," *IEEE Trans. Inf. Theory*, vol. 57, no. 7, pp. 4689–4708, Jul. 2011.

[10] A. Jameson, "Solution of the equation AX + XB = C by inversion of an $M \star M$ or $N \star N$ matrix," *SIAM J. Appl. Math.*, vol. 16, no. 5, pp. 1020–1023, Sep. 1968.

[11] *User Equipment Radio Transmission and Reception (Rel. 12), V12.9.0*, 3GPP Standard TS 36.101, Oct. 2015.