Compression Based LMMSE Channel Estimation with Adaptive Sparsity for Massive MIMO in 5G Systems

Li-Jun Ge, Yue Zhang, *Senior Member, IEEE*, Gao-Jie Chen, *Senior Member, IEEE*, Jun Tong, *Member, IEEE*

Abstract—**Massive multi-input multi-output (MIMO) has been regarded as one of the key technologies for fifth generation (5G) mobile communication systems, as it can significantly enhance the system capacity with high spectrum and energy efficiency. For massive MIMO systems, accurate channel estimation is a challenging problem, especially when the number of parameters to be estimated is large and the number of pilots is limited. In this paper, a compression based linear minimum mean square error (CLMMSE) channel estimation algorithm is proposed for massive MIMO in 5G systems. Compared with the traditional linear minimum mean square error (LMMSE) algorithm, the proposed approach calculates the channel autocorrelation matrix by investigating the channel prior information based on compressive sensing (CS) theory, utilizing the block sparsity of massive MIMO channels, to reduce the complexity for obtaining autocorrelation matrix. Then it substitutes matrix inverse operation by singular value decomposition (SVD) to further reduce the computational complexity. In addition, a block sparsity adaptive matching pursuit (BSAMP) method is also proposed to adaptively estimate the block sparsity of the channel in the first step of the proposed CLMMSE algorithm, which can make it more efficient. The sparsity-adaptive processing is achieved by setting a threshold and finding the position of the maximum backward difference, then using the regularized method to solve channel estimation as a convex optimization problem. Analyses and simulations indicate that with slight performance degradation, the proposed algorithm reduces the computational complexity significantly compared with the traditional LMMSE algorithm. And compared with pure CS methods, CLMMSE has an obviously better performance, which is benefit to solve the pilot pollution problem of massive MIMO in 5G systems. Furthermore, the BSAMP based CLMMSE algorithm has better performance and lower time complexity than the algorithm based on other CS methods, which further improves the system performance.**

Index Terms—**Massive MIMO, 5G, compressive sensing, channel estimation, CLMMSE, SVD, sparsity adaptive**

L. -J. Ge, and J. Tong are with the School of Electronics and Information Engineering, Tianjin Polytechnic University, Tianjin Key Laboratory of Optoelectronic Detection Technology and System, Tianjin, China. Email: {gelj_001@hotmail.com, eejtong@gmail.com}

Y. Zhang, and G. -J. Chen are with the Department of Engineering, University of Leicester, Leicester, LE1 7RH, United Kingdom. E-mail: {yue.zhang@leicester.ac.uk, Gaojie.chen@leicester.ac.uk}

Y. Zhang is the corresponding author.

I. INTRODUCTION

In the area of mobile communications, new technologies being able to increase the capacity and spectrum efficiency are forever needed to satisfy the increasing data rate demand from the users. The existing 3G CDMA (Code Division Multiple Access) and 4G LTE (Long Term Evolution) systems make the limited wireless bandwidth even more urgent. Multi-input multi-output (MIMO) is a key technology to increase the capacity and system reliability by exploring the spatial domain without any additional wireless bandwidth. In 4G communications, the standard named LTE-A (Long Term Evolution-Advanced) first adopts MIMO technology to boost the communication capacity [1]. Although MIMO can help increase the capacity over SISO with 4G systems, it cannot meet the expectation for the future mobile communication systems.

Massive MIMO can make a huge increase on system capacity with a much larger number of antennas configured at the base station (BS). Thus, it is regarded as one of the key technologies for future fifth generation (5G) wireless communication systems for its high spectrum and energy efficiency [2-3]. By utilizing spatial multiplexing, a massive MIMO base station can serve multiple users simultaneously through multi-user beamforming [4-6]. Recently, some proposals using massive MIMO technology for 5G standard have been presented in 3GPP release 15 for one of the key enabling technologies for 5G systems [7].

Baseband signal processing such as channel estimation, channel equalization, channel encoding, etc, are the important parts of wireless communication systems to resist fading channel [8-11], where channel estimation is especially critical for massive MIMO systems. In massive MIMO systems, accurate and efficient channel estimation is a challenging problem and an open research issue, because the number of channel parameters to be estimated is very large as the antennas increase, while the number of pilots adopted by channel estimation is limited to make sure a high spectrum efficiency.

Linear minimum mean square error (LMMSE), known as a classic algorithm, has been widely used for channel estimation of traditional wireless communication systems, such as LTE which adopts orthogonal frequency division

multiplexing (OFDM) for baseband modulation, due to its good anti-noise performance. It may also be suitable for the LTE-A system when the scale of antennas is not large. However, it is a challenge for LMMSE to be used in 5G massive MIMO systems that the channel autocorrelation matrix is difficult to get and the computational complexity is pretty high brought by the massive antennas [12]. Therefore, how to keep the optimum balance between the performance and complexity of channel estimation is essential to be studied in 5G systems.

In order to optimize the performance of the LMMSE algorithm, [13] utilizes the reference signals, received by the downlink of MIMO systems, to obtain second-order statistical properties such as channel correlation matrix, and then estimated the channel parameters. Reference [14] proposes a channel estimation method for MIMO-OFDM systems based on weighted iteration. By performing weight iteration on the LMMSE algorithm, higher estimation accuracy is achieved. In [15], the complexity of the LMMSE algorithm for MIMO-OFDM systems is optimized under the assumption of channel uncorrelation, and the difficulty of obtaining the channel autocorrelation matrix was avoided. Also, a discrete Fourier transformation (DFT) based LMMSE algorithm is proposed in [16], which uses DFT transform property to estimate the channel correlation matrix, reducing the amount of parameter calculation and achieving higher channel estimation performance with lower computational complexity.

Though numerous efforts have been done to improve the performance of traditional channel estimation algorithms, all of them are restricted to the Nyquist-Shannon theorem, which may lead to extremely high pilot overheads as well as high computational complexity. Compressive sensing (CS) theory is proposed by Donoho et al. in [17] for sparse signal processing. In recent years, it has been widely used in sparse channel estimation for massive MIMO systems [18-22] and [20] has proven that the massive MIMO channel shows joint sparse characteristics. In this paper, a compression based LMMSE (CLMMSE) channel estimation algorithm is proposed to reduce the complexity of LMMSE algorithm and enhance the performance of pure CS based estimation for massive MIMO systems. It uses the channel prior information estimated by CS theory to gain channel autocorrelation matrix, this operation using optimal rank reduction compared with traditional channel autocorrelation matrix calculation. Meanwhile, it adopts singular value decomposition (SVD) to substitute channel matrix inversion to further reduce the computational complexity of the proposed algorithm.

In the proposed CLMMSE algorithm, if the channel sparsity can be estimated accurately, the dimension of the autocorrelation matrix can be reduced greatly, which will further reduce the computational complexity. Thus, in the process of CS based estimation in the first step of the proposed CLMMSE algorithm, sparsity adaptive determination methods are also discussed in this paper. The

existing common CS channel estimators, such as those based on orthogonal matching pursuit (OMP) [23], subspace pursuit (SP) [24], need the knowledge of the channel sparsity level, which limits their applications in practice. The sparsity adaptive matching pursuit (SAMP) [25] algorithm can recover the channel information without knowledge of the channel sparsity level, but the iteration step size of this algorithm is fixed at the initial stage. If the step size is much smaller than the signal sparsity, a large number of iterations will be required. It has been demonstrated in [20] that the sub-channels between different transmitting and receiving antenna pairs have the same sparsity support set in massive MIMO systems, and an adaptive structured subspace pursuit (ASSP) algorithm exploiting such joint sparsity has been proposed in [21]. However, fixed step size cannot estimate the real sparsity precisely because this method can only set the estimated sparsity to an integer multiple of step size.

In this paper, a block sparsity adaptive matching pursuit (BSAMP) algorithm is further proposed, exploiting the joint sparse characteristics of massive MIMO channels, to make the proposed CLMMSE algorithm more efficient. When the channel sparsity level is unknown, the proposed algorithm first selects the atoms by setting a proper threshold and finding the position of maximum backward difference position. Then the regularized method is used to improve the accuracy of atoms selection. The proposed BSAMP algorithm reduces the iteration times greatly by exploiting the joint sparsity of sub-channels. The sparsity adaptive processing is more flexible, for it does not rely on fixed step size, which reduces the computational complexity effectively.

A. Contributions of this paper

Based on all the descriptions above, we summarize and list all the contributions of this paper here:

First, by utilizing the sparsity of massive MIMO channels, a compression based LMMSE channel estimation algorithm named CLMMSE is proposed, which gains the channel autocorrelation matrix by CS estimated channel prior information, solving the problem of obtaining the autocorrelation matrix, reducing the complexity of traditional LMMSE based channel estimation, improving the spectrum utilization compared with pure CS estimation for 5G wireless communication systems.

Second, SVD operation is adopted in the proposed CLMMSE algorithm to substitute channel matrix inversion, which further reduces the computational complexity efficiently.

Third, a block sparsity adaptive matching pursuit algorithm named BSAMP is proposed by exploiting the joint sparse characteristics of massive MIMO channels to make the proposed CLMMSE algorithm more efficient.
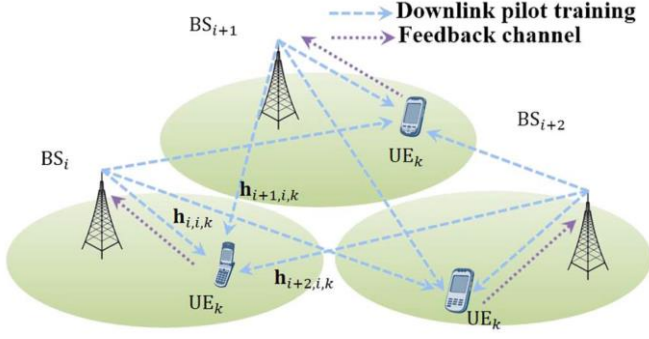
Fig. 1. Typical FDD 5G communication scenario

B. Organization of this paper

The reminder of this paper is organized as follows: Section II discusses the massive MIMO system model in typical 5G communication scenarios. Section III gives the detail description of the proposed CLMMSE algorithm and sparsity adaptive determination principle. Section IV gives the complexity analysis of the algorithms. In section V, a 5G massive MIMO system based on 3GPP release 15 evolved from LTE-A is established by MATLAB, and simulation results and discussions are presented to validate the proposed solutions. Section VI gives the conclusion of this paper.

II. SYSTEM MODEL

A typical frequency division duplexing (FDD) 5G communication scenario, based on 3GPP release 15 evolved from LTE-A, is shown in Fig. 1. The BS is equipped with M transmitting antennas, and OFDM modulation is adopted for downlink transmission. Suppose the total number of OFDM sub-carriers is N, and P sub-carriers ($P<N$) are selected as pilot sub-carriers. The length of the channel is L. Define the pilots of the ith transmission antenna as $\mathbf{p}(i)$, $i=1, 2, …, M$, and $\mathbf{p}(i) \cap \mathbf{p}(j) = \varnothing$, if $i \neq j$. Suppose $\mathbf{y}$ is the received signal after transmission, and $\mathbf{y}^{(i)}$ is the abbreviation of $\mathbf{y}(\mathbf{p}(i))$ corresponding to each antenna. The fundamental channel estimation model in the downlink can be formulated as

$$\mathbf{y}^{(i)} = \mathbf{P}^{(i)}\mathbf{F}^{(i)}\mathbf{h}^{(i)} + \mathbf{n}^{(i)} \tag{1}$$

where $\mathbf{P}^{(i)}=\text{diag}\{\mathbf{p}(i)\}$ is the diagonal matrix consisted of the selected pilots, $\mathbf{n}^{(i)}$ is the additive white Gaussian noise (AWGN), $\mathbf{n}^{(i)} \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I}_P)$, $\mathbf{I}_P$ is a $P \times P$ identity matrix. $\mathbf{F}^{(i)}$ is a $P \times L$ sub-matrix indexed by $\mathbf{p}(i)$ in row and [1, 2, …, L] in column from a standard $N \times N$ DFT matrix, where L is the length of the channel. $\mathbf{h}^{(i)}=[h^{(i)}(1), h^{(i)}(2), …, h^{(i)}(L)]^T$ denotes the channel impulsive response (CIR) of the ith transmitting antenna. Define $\mathbf{A}^{(i)}=\mathbf{P}^{(i)}\mathbf{F}^{(i)}$, then (1) can be written as

$$\mathbf{y}^{(i)} = \mathbf{A}^{(i)}\mathbf{h}^{(i)} + \mathbf{n}^{(i)} \tag{2}$$

Define $\text{supp}\{\mathbf{h}^{(i)}\}=\{l: |h^{(i)}(l)|>g_{th}, 1 \leq l \leq L\}$ represents the support index of the ith sub-channel and g_{th} is the noise threshold. It has been proven in [20] that sub-channels of different transmitting and receiving antenna pairs have approximately the same channel delay. Because the scale of the compact antenna array at the BS is relatively small compared with the long signal transmission distance in typical massive MIMO geometry, the channels associated with different transmitting and receiving antenna pairs share the common scatters. In other words, sub-channels between different transmitting and receiving antenna pairs have a common support set. It can be expressed as

$$\text{supp}\{\mathbf{h}^{(i)}\}=\text{supp}\{\mathbf{h}^{(j)}\} \tag{3}$$

when $i \neq j$, which shows the joint sparse characteristics in massive MIMO systems.

In order to exploit the joint sparsity of MIMO channels in all the transmitting antennas, we further define $\mathbf{h}=[\mathbf{h}_1^T, \mathbf{h}_2^T, …, \mathbf{h}_L^T]^T$, where $\mathbf{h}_j=[h^{(1)}(j), h^{(2)}(j), …, h^{(M)}(j)]^T$, $j=1, 2, …, L$ is the jth block of $\mathbf{h}$. Based on the fact that the position of non-zero entries of different antennas are same, we transform the received pilot signal as follows. Define $\mathbf{y}=[\mathbf{y}_1^T, \mathbf{y}_2^T, …, \mathbf{y}_P^T]^T$, where $\mathbf{y}_k=[y^{(1)}(k), y^{(2)}(k), …, y^{(M)}(k)]^T$, $k=1, 2, …, P$, and $\mathbf{n}=[\mathbf{n}_1^T, \mathbf{n}_2^T, …, \mathbf{n}_P^T]^T$, where $\mathbf{n}_k=[n^{(1)}(k), n^{(2)}(k), …, n^{(M)}(k)]^T$. Considering all the transmitting antennas, the received signal can be expressed as

$$\mathbf{y} = \mathbf{A}\mathbf{h} + \mathbf{n} \tag{4}$$

where $\mathbf{A}=[\mathbf{A}_1, \mathbf{A}_2, …, \mathbf{A}_L]$ is a $MP \times ML$ matrix, and M, P, L separately represent the number of transmitting antennas, pilot sub-carriers and channel length. $\mathbf{A}_j=[\mathbf{A}^{(1)}(j), \mathbf{A}^{(2)}(j), …, \mathbf{A}^{(M)}(j)]$ is the jth sub-matrix of $\mathbf{A}$ where $\mathbf{A}^{(i)}(j)$ is the jth column of $\mathbf{A}^{(i)}$. Formula (4) is the signal model of massive MIMO systems, where the CIR $\mathbf{h}$ is need to be estimated precisely by channel estimation algorithms.

III. COMPRESSION BASED LMMSE CHANNEL ESTIMATION WITH ADAPTIVE SPARSITY

A. CLMMSE Channel Estimation

In this section, we give the description of the proposed CLMMSE algorithm to solve the channel estimation problem of massive MIMO systems.

In the traditional LMMSE algorithm, the estimated result can be obtained as

$$\hat{\mathbf{H}}_{\text{LMMSE}} = \mathbf{R}_{\text{HH}}(\mathbf{R}_{\text{HH}} + \frac{\beta}{SNR}\mathbf{I}_N)^{-1}\hat{\mathbf{H}}_{\text{LS}} \tag{5}$$

where $\mathbf{R}_{\text{HH}}$ is the channel autocorrelation matrix. $\beta = \text{E}[|\mathbf{X}|^2]\text{E}[|1/\mathbf{X}|^2]$ is a constant determined by the modulation method, where $\mathbf{X}$ is the frequency domain data information and E[.] represents average energy calculation. $SNR = \text{E}[\mathbf{X}^2]/\sigma^2$ is the average signal-to-noise ratio (SNR), where σ^2 is the variance of the noise. $\mathbf{I}_N$ is a unit matrix. $\hat{\mathbf{H}}_{\text{LS}}$ is the channel frequency domain response estimated by

least square (LS) algorithm.

In (5), traditional LMMSE algorithm gets the channel autocorrelation matrix $\mathbf{R}_{HH}$ on the assumption that the power of the multi-path channel taps obeys negative exponential distribution, that the elements in $\mathbf{R}_{HH}$ can be calculated as

$$r_{m,n} = \frac{1-\exp[1-\tau_{\max}(j\frac{2\pi}{N}\tau_l(m-n)+\frac{1}{\tau_{rms}})]}{\tau_{rms}(1-\exp(\frac{\tau_{\max}}{\tau_{rms}}))[\frac{1}{\tau_{rms}+j\frac{2\pi}{N}\tau_l(m-n)}]} \quad (6)$$

where $\tau_{\max}$ is the largest delay of the channel, τ_l is the delay of the lth path and τ_{rms} is the average delay of the channel. However, in actual scenarios, prior information of the channel is hard to know, consequently the channel autocorrelation matrix is difficult to get. Although we can estimate the channel information, high computational complexity will be brought in massive MIMO systems for its large amount of antennas.

The problem stated above can also be solved in another way that the elements in $\mathbf{R}_{HH}$ can also be calculated by the power delay spectrum of the channel as

$$r_{m,n} = \sum_{l=0}^{L-1} \alpha_l^2 \exp(-j\frac{2\pi}{N}\tau_l(m-n)) \quad (7)$$

where α_l^2 is the power of the lth path tap and τ_l is the lth path delay. Based on formula (7), we obtain the prior information by CS based initial channel estimation in this paper, to improve the $\mathbf{R}_{HH}$ calculation efficiency, which is given in the next part.

In order to further reduce the complexity of the proposed algorithm, we also adopt SVD substitute matrix inversion and the channel autocorrelation matrix can be expressed as

$$\mathbf{R}_{HH} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H \quad (8)$$

where $\mathbf{U}$ is a unitary matrix, $\mathbf{\Lambda}$ is a diagonal matrix containing K singular values arranged in descending order from large to small that $\lambda(0) > \lambda(1)... > \lambda(K-1)$, and K is the channel sparsity level.

By substituting (8) into (5), and replacing $\hat{\mathbf{H}}_{LS}$ by CS based estimation result $\hat{\mathbf{H}}_{CS}$, the improved estimation formula can be expressed as

$$\hat{\mathbf{H}}_{CLMMSE} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H(\mathbf{U}\mathbf{\Lambda}\mathbf{U}^H + \sigma^2\mathbf{U}\mathbf{U}^H)^{-1}\hat{\mathbf{H}}_{CS} \quad (9)$$

Using unitary matrix property, the formula can be written as

$$\hat{\mathbf{H}}_{CLMMSE} = \mathbf{U}\text{diag}(\mathbf{\Delta},...,0)\mathbf{U}^H\hat{\mathbf{H}}_{CS} \quad (10)$$

where $\mathbf{\Delta} = \text{diag}(\frac{\lambda(0)}{\lambda(0)+\sigma^2}, \frac{\lambda(1)}{\lambda(1)+\sigma^2}, ..., \frac{\lambda(K-1)}{\lambda(K-1)+\sigma^2})$.

Considering all the transmitting antennas, the channel autocorrelation matrix of massive MIMO systems is expressed as

$$\mathbf{R}_{HH} = \begin{bmatrix} \mathbf{R}_{H_1H_1} & \mathbf{R}_{H_1H_2} & ... & \mathbf{R}_{H_1H_M} \\ \mathbf{R}_{H_2H_1} & \mathbf{R}_{H_2H_2} & ... & \mathbf{R}_{H_2H_M} \\ ... & ... & ... & ... \\ \mathbf{R}_{H_MH_1} & \mathbf{R}_{H_MH_2} & ... & \mathbf{R}_{H_MH_M} \end{bmatrix} \quad (11)$$

According to the channel autocorrelation property, highly correlated elements have a great impact on the performance of channel estimation, while the effect from lowly correlated elements is slight. Hence, $\mathbf{R}_{HH}$ can be approximately expressed in terms of

$$\mathbf{R}_{HH} = \text{diag}(\mathbf{R}_{H_1H_1}, \mathbf{R}_{H_2H_2}, ..., \mathbf{R}_{H_MH_M}) \quad (12)$$

Based on the former analysis, considering all the transmitting antennas, the ultimate channel vector is expressed as

$$\text{vec}(\hat{\mathbf{H}}_{CLMMSE}) = [\hat{\mathbf{H}}_{1,CLMMSE}^T, \hat{\mathbf{H}}_{2,CLMMSE}^T, ..., \hat{\mathbf{H}}_{M,CLMMSE}^T]^T \quad (13)$$

Formula (13) can further be transformed to

$$\text{vec}(\hat{\mathbf{H}}_{CLMMSE}) = \begin{bmatrix} \mathbf{U}_1\text{diag}(\mathbf{\Delta}_1,...,0)\mathbf{U}_1^H & & & \\ & \mathbf{U}_2\text{diag}(\mathbf{\Delta}_2,...,0)\mathbf{U}_2^H & & \\ & & ... & \\ & & & \mathbf{U}_M\text{diag}(\mathbf{\Delta}_M,...,0)\mathbf{U}_M^H \end{bmatrix}\hat{\mathbf{H}}_{CS} \quad (14)$$

where $\hat{\mathbf{H}}_{i,CLMMSE} = \mathbf{U}_i\text{diag}(\mathbf{\Delta}_i,...,0)\mathbf{U}_i^H\hat{\mathbf{H}}_{i,CS}$ is the estimated channel in frequency-domain form of the ith transmitting antenna. The detailed description of the proposed CLMMSE algorithm is shown in Table I.

TABLE I
THE PROPOSED CLMMSE ALGORITHM FOR CHANNEL ESTIMATION

Input: received pilot signal $\mathbf{y}$, observation matrix $\mathbf{A}$, number of antennas M.
Output: frequency domain channel estimation vector $\hat{\mathbf{H}}_{CLMMSE}$.
Step:
(1) Calculate the CIR vector $\mathbf{h}_i$ based on compressive sensing according to the signal model (4).
(2) Calculate the frequency domain channel autocorrelation matrix $\mathbf{R}_{H_iH_i}$ by using the obtained channel sub-path delay and tap information.
(3) Perform SVD operation on the matrix $\mathbf{R}_{H_iH_i}$ according to (10) to obtain the diagonal matrix $\mathbf{\Lambda}_i$, and then calculate $\mathbf{\Delta}_i$ by using the statistical characteristics of the noise.
(4) Perform Fourier transform on $\mathbf{h}_i$ to obtain the frequency domain channel information $\hat{\mathbf{H}}_{i,CS}$.
(5) Obtain the final channel estimation result $\hat{\mathbf{H}}_{CLMMSE}$ of the whole antennas according to (14).

B. CS Based Initial Channel Estimation

Since the massive MIMO channel shows sparse characteristics [20], compressive sensing algorithm is used in this paper to perform the initial channel estimation in the step (1) of the proposed CLMMSE algorithm.

In the basic compressive sensing principle, it recovers the channel $\mathbf{h}$ by calculating the correlation coefficient $\mathbf{u}$ between the measurement matrix $\mathbf{A}$ and the observation vector $\mathbf{y}$, which can be expressed as

$$\mathbf{u} = \left\{ u_j \middle| u_j = \left| \left\langle \mathbf{r}, \mathbf{A}_j \right\rangle \right| \right\} \tag{15}$$

where $\mathbf{r}$ is the residual of signal recovery, and j=1, 2, ..., N. When the atoms conform to the relevant condition, they will be selected into the support set $\mathbf{S}$. Then, the LS method is performed to recover the original signal and update the residual. The estimated signal can be expressed as

$$\hat{\mathbf{h}} = \arg \min \left\| \mathbf{y} - \mathbf{A}_S \mathbf{h} \right\|_2 \tag{16}$$

where $\left\| . \right\|_2$ denotes the l_2-norm, and $\mathbf{A}_S$ is derived from the observation matrix $\mathbf{A}$ by seeking the columns corresponding to $\mathbf{S}$. Then the new residual is expressed as

$$\mathbf{r} = \mathbf{y} - \mathbf{A}_S \hat{\mathbf{h}} \tag{17}$$

Repeat the above process of (15)-(17) until the residual meets the iteration quit condition.

Generally, to calculate the result of (16), greedy algorithms can be adopted to solve the following approximately equivalent l_0-norm minimization problem

$$\hat{\mathbf{h}} = \arg \min \left\| \mathbf{h} \right\|_0 \text{ subject to } \left\| \mathbf{y} - \mathbf{A}_S \mathbf{h} \right\|_2 \leq \varepsilon \tag{18}$$

where $\left\| . \right\|_0$ denotes the l_0-norm, i.e., the number of non-zero elements in the channel vector $\mathbf{h}$.

1) Sparsity Adaptive Determination: In this section, the proposed BSAMP algorithm is given to adaptively determine the sparsity of the CS algorithm, which can make the CLMMSE estimation more efficient.

In (18), reference [26] has proven that when

$$\left\| \mathbf{h} \right\|_0 < \frac{1}{2} \operatorname{spark}(\mathbf{A}) \tag{19}$$

is satisfied, unique recovery of $\mathbf{h}$ can be guaranteed, where spark($\mathbf{A}$) is the minimum number of columns with linear correlation of matrix $\mathbf{A}$. It can easily be seen that $2 \leq \operatorname{spark}(\mathbf{A}) \leq \operatorname{rank}(\mathbf{A}) + 1$. Since $\mathbf{A}$ is a $P \times L$ partial Fourier matrix and $P < L$, it can be found that $\left\| \mathbf{h} \right\|_0 < (P+1)/2$.

Since the CIR of the wireless channel shows sparse characteristics, the maximum number of the non-zero channel taps is small compared with the channel length L and (19) is satisfied. Denote K' as

$$K' = \begin{cases} \dfrac{P}{2}, & P \text{ is even} \\[2mm] \dfrac{P\text{-}1}{2}, & P \text{ is odd} \end{cases} \tag{20}$$

the L-K' elements of CIR with the smallest power can be considered as noise based on the analysis above.

Greedy algorithms select a fixed number of atoms during each iteration. As the number of antennas increases in massive MIMO systems, the number of iterations rises dramatically, and it incurs an intractably high computational complexity. Usually, the power of CIR is higher than that of AWGN. A clever thing to do to make a distinguish between the non-zero taps and the noise is to find a position with the fastest change of power. The selection of atoms can be more flexibility and efficiency by finding this position. Therefore, we calculate the backward difference between two adjacent elements to choose the appropriate entries. Those two elements with the maximum backward difference are considered as the boundary of non-zero taps and noise.

Much work has been done to show that convex optimization methods can be used in signal recovery to further improve the performance [23]. If for the k-sparse vector $\mathbf{h}$ there is

$$(1\text{-}\delta) \left\| \mathbf{h} \right\|_2 \leq \left\| \mathbf{A}\mathbf{h} \right\|_2 \leq (1+\delta) \left\| \mathbf{h} \right\|_2 \tag{21}$$

where $\delta \in (0,1)$, it is shown that the measurement matrix $\mathbf{A}$ satisfies the restricted isometry property (RIP). [27] shows that when matrix $\mathbf{A}$ satisfies the RIP and $\delta < \sqrt{2} - 1$ at the same time, the convex program can exactly recover the k-sparse signal. Since the partial Fourier matrix satisfies the RIP with the parameter $\delta < 0.5$ [28], the sparse recovery problem is equivalent to the convex program, expressed as

$$\hat{\mathbf{h}} = \arg \min \left\| \mathbf{h} \right\|_1 \text{ subject to } \left\| \mathbf{y} \text{ - } \mathbf{A}\mathbf{h} \right\|_2 \leq \varepsilon \tag{22}$$

Greedy methods are usually fast and easy to implement. Numerical evidence [29] suggests that the convex program should be more stable than the greedy algorithms. In order to guarantee the accuracy of the selection of the atoms, we further adopt the regularized method proposed in [30] to enhance the performance of the proposed algorithm. [30] has proved that the bound for the stability of the regularized method has the same form as that of the convex problem. The regularized method succeeds with absolutely no prior knowledge about the error noise, which magnitude can be arbitrary [30], and the method identifies at least one atom of the support set during each iteration.

2) BSAMP Based Initial Channel Estimation: In this subsection, the proposed BSAMP algorithm is applied to massive MIMO systems to do sparsity adaptive initial channel estimation.

Eq. (4) can be solved by CS exploiting the joint sparse characteristics of $\mathbf{h}$. Let both sides of (4) multiply $\mathbf{A}^H$, where $\mathbf{A}^H$ is the conjugate transpose of matrix $\mathbf{A}$, then

$$\begin{aligned} \mathbf{A}^H \mathbf{y} &= \mathbf{A}^H (\mathbf{A}\mathbf{h} + \mathbf{n}) \\ &= (\mathbf{I} + \mathbf{A}^H \mathbf{A} - \mathbf{I})\mathbf{h} + \mathbf{A}^H \mathbf{n} \\ &= \mathbf{h} + (\mathbf{A}^H \mathbf{A} - \mathbf{I})\mathbf{h} + \mathbf{A}^H \mathbf{n} \end{aligned} \tag{23}$$

where $\mathbf{I}$ is a $ML \times ML$ unit matrix. Since the observation matrix $\mathbf{A}$ is not strictly orthogonal generally, $\mathbf{A}^H \mathbf{A} - \mathbf{I}$ is a non-zero matrix with small element values and $(\mathbf{A}^H \mathbf{A} - \mathbf{I})\mathbf{h}$ represents the energy dispersion caused by the

non-orthogonality. Let $\mathbf{n}' = (\mathbf{A}^H\mathbf{A} - \mathbf{I})\mathbf{h} + \mathbf{A}^H\mathbf{n}$, then formula (23) can be rewritten as

$$\mathbf{A}^H\mathbf{y} = \mathbf{h} + \mathbf{n}' \qquad (24)$$

Then, calculate $\mathbf{R} = |\mathbf{A}^H\mathbf{r}|$, where $\mathbf{r}$ is the iteration residual indicating the residual energy of CIR after each iteration, and |.| means taking element-wise absolute values. Let $\mathbf{T}$ be a $L \times 1$ vector with its ith entry computed as

$$T(i) = \sum_{(j-1)\times M+1}^{j\times M} |R(i)|^2, \quad i = 1, 2, ..., ML, \qquad (25)$$
$$j = 1, 2, ..., L$$

where $R(i)$ is the ith element of $\mathbf{R}$. Then sort $\mathbf{T}$ in descending order to obtain $\mathbf{T}_s$, and denote the corresponding permutation vector by $\mathbf{S}_1$.

Since a recovery algorithm should be stable, it should be able to approximately recover the original signal from these perturbed measurements. From the analysis above, set $f = \mathrm{E}\{[T_s(i)]^2, i = K'+1, K'+2, ..., L\}$ as a threshold, then the elements in $\mathbf{T}_s$, whose power is not less than f, are selected into the support set.

At each iteration of the proposed BSAMP algorithm, we calculate the maximum backward difference of the elements in $\mathbf{T}_s$ and denote its position as t. Then the non-zero taps can be selected by t, and then regularize this set to improve its accuracy [30]. In this case, the power of the selected atoms is much larger than that of the unselected ones. The detailed description of the proposed BSAMP algorithm is shown in Table II.

TABLE II
THE PROPOSED BSAMP ALGORITHM FOR INITIAL CHANNEL ESTIMATION

Input: received pilot signal $\mathbf{y}$, observation matrix $\mathbf{A}$, number of antennas M.
Output: channel estimation vector $\hat{\mathbf{h}}$.
Initialization: block support set index $\mathbf{S}_1 = \varnothing$, support set index $\mathbf{S}_2 = \varnothing$, $\hat{\mathbf{h}} = \mathbf{0}$, threshold $f = \mathrm{E}\{[T_s(i)]^2, i = K'+1, K'+2, ..., L\}$, residual $\mathbf{r} = \mathbf{y}$.
Iteration:
(1) Calculate vector $\mathbf{T}$ according to (25) and sort it in descending order to obtain vector $\mathbf{T}_s$ and $\mathbf{S}_1$.
(2) Select the elements in $\mathbf{T}_s$ whose energy is larger than the threshold f, and set the number of elements be m. If $m = 0$, exit; otherwise, go to step (3).
(3) Calculate the maximum backward difference between adjacent elements in $\mathbf{T}_s(1:m+1)$, and denote its position by t.
(4) Do regularization to $\mathbf{T}_s(1:t)$. Define $\mathbf{e} = \mathbf{T}_s(1:t)$ and $\mathbf{J} = \mathbf{S}_1(1:t)$. Sort $\mathbf{e}$ into some groups according to $|e(i)| \le 2|e(j)|$, $i, j \in \mathbf{J}$, and choose the group with the maximal energy. Define $\mathbf{V}$ as the index vector of the selected atoms. Then $\mathbf{S}_2 = \mathbf{S}_2 \cup [(\mathbf{V}(k)-1)M+1 : \mathbf{V}(k)M]$, $k = 1, 2, ..., U$, where U is the length of the vector $\mathbf{V}$.
(5) Seek the columns in the observation matrix $\mathbf{A}$ corresponding to $\mathbf{S}_2$, and denote them as $\mathbf{A}_{S_2}$.
(6) Calculate the channel estimation vector by the LS algorithm, where $\hat{\mathbf{h}} = (\mathbf{A}_{S_2}^H \mathbf{A}_{S_2})^{-1} \mathbf{A}_{S_2}^H \mathbf{y}$.
(7) Update the residual by $\mathbf{r} = \mathbf{y} - \mathbf{A}_{S_2}\hat{\mathbf{h}}$, set $\mathbf{S}_1 = \varnothing$, $\mathbf{V} = \varnothing$ and return to step (1).
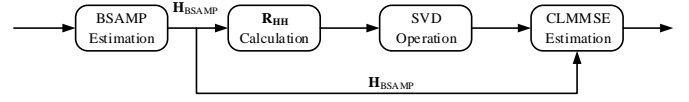

Fig. 2. Diagram of proposed BSAMP based CLMMSE algorithm

IV. ALGORITHM DIAGRAM AND COMPLEXITY ANALYSIS

A. BSAMP based CLMMSE Channel Estimation

In this section, we give the intact diagram of the proposed BSAMP based CLMMSE algorithm stated above, as show in Fig. 2, and compare its complexity with the traditional LMMSE algorithm and the OMP (typical CS algorithm with low computational complexity) based CLMMSE algorithm, to indicate the implementation difficulty, in terms of times of multiplication and addition operation.

Based on the assumption that the channel length is L and the channel sparsity level is K ($K \ll L$), the complexity of the traditional LMMSE algorithm is $O(2L^3)$, which mainly comes from the calculation of the channel autocorrelation matrix and the matrix inversion operation, and $O(.)$ represents the main complexity of the algorithm. Specifically, the number of multiplication in the LMMSE algorithm can be calculated as $2L^3 + 4L^2 + PL$. Since the CLMMSE algorithm uses the priori channel information provided by CS theory, when the CIR is initially estimated by OMP algorithm with $P/2$ iterations (a fixed experience value), the computational complexity of OMP based CLMMSE is $O(P^2L/2)$, and the specific number of multiplication is $2K^3 + 4K^2 + PK + PL(P+1)/2$. When the CIR is initially recovered by using the proposed BSAMP algorithm with R iterations ($R \le K \le P/2$), the calculation complexity is reduced to $O(RPL)$. Specifically, the number of multiplication in the BSAMP based CLMMSE algorithm is $2K^3 + 4K^2 + PK + RL(P+1)$. The complexities of the proposed and traditional algorithms are listed in Table III.

TABLE III
COMPLEXITIES OF THE PROPOSED AND TRADITIONAL ALGORITHMS

Algorithms	Multiplications	Additions
LMMSE	$2L^3 + 4L^2 + PL$	$2L^3 + 4L^2 + (P+4)L$
OMP-CLMMSE	$2K^3 + 4K^2 + PK + PL(P+1)/2$	$2K^3 + 4K^2 + (P-2)K + P^2(L-1)/2$
BSAMP-CLMMSE	$2K^3 + 4K^2 + PK + RL(P+1)$	$2K^3 + 4K^2 + (P-2)K + RP(L-1)$

From above analysis, it can be seen that since $K \ll L$ and $R \le K \le P/2$, the complexity of BSAMP based CLMMSE is much lower than that of LMMSE, and also lower than that of OMP based CLMMSE. The complexity of the proposed channel estimation algorithm is based on the block channel sparsity K, the channel length L, the pilot number P and the

iteration time *R*. As the four parameters being larger, the complexity of the algorithm increases correspondingly.

B. BSAMP Based Initial Channel Estimation

In this section, we separately give the complexity analysis of the proposed BSAMP algorithm and other reconstruction algorithms of CS in terms of iteration times and atoms searching time.

OMP algorithm, the most representative CS algorithm, is simple to operate and has fast convergence speed. But the channel sparsity is needed as a priori information in order to recover the original signal accurately. During each iteration process, the atom with the biggest correlation coefficient is selected into the support set, and then LS algorithm is used to calculate the signal and update the residual. The channel state information is recovered after *K* iteration times.

The iteration process of SP algorithm, another representative CS reconstruction algorithm, is very similar to that of OMP algorithm. The difference is that the atoms with *K* biggest correlation coefficients will be selected into the support set during each iteration, which makes it more efficient than OMP.

Both OMP and SP need the channel sparsity level as a priori information to recover the original signal accurately. However, the channel sparsity level *K* is usually unknown in actual situations. Artificially setting a fixed and large experience value to *K*, will bring invalid iteration operation and a low computational efficiency.

To solve this problem, SAMP algorithm recovers the original signal by adjusting the iterative step adaptively, according to the change of the signal residuals during the iteration. Fig. 3 shows the conceptual diagram of the SAMP algorithm in the *k*th iteration. Here, r_k represents the residue, C_k and F_k represent the candidate set and the final support set of the estimated signal, respectively. It is quite similar to that of the SP algorithms except that the sizes of candidate set $|C_k|$ and final set $|F_k|$ are adaptive. The estimation process could be divided into multiple stages, each of which contains multiple iterations. $|F_k|$ is kept fixed for iterations in the same stage and increased by a step size $s \leq K$ between two consecutive stages. The number of atoms selected into the support set is equal to the iteration step size. At the same time, the algorithm can eliminate the atoms which are not conform to the iteration condition. If the residual calculated at one stage is bigger than that of the previous stage, the algorithm will enter the next stage and adjust the iteration step size accordingly. The size of the final support set matches the signal sparsity, and the signal recovery process is completed. The sparsity adaptive processing of ASSP, which utilizes the block sparse characteristics of massive MIMO channel, is like SAMP. However, these methods using fixed step size cannot estimate the real sparsity precisely, because they can only set the estimated sparsity to an integer multiple of the step size.
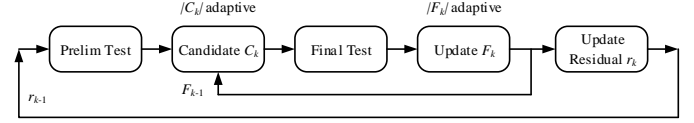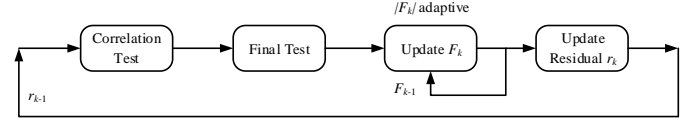

Fig. 3. Diagram of original SAMP algorithm

Fig. 4. Diagram of proposed BSAMP algorithm

The diagram of the proposed BSAMP algorithm in the *k*th iteration is shown is Fig. 4, which combines the speed and ease of implementation of the greedy algorithms with the strong guarantees of the convex program methods. It takes advantages of two major approaches to channel estimation: firstly, the atoms with the biggest power are picked by setting the proper threshold and finding the boundary position in $\mathbf{A}^H\mathbf{r}$; secondly, the regularized method based on convex program during each iteration ensures that the selected atoms are disjoint from those selected from the last one, which further make sure the support is selected correctly.

Besides, the proposed BSAMP algorithm reduces the iteration times greatly by exploiting the joint sparse characteristics of sub-channels in massive MIMO systems. The sparsity adaptive processing is more flexible and it does not rely on fixed step size. During each iteration, at least *M* atoms can be selected into the support set, which reduces the computational complexity effectively.

V. SIMULATION EXPERIMENT

In this section, the performance of the proposed channel estimation algorithm is evaluated via MATLAB simulation which offers a theoretical basis for hardware implementation. In this simulation, a massive MIMO-OFDM system with 64 transmitting antennas is constructed. The main system parameters are shown in Table IV, and they are forward compatible with the existing FDD-LTE standard defined by LTE-A and suitable for 3GPP release 15. Here, 64QAM modulation and low density parity check (LDPC) encoding are adopted. The number of total sub-carriers is 256, where 16 are the pilot sub-carriers. The sub-carrier frequency interval is 15 KHz. The length of cyclic prefix of transmitted OFDM symbols is 64. In addition, although channel models are recommended in the LTE standard, the proposed algorithm is more suitable and applicable to sparse massive MIMO channel recovery. Thus, according to [21], a channel model with length 60 is considered, which includes 4 to 8 valid paths dynamically changing as time. The positions of non-zero paths are uniformly distributed and the tap

coefficients follow Rayleigh distribution. To maximize the performance of channel estimation, pilot sub-carriers of the system are selected randomly for each antenna.

In the following part, the proposed channel estimation algorithm will be applied to the constructed massive MIMO system. By using Monte Carlo simulation, mean square error (MSE) curves and running times will be firstly given to show the estimation accuracy and efficiency separately. And then, constellations and bit error rate (BER) curves will also be given to indicate the performance of the system adopting the proposed algorithm. In order to make a contrast, other existing channel estimation algorithms are also simulated.
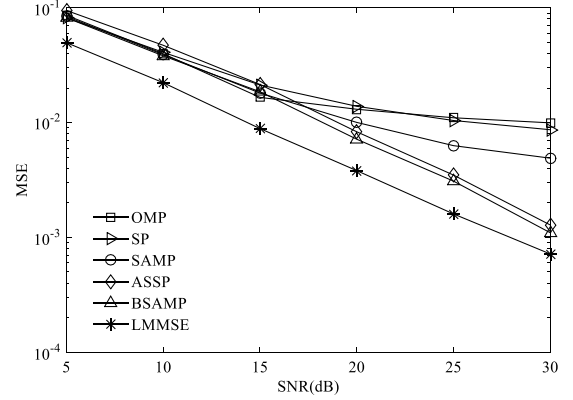

Fig. 5. MSE performance of different CS algorithms

Table V illustrates the average running time of different channel estimation algorithms with 1000 trails. We can see that although the OMP and SP algorithms cannot achieve sparsity adaptive processing, these two algorithms perform better in computational efficiency compared with the SAMP algorithm. The SAMP algorithm does not exploit the joint sparse characteristics of MIMO sub-channels and estimates the channel information individually by fixed step size, which causes high computational complexity due to the large scale of aggregate channel vector. By contrast, the ASSP algorithm and the proposed BSAMP algorithm utilize the joint sparsity of the sub-channels and can recover the information of multiple antennas at the same time during each iteration. Moreover, the sparsity adaptive processing of the BSAMP algorithm is more flexible that it does not rely on fixed step size which exhibits significant computational efficiency than other CS algorithms. From the table, it can be seen that the average running time of the BSAMP algorithm is 0.01284 s, which is only about 0.56% and 0.1% of that of OMP and ASSP separately.

TABLE IV
SYSTEM PARAMETERS

Parameters	Values
Transmitting antennas (M)	64
Total sub-carriers (N)	256
Pilot sub-carriers (P)	16
sub-carrier interval	15 KHz
Channel length (L)	60
Channel sparsity (K)	[4,8]
Digital modulation	64QAM
Channel encoding	LDPC

TABLE V
AVERAGE RUNNING TIME OF DIFFERENT ALGORITHMS

Channel Estimation Algorithms	Average Running Time (s)
OMP	2.2887
SP	1.1069
SAMP	93.393
ASSP	13.207
BSAMP	0.01284

A. Performance of BSAMP Based Initial Channel Estimation

Fig. 5 demonstrates the MSE performance of different CS based channel estimation algorithm versus the SNR through 1000 runs. The MSE of the traditional LMMSE algorithm is illustrated as a benchmark, which shows the optimal performance, for the exact locations of non-zero channel taps, the channel autocorrelation matrix and the noise variance are known as prior information. It can be observed that the algorithms exploiting joint sparse characteristics of sub-channels outperform the algorithms using individual sparse channel estimation. OMP and SP algorithms need to set the iteration times as half of the channel length when the channel sparsity is unknown, which reduces the precision of the algorithms greatly. SAMP algorithm does not need the priori knowledge of the channel sparsity level, but the fixed step size will degrade the estimation accuracy. By exploiting the joint characteristics of sub-channels, ASSP and BSAMP algorithms outperform others under the same SNR. BSAMP algorithm uses the regularized method to further enhance the channel estimation and eliminates the precision descending caused by the fixed step size, thus it has better performance than ASSP algorithm. The MSE of the proposed BSAMP algorithm is about 3×10^{-3} when the SNR is 25 dB.

B. Performance of BASMP Based CLMMSE Channel Estimation

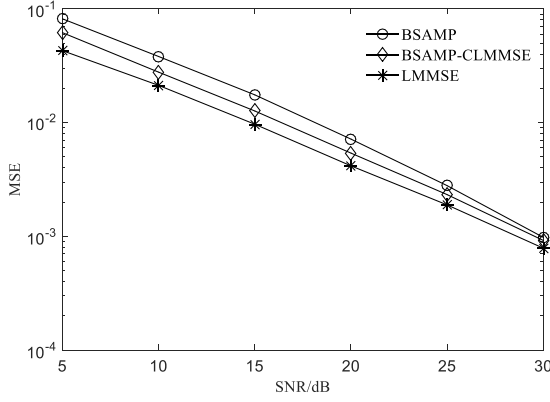Fig. 6 shows the MSE performance of the proposed BSAMP, BSAMP based CLMMSE and traditional LMMSE

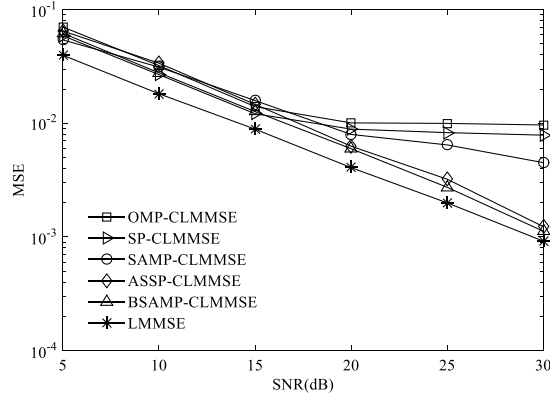Fig. 6. MSE performance of proposed CLMMSE algorithm

Fig. 7. MSE performance of different CS algorithms based CLMMSE

(a)   (b)

(c)

Fig. 8. Constellations of received data using different CS based CLMMSE. (a) SAMP-CLMMSE. (b) ASSP-CLMMSE. (c) BSAMP-CLMMSE.

versus the SNR. It can be seen that compared with the pure BSAMP, the BSAMP based CLMMSE algorithm has better performance, especially in lower SNR conditions. As the SNR increasing, the performance of BSAMP based CLMMSE gets close to that of pure BSAMP, and both of them are close to that of LMMSE. When the SNR is 25 dB, the MSE of BSAMP based CLMMSE gets 2×10^{-3}.

It investigates that the anti-noise performance of the proposed CLMMSE algorithm outperforms existing pure CS based channel estimation that is benefit for solving the pilot pollution problem. Although the performance of the proposed algorithm is slightly lower than that of the traditional LMMSE, it is in exchange for a significant increase in computational efficiency at an acceptable performance cost.

Fig. 7 shows the MSE performance of channel estimation algorithms versus the SNR, after applying different CS algorithms to CLMMSE. As can be seen, the performances of all the CLMMSE algorithms are lower than that of traditional LMMSE. However, the proposed BSAMP based CLMMSE has the most similar performance, which is obviously better than other ones, with that of LMMSE in all the CLMMSE algorithms.
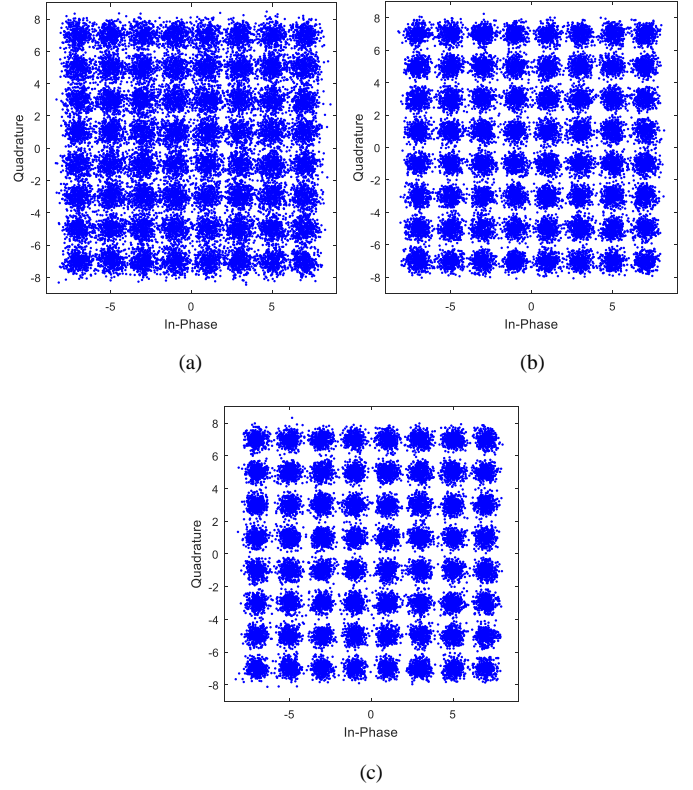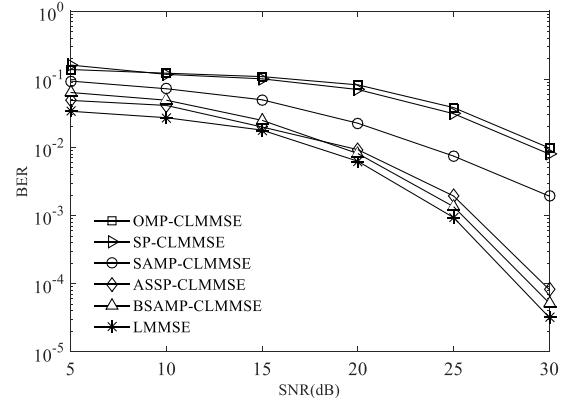

Fig. 9. BER performance of system using different CS based CLMMSE

C. Performance of CLMMSE Based 5G System

Fig. 8 (a) (b) and (c) separately shows the received signal constellation of the constructed system with 64QAM, by using CLMMSE channel estimation based on different sparsity adaptive CS algorithm (SAMP, ASSP, and BSAMP), when the SNR is 25 dB. It can be seen that the BSAMP based CLMMSE algorithm can recover the data with higher quality, compared with the ASSP based algorithm, while it is a little

serious for SAMP that the constellation obviously diffuses.

In Fig. 9, the BER performance of the system versus the SNR is given, using different channel estimation algorithms. As can be seen, the BERs of the system using different CS based CLMMSE are lower than that of the system adopting traditional LMMSE. However, the system utilizing the proposed BSAMP based CLMMSE has the most similar BER performance, which is obviously better than other ones, with that of the LMMSE based system. When the SNR is 25 dB, the BER of the system with the proposed algorithm is 1×10^{-3}, which guarantees the communication reliability of the system.

D. Application and Implementation Introduction

From the above simulation experiments, the proposed sparsity adaptive CLMMSE has great advantages in estimation accuracy and computational complexity, compared with other existing algorithms. By applying it to the existing FDD-LTE communication system, the pilot overhead can be reduced and the spectrum efficiency can be enhanced. For the further 5G systems with larger antenna scale, it can also be utilized as an economic estimator. In the FDD massive MIMO system shown in Fig. 1, the pilots are transmitted in the downlink, and the proposed channel estimation algorithm is performed at the UE side. To obtain a high spectrum efficiency, the design of pilot number should be taken into consideration, meanwhile the channel estimation performance should be continuously improved under limited pilot overhead.

On the other hand, to realistically implement the proposed channel estimation algorithm, the issue of hardware complexity should be carefully taken into account just as the analysis given in part IV, and the power consumption is normally proportional to the computational complexity. In order to show the implementation difficulty, based on the system parameters given in table IV, all the variances in table III are set to be specific values that $K=8$, $L=60$, $P=16$ and R is supposed to be the upper limit 8. Therefore, it can be calculated that the number of multiplications and additions of traditional LMMSE algorithm is 447360 and 447600 separately, while that of the proposed BSAMP based CLMMSE algorithm is 9568 and 8944 separately, which is only about 2.1% and 2.0% of that of LMMSE.

VI. CONCLUSION

In this paper, a compression based LMMSE algorithm named CLMMSE is proposed to recover channel state information, which is a challenging problem for massive MIMO in 5G systems. The channel is firstly estimated by applying CS theory and the channel autocorrelation matrix is calculated based on the initial estimation. SVD is also utilized to substitute the matrix inversion operation. Furthermore, in the process of CS based initial channel estimation, a sparsity

adaptive determination method named BSAMP is also proposed, which is achieved by setting a threshold and finding the position of the maximum backward difference. Theoretical and simulation analyses demonstrate that in the constructed 5G system, the proposed CLMMSE algorithm can keep the balance between performance and complexity. On condition of same pilots overhead, the proposed algorithm has an improved accuracy compared with existing pure CS based algorithms, which is benefit to improve the system spectrum utilization. When the SNR is 25 dB, the MSE of BSAMP based CLMMSE gets 2×10^{-3}. It is also investigated that the system utilizing the proposed BSAMP based CLMMSE channel estimation algorithm can achieve more excellent BER performance and much higher computational efficiency than those using other CS based CLMMSE. When the SNR is 25 dB, the BER of the system with the proposed algorithm can get 1×10^{-3}, which guarantees the communication reliability of the system. On the other hand, with a little loss of performance compared with the traditional LMMSE algorithm, the computational complexity is reduced from $O(2L^3)$ to $O(RPL)$, which is very important for massive MIMO systems to maintain an acceptable hardware cost.

In the future work, the proposed algorithm is planned to be implemented in an USRP based hardware experiment platform and the actual performance will be tested. Meanwhile, this algorithm will be optimized by furtherly exploring the sparsity of the massive MIMO channel and utilizing the correlation of the channel in time, frequency and spatial domain.
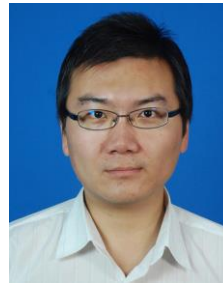
REFRENCES

[1] W. Guo, J. Fan, G. Y. Li, Q. Yin, X. Zhu and Y. Fu, "MIMO transmission with vertical sectorization for LTE-A downlink," *IEEE Wirel. Commun. Letters*, vol. 5, no. 4, pp. 372-375, Aug. 2016.

[2] T. L. Marzetta, "Non-cooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wirel. Commun.*, vol. 9, no. 11, pp. 3590-3600, Nov. 2010.

[3] H. Luo, Y. Zhang, L. Huang, J. Cosmas and A. Aggoun, "A closed-loop reciprocity calibration method for massive MIMO in terrestrial broadcasting systems," *IEEE Trans. on Broadcasting*, vol. 63, no. 1, pp. 11-19, March 2017.

[4] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Comm. Mag.*, vol. 52, no. 2, pp. 74-80, Feb. 2014.

[5] E. Larsson, O. Edfors, F. Tufvesson and T. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Comm. Mag.*, vol. 52, no. 2, pp. 186-195, Feb. 2014.

[6] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors and F. Tufvesson, "Scaling up MIMO: opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40-60, Jan. 2013.

[7] 3GPP, "3GPP release 15," *3GPP The Mobile Broadband Standard*. 2017. [Online]. Available: http://www.3gpp.org/release-15. [Accessed: Nov. 18, 2018]

[8] W. Li, Y. Zhang, L. Huang, C. Maple and J. Cosmas, "Implementation and co-simulation of hybrid pilot-aided channel estimation with decision feedback equalizer for OFDM Systems," *IEEE Trans. on Broadcasting*, vol. 58, no. 4, pp. 590-602, Dec. 2012.

[9] J. Zhong, G. Chen, J. Mao, S. Dang and P. Xiao, "Iterative Frequency Domain Equalization for MIMO-GFDM Systems," *IEEE Access*, vol. 6,

no. 4, pp. 19386-19395, April 2018.

[10] H. Luo, Y. Zhang, W. Li, L. Huang, J. Cosmas, D. Li, C. Maple and X. Zhang, "Low latency parallel turbo decoding implementation for future terrestrial broadcasting systems," *IEEE Trans. on Broadcasting*, vol. 64, no. 1, pp. 96-104, March 2018.

[11] W. Li, Y. Zhang, L. Huang, J. Cosmas, C. Maple and J. Xiong, "Self-IQ-demodulation based compensation scheme of frequency-dependent IQ Imbalance for wideband direct-conversion transmitters," *IEEE Trans. on Broadcasting*, vol. 61, no. 4, pp. 666-673, Sept. 2015.

[12] L. C. Fang and D. F. Huang, "Neumann series expansion based LMMSE channel estimation for OFDM systems," *IEEE Commun. Letters*, vol. 20, no. 4, pp. 748-751, Feb. 2016.

[13] C. Y. Hsieh, D. W. Lin, and C. Ma, "LMMSE-based channel estimation for LTE-advanced MIMO downlink employing UE-specific reference signals," in *VTC'15*, July 2015, pp. 1-5.

[14] M. Kashoob and Y. Zakharov, "Data-aided iterative reweighted LMMSE channel estimation for MIMO OFDM," in *ISSPIT'15*, Feb. 2015, pp. 663-667.

[15] A. Antonio, D. A. Emiliano and P. Silvano, "On the robustness of MIMO LMMSE channel estimation," *IEEE Trans. on Wirel. Commun.*, vol. 9, no. 11, pp. 3313-3319, Sept. 2010.

[16] D. Neumann, M. Joham, L. Weiland and W. Utschick, "Low-complexity computation of LMMSE channel estimates in Massive MIMO," in *WSA'15*, March 2015, pp. 1-6.

[17] D. L. Donoho, "Compressed sensing," *IEEE Trans. on Information Theory*, vol. 52, no. 4, pp. 1289-1306, April 2006.

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

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

[20] Z. Gao, L. L. Dai, Z. H. Lu, C. Yuen and Z. C. Wang, "Super-resolution sparse MIMO-OFDM channel estimation based on spatial and temporal correlations," *IEEE Commun. Letters*, vol. 18, no. 7, pp. 1266-1269, May 2014.

[21] Z. Gao, L. L. Dai, W. Dai, B. Shim and Z. C. Wang, "Structured compressive sensing-based spatio-temporal joint channel estimation for FDD Massive MIMO," *IEEE Trans. on Commun.*, vol. 64, no. 2, pp. 601-617, Dec. 2016.

[22] J. Shen, J. Zhang, K. Chen and K. B. Letaief, "High-dimensional CSI acquisition in massive MIMO: sparsity-inspired approaches," *IEEE Systems Journal*, vol. 11, no. 1, pp. 32-40, March 2017.

[23] J. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. on Information Theory*, vol. 53, no. 12, pp. 4655- 4666, Dec. 2007.

[24] J. J. Liu, C. Zhang and C. Y. Pan, "Priori-information hold subspace pursuit: a compressive sensing-based channel estimation for layer modulated TDS-OFDM," *IEEE Trans. on Broadcasting*, vol. 64, no. 1, pp. 119-127, June 2018.

[25] Y. Zhang, R. Venkatesan, O. A. Dobre and C. Li, "An adaptive matching pursuit algorithm for sparse channel estimation," in *WCNC'15*, March 2015, pp. 626-630.

[26] D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via l1 minimization," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 5, pp. 2197-2202, May 2003.

[27] E. Candes and T. Tao, "Decoding by linear programming," *IEEE Trans. Information Theory*, vol. 51, no. 12, pp. 4203-4215, Dec. 2005.

[28] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete Fourier information," *IEEE Trans. Information Theory*, vol. 52, no. 2, pp. 489-509, Feb. 2006.

[29] D. L. Donoho, M. Elad and V. N. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Trans. Information Theory*, vol. 52, no. 1, pp. 6-18, Jan. 2006.

[30] D. Needle and R. Vershynin, "Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp.

310-316, April 2010.

Li-Jun Ge received the B.E. degree in the major of Electronics Science and Technology in Nankai University, Tianjin, China, 2006. Subsequently, he did graduate study in the major of Communication and Information Systems in the same university. In 2011, he got the Ph.D. degree in the same major after five-year graduate study in Nankai University.
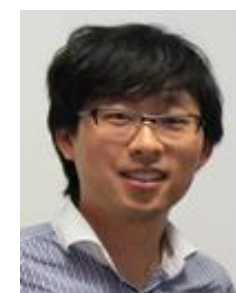
From 2008 to 2010, he was a Teaching Assistant in Nankai University teaching communication related experiment courses. In 2011, he joined the Department of Communication Engineering, School of Electronics and Information Engineering, Tianjin Polytechnic University, China, as a Lecture in Communication and Information Systems. From 2013, he has been an Associate Professor in the same academic field.

His research interests include OFDM and MIMO-OFDM wireless communication technologies, FPGA technologies and applications, and development of communication and information systems. During the past years, he was in charge of eight projects supported by the nation or the city or some companies and published more than forty academic papers.

Yue Zhang (M'06-SM'17) Associate Professor in Department of Engineering, University of Leicester, obtained his B.E and M.E degree in 2001 and 2004 respectively at Beijing University of Post and Telecommunications, China. In 2008 he received his PhD degree in Brunel University, UK, where he also worked as a Research Engineer for the EU FP6 project- PLUTO. From 2008, he was a Signal Processing Design Engineer in Microwave Measurement Division-Europe, Anritsu Corp. He was responsible for the RF/IF, digital and DSP design for the measurement instruments for various wireless and broadcasting systems. From 2010, he joined Department of Computer Science and Technology at the University of Bedfordshire, UK as Reader in Signal Processing. He also worked as a Royal Academy of Engineering, UK, Industrial Fellowship with Aeroflex Ltd. He currently leads EU Horizon 2020 5GPPP project IoRL as the Deputy Technique Manager. He is also one of the committee members of 5G PPP pre-standardization and 5G Architecture WG. His research interests are signal processing for 5G wireless and mobile systems, radio propagation model and multimedia and wireless networks. Dr. Zhang currently serves as an Associate Editor for IEEE Transactions on Broadcasting and IEEE Access.

Gaojie Chen (S'09-M'12-SM'18) received the B.Eng. and B.Ec. degrees in electrical information engineering and international economics and trade from Northwest University, China, in 2006, and the M.Sc. (Hons.) and Ph.D. degrees in electrical and electronic engineering from Loughborough University, Loughborough, U.K., in 2008 and 2012, respectively. From 2008 to 2009, he was a Software Engineering with DTmobile, Beijing, China, and from 2012 to 2013, he was a Research Associate with the School of Electronic, Electrical and Systems Engineering, Loughborough University. He was a Research Fellow with 5GIC, Faculty of Engineering and Physical Sciences, University of Surrey, U.K., from 2014 to 2015. Then he was a Research Associate with the Department of Engineering Science, University of Oxford, U.K., from 2015 to 2018. He is currently a Lecturer with the Department of Engineering, University of Leicester, U.K. He has served as an Editor for IET Electronics Letters (2018-present). His current research interests include information theory, wireless communications, cooperative communications, cognitive radio, secrecy communication, and random geometric networks.

Jun Tong received the B.E. and M.E. degrees from University of Electronic Science and Technology of China in 2001 and 2004, respectively, and got the Ph.D. degree from City University of Hong Kong in 2009. From 2009 to 2011, he did postdoctoral research in the Department of Electronics and Computers, University of Newcastle, United Kingdom. From 2011 to 2013, he was an Assistant Professor in the School of Electronic Information, University of Paderborn, Germany. At present, he is a Senior Lecturer with University of Wollongong, Australia, and a visiting lecturer at Tianjin Polytechnic University, China. His research interests include signal processing and communication systems.