Channel Estimation Based on Compressive Sensing for Multi-User Massive MIMO Systems

Qiang Chen1,*, Sunqing Su2 and Guowei Lei2

1School of information engineering, Jimei University, Xiamen, China
2School of science, Jimei University, Xiamen, China

*Corresponding author: qchen1@jmu.edu.cn

Abstract. Massive multiple-input multiple-output (MIMO) is becoming a key technology for future wireless communications. Channel feedback for massive MIMO is a challenging task due to the increased dimension of MIMO channel matrix. By exploiting the channel sparsity, channel estimation based on compressive sensing (CS) aims to reduce the feedback overhead in massive MIMO systems. In this paper, various CS algorithms for channel estimation in massive MIMO systems are summarized, and a novel CS algorithm, i.e. modified sparsity adaptive matching pursuit (MSAMP), is proposed hereupon. Moreover, various measurement matrices are introduced in the CS scheme. The performances of channel estimation and recovery are simulated and compared. It is inferred from simulation that, SAMP is very appropriate for reconstructing sparse channel information in massive MIMO system, especially the modified SAMP can give higher reconstruction, and some measurement matrix is suitable for certain optimal CS algorithms.

Keywords: massive multiple-input multiple-output, compressive sensing, measurement matrix, normalized mean square error.

1. Introduction

Massive multiple-input multiple-output (MIMO) is an emerging technology that provides unprecedented spectral efficiency by deploying a large amount of antennas at the base station (BS) [1]. Notwithstanding the number of BS and symbols of pilots are limited because of the mobility of UT and the length of coherent intervals, it can still be effective to increase number of antennas of BS under low signal interference noise ratio (SINR) and poor Channel State Information (CSI)[1]. One of foremost challenges which massive MIMO should face is to acquire CSI at BS. Taking into account the training overhead of collecting CSI at the transmitters [2], frequency-division duplexing (FDD) systems use downlink pilots for channel estimation and feed back the estimated CSI via uplink channels, while time-division duplexing (TDD) systems acquire uplink CSI via channel reciprocity [3].

To reduce the overhead, compressive sensing (CS) techniques have been introduced into CSI acquisition by exploiting the channel sparsity [4]. To date, many CS algorithms have been proposed for the reconstruction of CSI in massive MIMO system. The conventional algorithms include greedy methods, convex optimization, and so on. Matching pursuit (MP) is a type of fundamental greedy algorithms [5]. For each iteration, it chooses the column vector from the measurement matrix so as to
achieve the maximal inner product with the redundancy matrix. Each chosen vector is used as one element of the support set, and finally the set of column vectors is acquired to get the nearly sparse solution. Since the observation matrix is not square, the projection of signal on the observation will not be orthogonal. Further, orthogonal matching pursuit (OMP) is proposed to solve the problem [6]. To harness both spatial and temporal sparsity of the channel impulse response (CIR) for FDD massive MIMO channels, orthogonal matching pursuit (OMP) [7] and subspace pursuit (SP) [8] have been introduced into single-user MIMO channel estimation. For downlink massive multiuser (MU) MIMO system, an efficient estimation of CSI is proposed based on block subspace pursuit (SP) [9]. By exploiting the spatial and temporal correlations of large scale MIMO channels, the structured compressive sampling matching pursuit (CoSaMP) algorithm is proposed to recover multiple channel information [10]. To overcome the limit of a single column vector to the set of selected vectors at each stage, the stagewise OMP (StOMP) is employed in MIMO-OFDM channel estimation [11].

The above literatures mainly consider the base station with one user in MIMO system, or investigate just one of the CS algorithms in massive MIMO system. In our paper, we study the channel estimation based on CS for multi-user massive MIMO systems. Not only we conclude the application of various CS algorithms and propose a modified CS algorithm in channel estimation, but we introduce various measurement matrices and compare them. Finally, we draw comparisons among the CS algorithms and measurement matrices.

The rest of this paper is organized as follows. In section II, the multi-user massive MIMO system is introduced. The tranceiver model and massive MIMO channel are briefly described, including the measurement matrices. In section III, the CS optimal algorithms are summarized. In section IV, simulations and discussions are presented to verify the analysis. Section V concludes this paper.

To avoid confusion, the following notations are used throughout the manuscript. Superscripts $(\cdot)^*$, $(\cdot)^T$ and $(\cdot)^H$ denote matrix conjugate, matrix transpose and conjugate transpose. $|\cdot|$ denotes the absolute value, tr$(\cdot)$ denotes the trace of one matrix. $<\cdot, \cdot>$ denotes the inner product. In addition to these, $\|\cdot\|$ is the Euclidian vector norm, $\|\cdot\|_F$ denotes the Frobenius norm.

2. System model

The multi-user massive MIMO system is described in Fig. 1. The system is composed of one base station (BS) with $N$ antennas, and $L$ users each with $M$ antennas. In the channel estimation, a training sequence with the measured length of $T$ is transmitted by BS. It is assumed that the training sequence $X \in \mathbb{C}^{N \times T}$, channel matrix $H \in \mathbb{C}^{ML \times N}$, and the added white Gaussian noise (AWGN) $W \in \mathbb{C}^{ML \times T}$, then the received signals can be expressed by

$$Y =HX+W$$

In channel estimation, $Y$ denotes the measurement signal, $X$ denotes the measurement matrix, $H$ denotes the sparse channel matrix to be estimated. Various users have their respective channel matrices. Thereafter, the $ML$ antennas feed back their observations $\{Y_i\}_{i=1,2,\ldots,ML}$ to the BS side through error-free uplink channels.
Based on the representation of virtual angular domain [12], the channel matrix $H$ is rewritten as

$$H_i = A_R^H H_i^a A_R$$  (2)

where $A_R$ and $A_R^H$ are unitary matrices that represent the angular domain transformation at the user and BS side respectively, and $H_i^a \in \mathbb{C}^{ML \times N}$ is the angular domain channel matrix. In multi-user massive MIMO system, $H_i^a$ reflects sparse in general, as indicated by experimental study [13]. Hence, compressive sensing can be applied in the sparse measurement of massive MIMO channels.

Based on (2), the model in (1) can be transformed into a standard CS measurement model as follows

$$\hat{Y}_i = \hat{X} \hat{H}_i + \hat{W}_i$$  (3)

Where

$$\hat{Y}_i = \sqrt{\frac{N}{PT}} Y_i^H A_R$$  (4)

$$\hat{X} = \sqrt{\frac{N}{PT}} X_i^H A_R$$  (5)

$$\hat{H}_i = (H_i^a)^H$$  (6)

$$\hat{W}_i = \sqrt{\frac{N}{PT}} W_i^H A_R$$  (7)

It is noted that $\text{tr}(XX) = PT$, where $P$ is the transmit signal-to-noise ratio (SNR) per time slot at the BS. Therefore, the problem is equivalent to reconstructing the sparse matrix (6) with the received measurements (4) and the measurement matrix (5).

For different users, they share some common local scatterers at the BS [14]. Thus, their channel matrices $\{H_i^a\}$ have a partial common support. Common CS measurement matrices include sparse
random matrix, *Gaussian* random matrix, *Bernoulli* random matrix, *Toeplitz* matrix [15], partial *Fourier* matrix [16], partial Hadamard matrix [17], circulant matrix [18]. It is demonstrated from CS theory that, original signals with *N* dimension can be mapped into the signals with *M* dimension via the measurement matrix of size *M*×*N* (*M*<<*N) from sparse support design. When the measurement matrix satisfies the restricted isometry property (RIP), the original signals can be reconstructed from optimization algorithms [19]. For the measurement matrices, the correlations between the measured length *T* and the sparsity *K* are given in Table 1, where *N* denotes the length of measurement matrix, *c* is a constant value, *ε* is a constant value that lies in (0,1).

| matrices             | the correlation between *T* and *K*                                      |
|----------------------|------------------------------------------------------------------------|
| *Gaussian* random matrix | *T* ≥ *cK log (N / K)                                                  |
| *Bernoulli* random matrix  | *T* ≥ *cK log (N / K)                                                  |
| *Toeplitz* matrix      | *T* ≥ *cK log (N / ε)                                                  |
| partial *Fourier* matrix | *T* ≥ *cK(\log N)^6*                                                 |
| partial Hadamard matrix | *T* ≥ *cK^2 log (2N^2 / ε)                                          |
| circulant matrix       | *T* ≥ *cK^2 log (2N^2 / ε)                                          |

3. Compressive sensing algorithms

The reconstruction of measurement signal is the kernel step. Some robust and low complexity algorithms have been reported. In this section, we employ them in channel estimation for multi-user massive MIMO systems. To date, some main algorithms have been summarized as follow. Moreover, a modified algorithm is proposed to improve the performance.

1) Orthogonal matching pursuit (OMP) can guarantee the orthogonality of each element via recursion. In Table 2, *t* denotes the iterations, *K* denotes the total sparsity level. The estimated channel matrices \{\(H_i\)\} (i=1,2,…,ML)

| Procedure | Operations                                                                 |
|-----------|-----------------------------------------------------------------------------|
| **Input:** | The received measurements of all users \{\(Y_i\)\}, the pilot training matrix \(X\), the total sparsity level *K*. |
| **Output:** | The estimated channel matrices \{\(H_i\)\}={\(h_1\), …,\(h_N\)}. |
| 1         | Compute \{\(\hat{Y}_i\)\} and \(\hat{X}\) from \{\(Y_i\)\} and \(X\) as in (4a) and (4b). |
| 2         | Initialize, residual \(r_0=\hat{y}\), \(Λ_0=\emptyset\), \(ψ_0=\emptyset\), \(t=1\). |
| 3         | find index \(λ_t\), such that \(\hat{λ}_t = \arg \max_{j=1...N} \left| r_{t-1,λ} \right| \) |
| 4         | Let \(Λ_t=Λ_{t-1} \cup \{λ_t\}\), \(ψ_t=ψ_{t-1} \cup \{\hat{X}_{\hat{λ}_t}\}\). |
| 5         | Compute \(\hat{h}_t = \arg \min_h \|\hat{y} - ψ_t h\| = (ψ_t^T ψ_t)^{-1} ψ_t^T \hat{y}\) |
| 6         | Refresh \(r_t = \hat{y} - ψ_t \hat{h}_t = \hat{y} - ψ_t (ψ_t^T ψ_t)^{-1} ψ_t^T \hat{y}\) |
| 7         | \(t = t + 1\), if \(t ≤ K\), return to step 4, else continue. |
| 8         | Return \(H_i\). |
2) Regularized orthogonal matching pursuit (ROMP) adds the regularization to OMP. The purpose is to combine the speed of the greedy methods and ease of implementation with the convex programming methods [20]. The estimated channel matrices \( \{ H_i \} (i=1,2,\ldots,ML) \)

| Table 3. Algorithm of regularized orthogonal matching pursuit |
|-------------------------------------------------------------|
| **Procedure**                                                                 | **Operations**                                      |
| **Input:** The received measurements of all users \( \{ Y_i \} \), the pilot training matrix \( X \), the total sparsity level \( K \). |                                      |
| **Output:** The estimated channel matrices \( \{ H_i \} = \{ h_1, \ldots, h_N \} \). |                                      |
| 1 Compute \( \{ \hat{Y}_i \} \) and \( \hat{X} \) from \( \{ Y_i \} \) and \( X \) as in (4a) and (4b). |                                      |
| 2 Initialize, residual \( r_0 = \hat{y} \), \( \Lambda_0 = \emptyset \), \( \psi_0 = \emptyset \), \( t=1 \). |                                      |
| 3 Compute \( u^T X_t \hat{r} \), select the \( K \) biggest from \( u \), and construct the set \( J \) with the corresponding index \( j \). |                                      |
| 4 Let \( \Lambda_t = \Lambda_{t-1} \cup \{ J_0 \}, \psi_t = \psi_{t-1} \cup \{ \hat{X}_{j_0} \} \) (for all \( j \in J_0 \)). |                                      |
| 5 Compute \( \hat{h}_t = \arg \min_h \| \hat{y} - \psi_t h_t \| = (\psi_t^T \psi_t)^{-1} \psi_t^T \hat{y} \) |                                      |
| 6 Refresh \( r_t = \hat{y} - \psi_t \hat{h}_t = \hat{y} - \psi_t (\psi_t^T \psi_t)^{-1} \psi_t^T \hat{y} \) \( t=t+1 \), if \( t \leq K \), return to step 4, else continue. |                                      |
| 9 Return \( \hat{H}_i \) |                                      |

3) Compressive sampling matching pursuit (CoSaMP) [21] is extremely efficient for practical problems, since it requires only matrix or vector multiplied with the sampling matrix. The estimated channel matrices \( \{ H_i \} (i=1,2,\ldots,ML) \)

| Table 4. Algorithm of compressive sampling matching pursuit |
|-------------------------------------------------------------|
| **Procedure**                                                                 | **Operations**                                      |
| **Input:** The received measurements of all users \( \{ Y_i \} \), the pilot training matrix \( X \), the total sparsity level \( K \). |                                      |
| **Output:** The estimated channel matrices \( \{ H_i \} = \{ h_1, \ldots, h_N \} \). |                                      |
| 1 Compute \( \{ \hat{Y}_i \} \) and \( \hat{X} \) from \( \{ Y_i \} \) and \( X \) as in (4a) and (4b). |                                      |
| 2 Initialize, residual \( r_0 = \hat{y} \), \( \Lambda_0 = \emptyset \), \( \psi_0 = \emptyset \), \( t=1 \). |                                      |
| 3 Compute \( u = \{ r_{t-1}, \hat{X} \} \), select the 2\( K \) biggest from \( u \), and construct the set \( J_0 \) with the corresponding index \( j \). |                                      |
| 4 Let \( \Lambda_t = \Lambda_{t-1} \cup \{ J_0 \}, \psi_t = \psi_{t-1} \cup \{ \hat{X}_{j_0} \} \) (for all \( j \in J_0 \)). |                                      |
| 5 Compute \( \hat{h}_t = \arg \min_h \| \hat{y} - \psi_t h_t \| = (\psi_t^T \psi_t)^{-1} \psi_t^T \hat{y} \) |                                      |
| 6 Choose \( \hat{h}_{k}(k=1,\ldots,K) \) with the biggest from \( \hat{h}_t \), and construct the corresponding \( \psi_{ik} \), refresh \( \Lambda_t = \Lambda_{ik} \). |                                      |
| 7 Refresh \( r_t = \hat{y} - \psi_{ik} \hat{h}_{ik} = \hat{y} - \psi_{ik} (\psi_{ik}^T \psi_{ik})^{-1} \psi_{ik}^T \hat{y} \) \( t=t+1 \), if \( t \leq K \), return to step 4, else continue. |                                      |
| 9 Return \( \hat{H}_i \) |                                      |
4) Stagewise orthogonal matching pursuit (StOMP) [22] is significantly faster than OMP on large scale problems with sparse solutions owing to a number of stages. The estimated channel matrices \( \{ H_i \} \) \((i=1,2,\ldots,ML)\)

Table 5. Algorithm of stagewise orthogonal matching pursuit

| Procedure | Operations |
|------------|------------|
| **Input:** | The received measurements of all users \( \{ Y_i \} \), the pilot training matrix \( X \), the total sparsity level \( K \). |
| **Output:** | The estimated channel matrices \( \{ H_i \} = \{ h_i^1, \ldots, h_i^N \} \). |
| 1 | Compute \( \hat{Y}_i \) and \( \hat{X} \) from \( \{ Y_i \} \) and \( X \) as in (4a) and (4b). |
| 2 | Initialize, residual \( r_0 = \hat{y} \), \( \Lambda_0 = \emptyset \), \( \psi_0 = \emptyset \), \( t = 1 \). |
| 3 | Compute \( u = \left\{ r_{t-1}, \hat{X} \right\} \), select those that are bigger than threshold \( T_h \) from \( u \), and construct the set \( J_0 \) with the corresponding index \( j \) |
| 4 | Let \( \Lambda_t = \Lambda_{t-1} \cup \{ J_0 \}, \psi_t = \psi_{t-1} \cup \{ \hat{X}_{j_0} \} \) (for all \( j \in J_0 \)). if \( \Lambda_t = \Lambda_{t-1} \), then goto step 8. |
| 5 | Compute \( \hat{h}_i = \arg \min_k \| \hat{y} - \psi_i \hat{h}_i \| = \left( \psi_i^T \psi_i \right)^{-1} \psi_i^T \hat{y} \) |
| 6 | Refresh \( r_t = \hat{y} - \psi_i \hat{h}_i = \hat{y} - \psi_i \left( \psi_i^T \psi_i \right)^{-1} \psi_i^T \hat{y} \) |
| 7 | \( t = t + 1 \), if \( t \leq K \), return to step 4, else continue. |
| 8 | Return \( \hat{H}_i \) |

5) Subspace pursuit (SP) is similar to CoSaMP. The main difference between them is in the manner where new candidates are added to the list [23]. In each iteration, in SP algorithm, only \( K \) new candidates are added, while CoSaMP algorithm adds \( 2K \) vectors. This makes SP computationally more efficient than CoSaMP. The estimated channel matrices \( \{ H_i \} \) \((i=1,2,\ldots,ML)\)

Table 6. Algorithm of subspace pursuit

| Procedure | Operations |
|------------|------------|
| **Input:** | The received measurements of all users \( \{ Y_i \} \), the pilot training matrix \( X \), the total sparsity level \( K \). |
| **Output:** | The estimated channel matrices \( \{ H_i \} = \{ h_i^1, \ldots, h_i^N \} \). |
| 1 | Compute \( \hat{Y}_i \) and \( \hat{X} \) from \( \{ Y_i \} \) and \( X \) as in (4a) and (4b). |
| 2 | Initialize, residual \( r_0 = \hat{y} \), \( \Lambda_0 = \emptyset \), \( \psi_0 = \emptyset \), \( t = 1 \). |
| 3 | Compute \( u = \left\{ r_{t-1}, \hat{X} \right\} \), select the \( K \) biggest from \( u \), and construct the set \( J_0 \) with the corresponding index \( j \) |
| 4 | Let \( \Lambda_t = \Lambda_{t-1} \cup \{ J_0 \}, \psi_t = \psi_{t-1} \cup \{ \hat{X}_{j_0} \} \) (for all \( j \in J_0 \)). |
| 5 | Compute \( \hat{y}_r = \left( \psi_i^T \psi_i \right)^{-1} \psi_i^T \hat{y} \), sort \( \hat{y}_r \) in descend order and choose the \( K \) biggest so as to form \( \hat{h}_{ik} \). |
| 6 | Refresh \( r_t = \hat{y} - \psi_i \hat{h}_{ik} \), if \( \| r_t \|_2 \geq \| r_{t-1} \|_2 \), then quit the iteration |
| 7 | \( t = t + 1 \), if \( t \leq K \), return to step 4, else continue. |
| 8 | Return \( \hat{H}_i \) |
6) Sparsity adaptive matching pursuit (SAMP) is a generalization of existing greedy algorithms as both the OMP and the SP can be viewed as its special cases. It follows the “divide and conquer” principle through stage by stage estimation of the sparsity level and the true support set of the target signals [24]. The estimated channel matrices \( \{ H_i \} (i=1,2,...,ML) \)

**Table 7. Algorithm of sparsity adaptive matching pursuit**

| Procedure | Operations |
|-----------|------------|
| **Input:** | The received measurements of all users \( \{ Y_i \} \), the pilot training matrix \( X \), the total sparsity level \( K \), the step size \( s \). |
| **Output:** | The estimated channel matrices \( \{ H_i \} = \{ h_i^1, ..., h_i^N \} \). |
| 1 | Compute \( \{ Y_i^\hat{\imath} \} \) and \( \hat{X} \) from \( \{ Y_i \} \) and \( X \) as in (4a) and (4b). |
| 2 | Initialize, residual \( r_0 = \hat{y} \), \( \Lambda_0 = \emptyset \), \( \psi_0 = \emptyset \), \( t=1 \), step \( L=s \). |
| 3 | Compute \( u = \left\{ r_{t-1}^i, \hat{X} \right\} \), select the \( L \) biggest from \( u \), and construct the set \( J_0 \) with the corresponding index \( j \). |
| 4 | Let \( \Lambda_t = \Lambda_{t-1} \cup \{ J_0 \}, \psi_t = \psi_{t-1} \cup \{ \hat{X}_{J_0} \} \) (for all \( j \in J_0 \)). |
| 5 | Compute \( \hat{h}_t = \arg \min_h \| \hat{y} - \psi_t h_t \| = \left( \psi_t^T \psi_t \right)^{-1} \psi_t^T \hat{y} \). |
| 6 | Choose the \( L \) biggest in \( | \hat{h}_t | \) so as to form \( \hat{h}_t^L \), the corresponding \( \psi_t^L \) and its index set \( F = \Lambda_t^L \). |
| 7 | Refresh \( r_t = \hat{y} - \psi_t^L \hat{h}_t = \hat{y} - \psi_t^L \left( \psi_t^L \psi_t^L \right)^{-1} \psi_t^L \hat{y} \). |
| 8 | If \( \| r_t \|_2 \geq \| r_{t-1} \|_2 \), then \( L = L+s \), goto step 3. |
| 9 | \( t = t+1 \), if \( t \leq K \), return to step 3, else continue. |
| 10 | Return \( \hat{H}_i \). |

7) By adding regularization into the SAMP, we propose an algorithm of modified sparsity adaptive matching pursuit (MSAMP). \( \{ H_i \} (i=1,2,...,ML) \) as shown in table 8.
Table 8. Algorithm of modified sparsity adaptive matching pursuit (MSAMP)

| Procedure | Operations |
|-----------|------------|
| **Input:** | The received measurements of all users \( \{ Y_i \} \), the pilot training matrix \( X \), the total sparsity level \( K \), the step size \( s \). |
| **Output:** | The estimated channel matrices \( \{ H_i \} = \{ h_1, \ldots, h_N \} \). |
| 1 | Compute \( \{ \hat{Y}_i \} \) and \( \hat{X} \) from \( \{ Y_i \} \) and \( X \) as in (4a) and (4b). |
| 2 | Initialize, residual \( r_0 = \hat{y} \), \( \Lambda_0 = \emptyset \), \( \psi_0 = \emptyset \), \( t = 1 \), step \( L = s \). |
| 3 | Compute \( u = \langle r_{t-1}, \hat{X} \rangle \), select the \( L \) biggest from \( u \), and construct the set \( J \) with the corresponding index \( j \). |
| 4 | Regularize: find subset \( J_0 \) in \( J \), such that \( | u(i) | \leq 2 | u(j) | \), for all \( i, j \in J_0 \), choose \( J_0 \) with maximal energy \( \sum_j | u(j) |^2 \). |
| 5 | Let \( \Lambda_t = \Lambda_{t-1} \cup \{ J_0 \}, \psi_t = \psi_{t-1} \cup \{ \hat{X}_{j_0} \} \) (for all \( j \in J_0 \)). |
| 6 | Compute \( \hat{h}_t = \arg \min_h \| \hat{y} - \psi_t h \| = \left( \psi_t^T \psi_t \right)^{-1} \psi_t^T \hat{y} \). |
| 7 | Choose the \( L \) biggest in \( | \hat{h}_t | \) so as to form \( \hat{h}_{1L} \), the corresponding \( \psi_{1L} \) and its index set \( F = \Lambda_{1L} \). |
| 8 | Refresh \( r_t = \hat{y} - \psi_{1L} \hat{h}_t = \hat{y} - \psi_{1L} \left( \psi_{1L}^T \psi_{1L} \right)^{-1} \psi_{1L}^T \hat{y} \). |
| 9 | if \( \| r_t \|_2 \geq \| r_{t-1} \|_2 \), then \( L = L + s \), goto step 3. |
| 10 | \( t = t + 1 \), if \( t \leq K \), return to step 3, else continue. |
| 11 | Return \( \hat{H}_t \). |

4. Simulation results
In this section, we present the performance of the multi-user massive MIMO system via simulation. We will compare the performances of various measurement matrices and various CS optimal algorithms in channel estimation. In the simulation, the system consists of one base station (BS) with \( N = 100 \) antennas, and \( L = 6 \) users each with 8 antennas. In the channel estimation, a training sequence with length of \( T = 60 \) is transmitted by BS. It is assumed that the training sequence \( X \in \mathbb{C}^{N \times T} \), channel matrix \( H \in \mathbb{C}^{ML \times N} \), and the added white Gaussian noise \( W \) with zero mean and 1/2 variance.

To evaluate the error-rate performances of various algorithms in channel estimation, the normalized mean square error (NMSE) is exploited as follows

\[
\text{NMSE} = \frac{1}{G} \sum_{n=1}^{G} \left\| H_n - \tilde{H}_n \right\|_F^2
\]

Where \( H_n \) and \( \tilde{H}_n \) denote the real channel matrix and estimated channel matrix respectively. \( \left\| (\cdot) \right\|_F \) denotes the Frobenius norm. \( G \) denotes the iterations of Monte-Carlo simulation.

In Fig. 2, it is observed that partial Hadamard and Toeplitz matrices outperform other measurement matrices. The sparse random matrix has the worst recovery performance in channel estimation. However, as SNR increases, their error–rate performances should tend to approach all together. Moreover, as SNR grows beyond some extent, the Gaussian and circulant matrix seem to get better.
In Fig. 3, it is observed that various algorithms have different recovery performances under the same number of antennas and the same number of measurements based on Toeplitz matrix. The OMP has the worst recovery performance in channel estimation. In addition, both SAMP and SP have nearly the same recovery performance in channel estimation. However, as SNR increases, the error–rate performances of all the algorithms should tend to approach all together.

Fig. 4 presents the error-rate performance versus the sparsity level $K$ based on Circulant matrix. Under the same channel sparsity, SAMP, SP, and CoSaMP have the nearly same recovery performances. MSAMP has the best property of NMSE among them. Meanwhile OMP is the worst among them. It can be seen that if sparsity $K$ gets bigger, there will be less statistic information about the common support, the probability of reconstruct the channel information should get smaller.
Thereby, the performance should deteriorate as the sparsity increase under the same number of antennas and the same number of measurements.

![Figure 4. NMSE v.s. sparsity level $K$ based on Circulant matrix](image)

In Fig. 4, it is shown that the performances of the SAMP far exceed that of other algorithms. On one hand, all the other algorithms start to fail when $K \geq 15$. On the other hand, the SAMP still can afford until sparsity $K \geq 19$. Meanwhile, it can be seen in SAMP that, the step $s=5$ should have the best recovery performance than the step $s=1$ and 10. MSAMP with $s=10$ is better than SAMP with $s=10$.

![Figure 5. The probability of exact reconstruction vs. sparsity level $K$ via Gaussian matrix](image)

In Fig. 5, it is shown that the performances of the SAMP far exceed that of other algorithms. On one hand, all the other algorithms start to fail when $K \geq 15$. On the other hand, the SAMP still can afford until sparsity $K \geq 19$. Meanwhile, it can be seen in SAMP that, the step $s=5$ should have the best recovery performance than the step $s=1$ and 10. MSAMP with $s=10$ is better than SAMP with $s=10$.

In Fig. 6, we can see that SAMP and MSAMP have the best recovery performance. It is also observed that when the number of measurement is insufficient to guarantee the exact recovery, the probability of exact recovery of SAMP depends on its step size. For Gaussian matrix, SAMP with a smaller step size gets a higher chance of reconstructing the channel information. In addition, it is
shown that other algorithms have nearly the same recovery performances when the number of measurements is lower than certain value.

![Figure 6](image)

**Figure 6.** The probability of exact reconstruction vs. number of measurements via Gaussian matrix

5. Conclusions
In this paper, we have investigated the error-rate performances of various measurement matrices and various CS optimization algorithms in multi-user massive MIMO system. In addition, we have proposed an algorithm of modified SAMP. In the reconstruction of channel information, we have drawn comparisons among all the CS optimization algorithms in terms of the sparsity level and the number of measurements. It is confirmed that SAMP is very appropriate for reconstructing sparse channel information in massive MIMO system. Meanwhile, it can be inferred that some measurement matrix is suitable for certain optimal CS algorithms in channel estimation.

Acknowledgments
This work is supported by the Natural Science Foundation of Fujian Province (2020J01711), Li Shangda Discipline Construction Fund (ZC2016008), and Research Start-up Fund of Jimei University (ZQ2019021).

References
[1] F. Rusek, D. Persson, B.K. Lau, *et al.*, Scaling up MIMO: Opportunities and challenges with very large arrays, IEEE Signal Processing Magazine, 30(1), 2013(40-60).
[2] S.A. Ramprashad, G. Caire, H.C. Papadopoulos, “Cellular and network MIMO architectures: MU-MIMO spectral efficiency and costs of channel state information,” IEEE ACSSC’09, CA, 2009, pp.1811-1818.
[3] J. Hoydis, S.T. Brink, M. Debbah, Massive MIMO in the UL/DL of cellular networks: How many antennas do we need? IEEE Journal on selected areas in communications, 31(2), 2013(160-171).
[4] Y. C. Eldar and G. Kutyniok, Compressed Sensing: Theory and Applications, Cambridge, U.K.: Cambridge University Press, 2012.
[5] S.G. Mallat and Z. Zhang, Matching pursuits with time-frequency dictionaries, IEEE Transactions on signal processing, 41(12), 1993(3397-3415).
[6] J.A. Tropp and A.C. Gilbert, Signal Recovery From Random Measurements via Orthogonal Matching Pursuit, IEEE Transactions on Information Theory, 53(12), 2007(4655-4666).
[7] Z. Gao, L. Dai, W. Dai, et al., “Block compressive channel estimation and feedback for FDD massive MIMO,” IEEE InfoCom Proceedings, Hong Kong, 2015, pp.49-50.

[8] L. Dai, Z. Gao, Z. Wang, “Joint channel estimation and feedback with low overhead for FDD massive MIMO systems,” IEEE/CIC ICCC 2015 Symposium on Signal Processing for Communications, Shenzhen, 2015, pp. 1-6.

[9] Y. Nan, L. Zhang, X. Sun, Efficient downlink channel estimation scheme based on block-structured compressive sensing for TDD massive MU-MIMO systems, IEEE Wireless communications letter, 4(4),2015(345-348).

[10] L. Dai, Z. Gao, Z. Wang, et al., “Spectrum-efficient superimposed pilot design based on structured compressive sensing for downlink large-scale MIMO systems,” 2014 XXXIth URSI General Assembly and Scientific Symposium (URSI GASS), Beijing, 2014, pp. 1-4.

[11] D. Lee, MIMO OFDM Channel estimation via block stagewise orthogonal matching pursuit, IEEE Communications letters, 20(10), 2016(2115-2118).

[12] D. Tse, P. Viswanath, Fundamentals of wireless communication, Cambridge, U.K.: Cambridge University Press, 2005.

[13] Y. Zhou, M. Herdin, A.M. Sayeed, “Experimental study of MIMO channel statistics and capacity via virtual channel representation,” IEE Proceedings—Radar, Sonar and Navigation, 2006, pp. 1-18.

[14] J. Hoydis, C. Hoek, T. Wild, et al., Channel measurements for large antenna arrays, Journal of applied physics, 91(12), 2012(811-815).

[15] H. Qiao, P. Pal, Generalized Nested Sampling for Compressing Low Rank Toeplitz Matrices, IEEE Signal processing letters, 22(11), 2015(1844-1848).

[16] Y. Chen, C. Cheng, Q.Y. Sun, Reconstruction of Sparse Wavelet Signals From Partial Fourier Measurements, IEEE Signal processing letters, 22(12), 2015(2299-2303).

[17] S. Haghighatshoar, E. Abbe, E. Telatar, “Adaptive sensing using deterministic partial Hadamard matrices,” 2012 IEEE International Symposium on Information Theory Proceedings, Cambridge USA, 2012, pp. 1842-1846.

[18] H. Rauhut, “Circulant and Toeplitz Matrices in Compressed Sensing,” Signal Processing with Adaptive Sparse Structured Representations (SPARS’09), 2009, pp.1-6.

[19] E. Candès, J. Romberg, T. Tao, Robust Uncertainty Principles: Exact Signal Reconstruction from Highly Incomplete Frequency Information, IEEE Transactions on Information Theory, 52(2), 2006(489-509).

[20] D. Needell, R. Vershynin, Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit, IEEE Journal of selected topics in signal processing, 4(2), 2010(310-316).

[21] D. Needell, J.A. Tropp, CoSaMP: Iterative signal recovery from incomplete and inaccurate samples, Applied and Computational Harmonic Analysis, 26(3), 2009(301-321).

[22] D.L. Donoho, Y. Tsaig, I. Drori, et al, Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit, IEEE Transactions on information theory, 58(2), 2012(1094-1121).

[23] W. Dai and O. Milenkovic, Subspace Pursuit for Compressive Sensing Signal Reconstruction, IEEE Transactions on Information Theory, 55(5), 2009(2230-2249).

[24] T.T. Do, L. Gan, N. Nguyen, et al., “Sparsity adaptive matching pursuit algorithm for practical compressed sensing.” 2008 42nd Asilomar Conference on Signals, Systems and Computers, Pacific Grove, USA, 2008, pp.581-587.