

# A Robust Recovery of Block Sparse Channels in Massive MIMO-OFDM Systems



Jing Zhang, Ying Su\*

Department of Communication Engineering, Shanghai Normal University, Shanghai 200234, China  
{jannety, yingsu}@shnu.edu.cn

Received 30 June 2020; Revised 16 November 2020; Accepted 6 January 2021

**Abstract.** The recovery algorithms of the finite impulse responses (FIRs) of multipath channels via a single-measurement-vector and a multiple-measurement-vector, respectively, are investigated in the massive multiple input multiple output orthogonal frequency division multiplexing (MIMO-OFDM) systems. Based on the characteristics of the spatial-temporal block sparse structure of the multipath channels, the phase-shift orthogonal comb pilots are designed and inserted into the symbols to ensure the column uncorrelatedness of the sensing matrix. These pilot tones are spread among all transmit antennas. For the purpose of resolving the difficulty of approximating the unknown sparsity without overestimation, the robust sparsity adaptive matching pursuit (RSAMP) algorithms to be used in both measurement scenarios are proposed. The residuals in these algorithms are smoothed within several iterations but achieving much more acceptable halting state. The proposed algorithms are proven to be immune to the size of the support set and a flexible number of measurement vectors. Simulations demonstrate that the normalized mean square error performance of these algorithms is comparable to those of the orthogonal matching pursuit algorithm and the subspace pursuit algorithm.

**Keywords:** massive MIMO, OFDM, channel estimation, compressive sensing, adaptive estimation

## 1 Introduction

The fifth generation (5G) mobile communication has strict requirements on the data rate, the time delay and the energy efficiency. Massive multiple input multiple output (MIMO), which refers to a base station equipped with dozens or even hundreds of antennas for carrying services to users, is the key scheme to meet these demands. Compared with the multi-antenna schemes that are widely used in the fourth generation systems, the massive MIMO will provide higher spectrum efficiency along with improved energy efficiency, spatial resolution and a more effective transceiver design [1]. Meanwhile, the orthogonal frequency division multiplexing (OFDM) is another effective candidate to achieve high data rate transmissions. It works better to resist the frequency selective fading and therefore acquire higher spectrum utilization. The combination of massive MIMO and OFDM working together can achieve unprecedented throughputs. Nevertheless, these advantages brought by massive MIMO-OFDM systems depend on the accuracy level of the estimation of the channel state information (CSI).

OFDM employs discrete Fourier transform (DFT) pairs and divides the wideband channel into multiple narrowband sub-channels to avoid channel fading, which makes each OFDM sub-channel to be regarded as nearly flat fading. The maximum likelihood (ML) algorithm is usually used for the joint channel estimation and the symbol detection [2]. Due to the lack of a channel correlation matrix plus the large number of subcarriers, the estimation with a minimum mean square error (MMSE) is difficult to implement. Furthermore, the high dimension of the massive MIMO channel limits these aspects: the direct calculations of the inversion or pseudo-inversion matrix, the autocorrelation matrix as well as the singular value decompositions in the processes of the least square (LS) estimation and the MMSE

---

\* Corresponding Author

estimation [3]. It should also be noted that the number of pilots is proportional to the number of users participating in the transmissions in the uplink and the number of transmit antennas of a base station in the downlink. All these methods will generate an overwhelming pilot overhead and thus severely reduce the data rate. Also because of the limitations of the channel coherence time and the correlation bandwidth, the pilot resources available in massive MIMO-OFDM systems are always not sufficient. The resultant channel estimation is therefore vulnerable to the non-orthogonal pilot sequence, which makes it difficult for these methods to achieve the best estimation performance.

The pursuing of an accurate channel estimation that has an acceptable pilot overhead and limited computational complexity is becoming a hot topic in current research activities. And the pilot-based and semi-blind approaches are the two main areas [4]. It is well known that the compressed sensing (CS) mainly uses a specific sensing or a dictionary matrix to project sparse or compressible high-dimensional signals into a low-dimensional space. This approach then uses a linear or a nonlinear reconstruction algorithm to recover the original signal based on the sparse prior knowledge. In broadband wireless communications, many real channels show sparsity in certain domains. Therefore, when the channel statistics are unknown, the CS is an attractive method for the estimation of a sparse channel by introducing a short pilot sequence.

The current channel estimation algorithms mainly focus on the improvement of the accuracy along with a pilot overhead reduction. By exploiting the space-time structure of a specific type of transmission frames in the massive MIMO-OFDM systems, the finite impulse response (FIR) of a sub-channel becomes sparse in its both spatial and temporal domains. While the recovery methods based on the CS are widely studied, the pilot allocation and then optimization are also investigated for the reduction of a pilot overhead and the improvement of the estimation performance. One distributed sparse channel estimation method is proposed to minimize the cross-correlation of the measurement matrix, in which the pilot position is optimized by a genetic algorithm [5]. In the downlink direction, the same pilot subcarriers instead of the orthogonal subcarriers can be allocated to each transmit antenna, thus reducing the pilot overhead. When the number of pilots is large enough, the channel can be reconstructed by minimizing the correlations of the Fourier sub-matrix [6]. In order to minimize the mean square error (MSE) of the spatial correlation channel in the orthogonal space-time block coding (OSTBC) systems, superimposed pilot sequences are designed [7]. The weighted homotopy algorithm can achieve decent performance by using the characteristics of the sparse channel with relatively low pilot overhead [8]. To relax the complexities of the pilot design, a semi-blind algorithm for multiuser MIMO-OFDM systems is proposed [9].

In general, the FIR energy of a sub-channel gathers on several taps, while the rest taps are either zero or close to zero due to the limited scattering effect at the end of the transmit antennas. This characteristics result in structured sparsity of the channel parameters in the massive MIMO-OFDM systems. When multiple measurements are available, the adjacent channel blocks in the spatial-temporal domain will share a common support set. By exploring the space-time correlation of the sub-channels, the channel can be reconstructed by several consecutive symbols, thus further reducing the pilot overhead. Based on the sparsity properties of the delay domain and spatial domain, the base expansion model (BEM) is used to reduce the estimated parameters [10]. A modified spatial BEM model is formulated to a block-sparse signal and a quasi-block simultaneous orthogonal matching pursuit (OMP) is introduced to recover the channel parameters [11]. Furthermore, one pilot search algorithm with the particle swarm optimization is specifically studied for a vector OFDM-MIMO system. An algorithm of a sparsity adaptive matching pursuit (SAMP) with a variable threshold achieves high recovery accuracy [12]. In addition, the sparse Bayesian learning is investigated for the joint channel estimation and sequence detection in the space-time trellis coded MIMO-OFDM systems [13]. By incorporating three-dimensional sparse representation into a tensor model, the OMP algorithm carries on greedy search on each dimension of the measured tensor data [14]. SAMP algorithm is an extension of the existing greedy algorithms in the cases of unknown sparsity [15]. When the measurement matrix is a partial discrete Fourier transform matrix in an OFDM system, an improved SAMP can optimize the initial support vector along with an iteration threshold to achieve better performance [16]. Adaptive algorithms without sparsity prior need to estimate the sparsity by residual projection. A regularization method of subspace tracking based on a variable step size is discussed in [17].

The sparse pilot design and channel recovery algorithms for the massive MIMO frequency division duplexing (FDD) systems in a CS frame are also studied. The uplink and downlink of the massive

MIMO-FDD systems use different frequency division channels thus the channel parameters do not have reciprocity. A two-stage weighted block  $l_1$ -norm minimization algorithm is examined for a block sparse channel in an angle domain [18]. The structured sparsity of a channel matrix can be utilized to design a joint training and feedback scheme by using the compressed sampling matching tracking (CoSaMP) algorithm that simultaneously reduces the pilot overhead of the downlink and the feedback overhead of the uplink [19]. The framework of a closed-loop estimation with 1-bit feedback is to learn the minimum required pilot overhead for a certain target mean square error (MSE) [20]. By a closed-loop framework of the pilot and CSI feedback bit adaptation, not only can the joint sparse characteristics of a multi-user massive MIMO channel be utilized that improves the CSI estimation performance, but also has the built-in learning ability to recover CSI successfully that uses less pilot and feedback resources under an unknown and time-varying channel sparse level [21]. In order to enhance the robustness of the sparse recovery algorithm, the partial support set information is extracted from the channel correlation characteristics. The minimum absolute contraction and selection (LASSO) algorithm is to alleviate the strict requirement for the sparsity [22]. When the information of the required training cost is formulated in the model of the weighted  $l_1$ -norm minimization, the pilot overhead can be significantly reduced by using the channel partial support set [23]. The approximate message passing (AMP) algorithm can also be used when the sparse prior information of channel parameters is assumed [24].

This paper studies the problem that exists in the sparse channel recovery from the aspects of pilot design and adaptive reconstruction algorithms in the MIMO-OFDM system. The multipath channel is considered as a structured sparse one in spatial and temporal domains. After the discussion and comparison of the random pilot along with general greedy algorithms, this paper focuses on the recovery method of FIR by using an adaptive algorithm with the phase-shift orthogonal pilot insertion. Robust SAMP (RSAMP) algorithms with an improved halting condition in single-measurement-vector and multiple-measurement-vector scenarios are examined to improve the MSE performance. Unlike the allocation of the random pilot on the same subcarriers, the phase-shift orthogonal pilot can ensure that the support vectors selected each time have no correlation with the previous selected support set. Therefore, the conditional number of the matrix composed of support vectors is relatively small that enhances numerical reliability. Moreover, compared with the methods that use multi-threshold to approximate the real sparsity, this RSAMP algorithm smoothens the residuals during the multi-step iterations in account of inevitable observation noise. Once the smoothed residuals are smaller than the SNR at the receiver, the algorithm stops and thus the over estimation of the sparsity is effectively avoided. Although the preset threshold results in a deviation in the approximation of sparsity, it still shows little effect on the normalized MSE.

This paper is organized as follows. Section 2 describes the spatial-temporal sparse channel and introduces the received signal model. Section 3 discusses the pilot pattern and phase-shift pilot. In section 4 the recovery algorithms with single and multiple measurement vectors are presented. The simulation results and analysis are included in section 5. And lastly section 6 has the conclusion.

## 2 Received Signal Model

### 2.1 Block Sparse MIMO Channel

Consider a massive MIMO system with the number of transmit antennas denoted as  $N_G$  and the number of receive antennas as  $N_Q$ . Its wireless space-time multipath channel is represented as a  $N_Q \times N_G$  matrix. The entry  $h_{q,g}(t)$ ,  $q = 1, \dots, N_Q$ ,  $g = 1, \dots, N_G$  represents one single sub-channel from transmit antenna  $g$  to receive antenna  $q$ , which is expressed as

$$h_{q,g}(t) = \sum_{l=0}^{L-1} \alpha_{q,g}^l \delta(t - \tau_{q,g}^l), \quad (1)$$

where  $\alpha_{q,g}^l$  denotes the complex fading of the  $l$ -th path, and  $\tau_{q,g}^l$  denote the corresponding time delay.

To recover one channel with a limited number of computations, its space-time structure of sub-channels needs to be explored. Firstly the space interval of the antenna array at a base station is relatively small compared with the transmission distance of signals. Secondly there exist common scatters among different transmit antennas and a certain receive antenna. More importantly, most of the channel impulse

responses (FIRs) overlap. Thus, the FIRs of all sub-channels have very similar path delay. And the lengths of subcarrier FIRs can be treated as the same length. Under these circumstances, the FIR of one sub-channel is further expressed as

$$\mathbf{h}_{q,g}(n) = [h_{q,g(0)}(n), h_{q,g(1)}(n), \dots, h_{q,g(L-1)}(n)]^T, \forall g. \quad (2)$$

Additionally, during the period of transmitting  $N$  symbols, the time delay of the sub-channels varies slowly. Hence, the support sets of sub-channels at different time instant are regarded as the same, and represented by

$$\text{supp}\{\mathbf{h}_{q,g}(n)\} = \text{supp}\{\mathbf{h}_{q,g}(n+1)\} = \dots = \text{supp}\{\mathbf{h}_{q,g}(n+N-1)\}. \quad (3)$$

Considering the fact that the antennas at one base station are usually placed compact and share common scatters, all the sub-channels among one receive antenna and multiple transmit antennas thus share a common space-domain sparse characteristic, i.e.,

$$\text{supp}\{\mathbf{h}_{q,1}(n)\} = \text{supp}\{\mathbf{h}_{q,2}(n)\} = \dots = \text{supp}\{\mathbf{h}_{q,N_G}(n)\}. \quad (4)$$

Assume the number of subcarriers of OFDM is  $K$ . Orthogonal subcarrier divides the multipath sub-channels into frequency-domain parallel sub-channels. Based on the DFT and the space-time analysis, it can be assumed that the first  $L$  element ahead of  $h_{q,g}(n)$  could be nonzero but the hinder  $K-L$  elements must be zeros. The channel also remains unchanged during successive symbol periods. This analysis shows that the channel is of structured-sparsity. That is, for one receive antenna, its channel has a fixed sparse structure as

$$\mathbf{h}_q = [h_{q,1(0)}, \dots, h_{q,1(L-1)}, 0, \dots, 0, h_{q,2(0)}, \dots, h_{q,2(L-1)}, 0, \dots, 0, \dots, h_{q,N_G(0)}, \dots, h_{q,N_G(L-1)}, 0, \dots, 0]^T. \quad (5)$$

Eq. (5) indicates that  $\mathbf{h}_{q,g}$  is a block sparse signal in which non-zero coefficients is clustered. In order to represent the block sparsity,  $\mathbf{h}_{q,g}$  can be regarded as a concatenation of blocks, i.e.,

$$\mathbf{h}_q = [\mathbf{h}_{q,1}^T, \mathbf{h}_{q,2}^T, \dots, \mathbf{h}_{q,N_G}^T]^T, \quad (6)$$

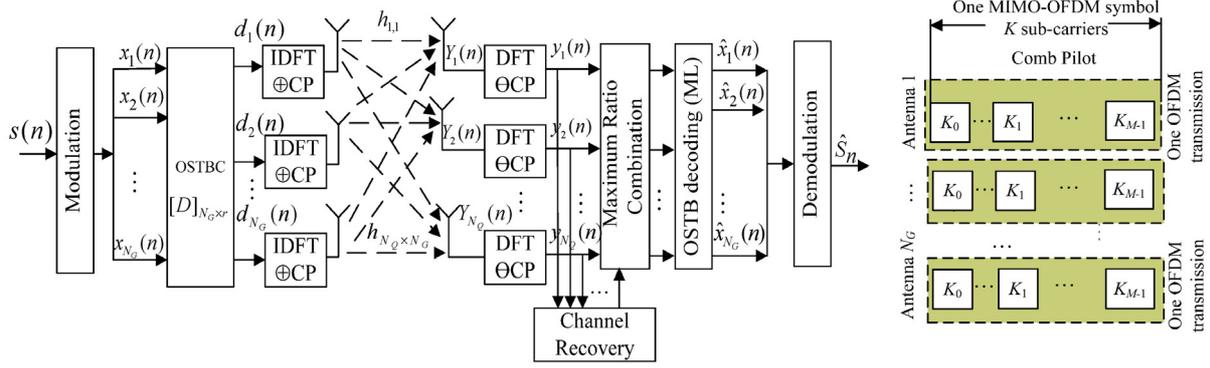
where  $\mathbf{h}_{q,j}$   $j = 1, \dots, N_G$  is a sub-block. Non-zero elements are concentrated in the sub-block.

## 2.2 Received Signal Model

In an end-to-end massive MIMO-OFDM system, data symbols are transmitted through a space-time multipath channel with  $K$  subcarriers of OFDM as shown in Fig. 1. The  $n$ th modulated high-speed data symbol  $S(n)$  is divided into multiple low-speed sub-blocks  $\{x_1(n), \dots, x_{N_G}(n)\}$ . After the orthogonal space-time block coding (OSTBC) with coding rate  $r$ , the signals  $\{d_1(n), \dots, d_{N_G}(n)\}$  at time instant  $n$  are obtained. Meanwhile, pilots are inserted into the sub-blocks with a proper interval. And after the insertion of cyclic prefix (CP), the signals are transformed via inverse DFT, and then transmitted through the corresponding antennas. Over the channel transmission, the signals at pilot tones are measured for the channel recovery. Upon received at each receive antenna, the operation of DFT along with the removal of CP at each receive antenna are performed. With further maximum ratio combination and space-time block decoding, the symbols are demodulated out at the receiver. The received signal after the CP removal can be described as

$$\mathbf{y}_q(n) = \sum_{g=1}^{N_G} \{s_g(n)\}_{diag} \mathbf{F} \mathbf{h}_{q,g}(n) + \boldsymbol{\xi}_q(n), \quad (7)$$

where  $s_g(n) = [s_g(1, n), \dots, s_g(K, n)]^T$  denotes the symbol on subcarriers of the transmit antenna  $g$  at the time instant  $n$ .  $\mathbf{F}$  is the DFT matrix.  $\mathbf{F} = [\boldsymbol{\Omega}(0), \dots, \boldsymbol{\Omega}(K-1)]^T$ ;  $\boldsymbol{\Omega}(i) = 1/\sqrt{K}[\omega_K^{i0}, \dots, \omega_K^{i(K-1)}]^T$ ;  $\omega_K^{ik} = e^{j2\pi ik/K}$ ;  $i, k = 0, \dots, K-1$ .  $\boldsymbol{\xi}_q(n)$  denotes the complex Gaussian noise with variance  $\sigma^2$ .



**Fig. 1.** Transmission flow scheme and the pilot pattern of the massive MIMO-OFDM system

The signals are processed at any arbitrary receive antenna. To simplify the expression, the subscript  $q$  is omitted, and (7) is rewritten in the general format below:

$$\mathbf{y}(n) = \mathbf{A}(n)\mathbf{h}(n) + \boldsymbol{\zeta}(n). \quad (8)$$

In (8),  $\mathbf{y}(n) = [y(n, k_0), \dots, y(n, K_{M-1})]^T$  is the measurement signal at pilot tones  $K_m$ ,  $m = 0, \dots, M-1$ ;  $\mathbf{A}(n) = [\mathbf{A}_1(n), \dots, \mathbf{A}_{N_G}(n)]^T$  is the sensing matrix with the dimension of  $M \times K \cdot N_G$ , where  $\mathbf{A}_g(n) = [s_g(n, k_0)\boldsymbol{\Omega}(k_0), \dots, s_g(n, k_{M-1})\boldsymbol{\Omega}(k_{M-1})]$ ;  $\mathbf{h}(n) = [\mathbf{h}_1, \dots, \mathbf{h}_{N_G}]^T$ , where  $\mathbf{h}_g$  represents the FIR between the transmit antenna  $g$  and the receive antenna; and  $\boldsymbol{\zeta}(n) = [\zeta(n, K_0), \dots, \zeta(n, K_{M-1})]^T$ .

$$\mathbf{A}_g(n) = \begin{bmatrix} S_{g,K_0}(n) & S_{g,K_1}(n)e^{-\frac{(2\pi)}{K}K_0} & \dots & S_{g,K_{M-1}}(n)e^{-\frac{(2\pi)}{K}K_0(L-1)} \\ \vdots & \ddots & & \vdots \\ S_{g,K_0}(n) & S_{g,K_1}(n)e^{-\frac{(2\pi)}{K}K_{M-1}} & \dots & S_{g,K_{M-1}}(n)e^{-\frac{(2\pi)}{K}K_{M-1}(L-1)} \end{bmatrix}. \quad (9)$$

We rewrite  $\mathbf{A}_g(n)$  in (9) as

$$\mathbf{A}_g(n) = \mathbf{s}_g(n) \otimes \boldsymbol{\Omega}_{ML}, \quad (10)$$

where  $\otimes$  denotes the Kronecker product,

$$\mathbf{s}_g(n) = \begin{bmatrix} s_{g,K_0}(n) & s_{g,K_1}(n) & \dots & s_{g,K_{M-1}}(n) \\ \vdots & \ddots & & \vdots \\ s_{g,K_0}(n) & s_{g,K_1}(n) & \dots & s_{g,K_{M-1}}(n) \end{bmatrix}, \quad (11)$$

and

$$\boldsymbol{\Omega}_{ML} = \begin{bmatrix} 1 & e^{-\frac{(2\pi)}{K}K_0} & \dots & e^{-\frac{(2\pi)}{K}K_0(L-1)} \\ \vdots & \ddots & & \vdots \\ 1 & e^{-\frac{(2\pi)}{K}K_{M-1}} & \dots & e^{-\frac{(2\pi)}{K}K_{M-1}(L-1)} \end{bmatrix}. \quad (12)$$

The signal to noise ratio (SNR) for the MIMO-OFDM system is defined as

$$SNR = \left\| \sum_{g=1}^{N_G} \mathbf{h}_g(n) \mathbf{s}_g(n) \right\| / \sigma^2. \quad (13)$$

### 3 Pilot Insertion Pattern

The pattern of the inserted signals has a significant influence on the solution of Eq. (8). Under the framework of CS, if the estimated  $\mathbf{h}(n)$  is  $s$ -sparse and the matrix  $\mathbf{A}(n)$  meets the  $s$  order condition of the restricted isometric property (RIP),

$$(1 - \delta_s) \|\mathbf{h}(n)\|_2^2 \leq \|\mathbf{A}_s(n)\mathbf{h}(n)\|_2^2 \leq (1 + \delta_s) \|\mathbf{h}(n)\|_2^2, \quad (14)$$

where  $\delta_s$  is a constant related to sparsity  $s$ ,  $0 \leq \delta_s < 1$ ;  $\mathbf{A}_s(n)$  is arbitrary  $s$  columns of  $\mathbf{A}(n)$ .

We can recover  $\mathbf{h}(n)$  by

$$\min \|\mathbf{h}(n)\|_1 \quad s.t. \quad \mathbf{y}(n) = \mathbf{A}(n)\mathbf{h}(n). \quad (15)$$

#### 3.1 Design Principle of Sensing Matrix

The validation of a sensing matrix that satisfies RIP is very complex. It is impractical to apply (14) for the pilot optimization. Derived from the condition of RIP, it is known that for the accurate recovery the projection of  $\mathbf{h}(n)$  on any column of  $\mathbf{A}(n)$  should be equal and the columns of  $\mathbf{A}(n)$  should be the least uncorrelated. The mutual irrelevant property (MIP) usually exists in RIP, but MIP is simpler and more suitable in practically measure the quality of the sensing matrix at present. The mutual correlation value for  $\mathbf{A}$  is defined as

$$\mu(\mathbf{A}) = \max_{\substack{1 \leq C_i, C_j \leq N \\ C_i \neq C_j}} \frac{|\langle \mathbf{A}_{C_i}^H, \mathbf{A}_{C_j} \rangle|}{\|\mathbf{A}_{C_i}\|_2 \cdot \|\mathbf{A}_{C_j}\|_2}, \quad (16)$$

where  $\mathbf{A}_{C_i}$  and  $\mathbf{A}_{C_j}$  represent the  $C_i$  and  $C_j$  columns of  $\mathbf{A}$ .

It can be seen that the minimization of (16) is also difficult since the number of pilots tends large. The position and the value of the pilots affect the magnitude of the mutual correlation value. To follow the principle that the pilot energies on pilot tones are equal, we define the average power of pilots as

$$E_{av} = \frac{1}{N_G} \sum_{g=1}^{N_G} \sum_{m=0}^{M-1} |s_{g,k_m}|^2. \quad (17)$$

Hence, the MIP condition is to minimize

$$\min \left\{ \sum_{m=0}^{M-1} s_{g,K_m(C_i)}(n) s_{g',K_m(C_j)}(n) e^{-\left(\frac{2\pi}{K}\right) K_m [C_i(l) - C_j(l)]} \right\} \quad s.t. \quad \sum_{m=0}^{M-1} |s_{g,k_m}|^2 = \text{constant}, \forall g, g', g \neq g'. \quad (18)$$

#### 3.2 Phase-shift Orthogonal Pilot

By design the multiple antennas share the same pilot tones. If the pilot tones are sufficient, i.e.,  $M \geq 2N_G L$ , at one time instant, i.e.  $n = 0$ , the column correlation can be calculated by

$$\mathbf{A}^H(0)\mathbf{A}(0) = \begin{bmatrix} \mathbf{A}_1^H(0)\mathbf{A}_1(0) & \cdots & \mathbf{A}_1^H(0)\mathbf{A}_{N_G}(0) \\ \vdots & \ddots & \vdots \\ \mathbf{A}_{N_G}^H(0)\mathbf{A}_1(0) & \cdots & \mathbf{A}_{N_G}^H(0)\mathbf{A}_{N_G}(0) \end{bmatrix}. \quad (19)$$

Every sub-block is  $\mathbf{A}_g^H(0)\mathbf{A}_{g'}(0)$  with the dimension of  $K_M \times L$ ,

$$\mathbf{A}_g^H(0)\mathbf{A}_{g'}(0) = \mathbf{\Omega}_{ML}^H \begin{bmatrix} s_{gK_0}^H(0) s_{g'K_0}(0) & 0 & 0 \\ 0 & \ddots & \\ 0 & 0 & s_{gK_{M-1}}^H(0) s_{g'K_{M-1}}(0) \end{bmatrix} \mathbf{\Omega}_{ML}. \quad (20)$$

In real situations, to make the sensing matrix to be the least correlated and even not uncorrelated is very important. Directly using random pilot may lead to correlated columns which deteriorate the later reconstruct procedure greatly. It is noted that there are two types of optimal pilot: diagonal pilot and phase-shift orthogonal pilot. The diagonal pilot is a scheme in which the pilot tones of each antenna are orthogonal in the frequency domain. The pilot symbols on one antenna are transmitted at a certain pilot tone. Meanwhile, other antennas at the subcarrier are kept silent to avoid interference. This scheme is similar to the pilot scheme of single input single output OFDM. It avoids the matrix inversion operations and reduces the complexity of channel estimations at the receiver end. This scheme, however, needs to estimate the time-domain FIR on multiple antennas one by one, and the pilot overhead is high.

Phase-shift orthogonal sequences can be distributed during the whole transmission period. The pilot sequence on each transmit antenna has a fixed phase deviation, while there is phase difference at the pilot tones at different impulse response taps. The latter scheme performs better for the estimation of time-varying channels and thus more applicable in the cases of large moving speed scenarios.

The phase-shift orthogonal pilot is chosen to fulfill (18). And the pilot tones are fixed with equal intervals. Then the allocation of the pilot on antenna  $g$  with pilot tone  $Kp$  is as

$$s_{Kp}^g = E_{av} / M e^{-j2\pi(g-1)LK_p/K_M}, \quad (21)$$

where  $g = (1, \dots, N_G)$ ;  $M$  is the number of pilot tones on each transmit antenna.

In this way the whole subcarriers are occupied by the comb pilots and traffic symbols. At the same time instant, each transmit antenna embeds pilots on the same orthogonal carriers. The pilot tones should be set strong such that it ensures that the sensing matrix has sufficient measurements by extracting the uncorrelated columns constructed by the phase-shift pilots and the partial DFT matrix. The pilot pattern is already shown in Fig. 1. Pilots immediately follow the OSTBC symbols. The frequency interval of the pilot tones is less than  $1/\tau_{\max}\Delta f$ , where  $\tau_{\max}$  is the maximum multipath delay and  $\Delta f$  is the sub-channel interval.

## 4 Robust SAMP Algorithm with MMV

After the modeling of the sparse structure of the massive MIMO-OFDM channel and the designing of the pilot pattern in the transmission frame, Eq. (15) needs to be solved using the recovery algorithms to find the optima. The evaluation criteria for a recovery algorithm generally include the reconstruction accuracy and the computational complexity.

It is known that the number of measurements on pilot tones should be at least equal to two times of sparsity. When there are multiple measurement vectors (MMV) and the unknown signal has a common support set, the performance of the reconstruction algorithm can be improved by using the joint spatial-temporal feature of the signal.

### 4.1 Multiple Measurement Vector Model

One special form of the block sparse models is the MMV model. In CS, if the unknown sparse signal is reconstructed from a single measurement vector (SMV), the measurement model is called SMV model. From this SMV model, the signal  $\mathbf{h}$  is able to be recovered when the measurement  $\mathbf{y}$  defined by the linear measurement matrix  $\mathbf{A}$  is known. The problem this model has, however, is usually NP hard. In ideal cases that the sparsity of a signal  $\mathbf{h}$  is known and the measurement is noiseless, there exists several reconstruction algorithms that ensure the accurate reconstruction of the signal  $\mathbf{h}$ . Practically when the  $\mathbf{h}$ 's sparsity is unknown even though it has a good sparse approximation, or if the linear measurement  $\mathbf{y}$  is polluted by noise, the reconstruction errors will be large. But under the noise-polluted measurement conditions, the MMV model is still able to function.

Suppose that the multiple measurement vectors be  $\mathbf{Y} = [\mathbf{y}(1), \dots, \mathbf{y}(n)]^T$ . Subsequently, a channel is described as  $\mathbf{H} = [\mathbf{h}(1), \dots, \mathbf{h}(n)]^T$ , where  $\mathbf{h}(1) = \dots = \mathbf{h}(n)$ . And the sensing matrix is  $\mathbf{A} = [\mathbf{A}(1), \dots, \mathbf{A}(n)]^T$ . The standard MMV problem can be expressed as the following optimization expression

$$\min \|\mathbf{H}\| \quad s.t. \quad \mathbf{Y} = \mathbf{A}\mathbf{H}. \quad (22)$$

## 4.2 Comparison of Recovery Algorithms

Typical reconstruction algorithms include two types of algorithms: greedy matching pursuit, and base pursuit (BP) that is aiming at  $l_1$ -norm minimization. These algorithms primarily focus on the effective implementation and the high-precision expansion. Within the type of greedy matching pursuit, the OMP algorithm is common when the sparsity is known. For the reconstruction of the  $s$ -sparsity signal, the OMP algorithm implements the  $s$  times of iterations, and during an iteration an atom is selected. Moreover, the regularized OMP (ROMP) algorithm firstly selects  $s$  candidates with the largest absolute value of inner product, and then decides the finalist from the  $s$  candidates according to its regularization principle. The subspace pursuit (SP) algorithm has lower computational complexity than the OMP algorithm. The stagewise orthogonal matching pursuit (StOMP) algorithm simplifies the OMP algorithm in order to improve the speed of calculations at the cost of the approximation accuracy. The CoSaMP algorithm introduces a backtracking to improve the approximation accuracy also with the reduction of the computational complexity. It can be observed that the algorithms with the backtracking need to select the finalist atoms contained in the candidate set stage by stage.

On the other hand, within the type of  $l_1$ -norm minimization, the algorithm of LASSO performs the least square (LS) estimation to minimize the MSE and constrain the  $l_1$ -norm of reconstruction vector to a certain upper bound. Moreover, the least angle regression (LARS) can be an efficient method to solve the LASSO problems. The Bregman iteration is another choice to find the optima under the  $l_1$ -norm minimization and Lagrange constraint [25].

Although the former type of algorithms seems irrelevant to the optimization of the problem of Eq. (15), this type actually can approach close to the solution in most cases. It calculates the LS solution by using the candidate sets formed by partial columns of the sensing matrix. When the backtracking iterative mechanism is accompanied with the selection of the finalist, the LS calculation is doubled to ensure that the correct atoms are added. Its computational cost is relatively high. In contrast to the greedy algorithms, the type of  $l_1$ -norm minimization algorithms does not need any prior knowledge of the sparsity. This type has less computational efforts but the convergence speed may be slow. The recovery accuracy of the latter can be worse than the former one, as the LS estimation is of the optimum solution in the scenario that measurement is polluted by noise.

## 4.3 Oracle LS

The common way of the reconstruction of sparse channels is to find out the positions of all non-zero elements in the sparse vector  $\mathbf{h}$  iteratively, and then give

$$\hat{\mathbf{h}} = \arg \min \|\mathbf{y} - \mathbf{A}\mathbf{h}\|_2. \quad (23)$$

Oracle LS, which can be referred to as the Gaussian elimination with a partial pivoting method, is to find the specific solution that has the most zero elements. And the null space of  $\mathbf{A}$  can be further used to find the general solution of the underdetermined equation. It is eventually to select the element that is of the largest absolute value from the elements in column  $C_k$ , and then exchange the position of this element with that of the principal element so as to carry out the desired elimination. In the process of this method, the unknown variables are eliminated in sequence. By selecting the principal component carefully, the upper bound of the acquisition accuracy can be achieved.

## 4.4 OMP Algorithm

OMP algorithm is specifically used to solve the problem of the general matching pursuit algorithm, i.e., atoms cannot be selected accurately. The solution is that all the selected atoms are kept orthogonalized at each step of the decomposition, which speeds the convergence. The central idea is to greedily select the columns of the sensing matrix. Through the orthogonal processing and the use of the selection principle of the special atoms, this algorithm finally obtains the expected atoms that best match either the redundant vector or the compressed measurement value. Afterwards, it starts to project the signal to the subspace of the orthogonal atoms. OMP algorithm only selects one atom at a time to expand the support set during each iteration, which will lead to multiple iterations in the process. Therefore, when the sparsity is at large, the reconstruction process becomes very time-consuming.

When the requirements for accuracy are the same, the algorithm is able to ensure the performance in its iterative process. In addition, OMP has the advantages of simple structure and low complexity. Its calculating steps with SMV are shown in Table 1. When the MMV model is adopted, the needed atoms of the sensing matrix can be selected using the means of continuously multiplying the residuals. The algorithm with MMV is shown in Table 2.

**Table 1.** OMP algorithm with SMV

Input: measurement matrix $A$ , measumrents $S$ , sparsity $S$
Output: sparse vector $h$
Initialization:
sparse vector $h = 0$ ; residue $y_0 = y$ ; support set $E_0 = \phi$ ; iteration index $k = 1$
Repeat
(1) fine the atom $S_k = \max  A_t, r_{k-1} $
(2) add the atoms into the support set $E_k = E_{k-1} \cup S_k$
(3) calculate the LS solution $\hat{h}_t = (A_t A_t^H)^{-1} A_t^H y$
(4) calculate the residue $r = y - A_E (A_E A_E^H)^{-1} A_E^H y$
(5) if $k = s$ , break; else $k = k + 1$

**Table 2.** OMP algorithm with MMV

Input: measurement matrix $A$ , measurement martis $Y$ , sparsity $S$
Output: sparse vector $h$
Initialization:
sparse vector $h = 0$ ; residue $r_0 = Y$ ; support set $E_0 = \phi$ ; iteration index $k = 1$
Repeat
(1) fine the atom $S_k = \max \prod_{j=1}^n  A_t, r_{k-1} $
(2) add the atoms into the support set $E_k = E_{k-1} \cup S_k$
(3) calculate the LS solution $\hat{h}_t = (A_t A_t^H)^{-1} A_t^H y$
(4) calculate the residue $r = Y - A_E (A_E A_E^H)^{-1} A_E^H Y$
(5) if $k = s$ , break; else $k = k + 1$

#### 4.5 SP Algorithm

In SP algorithm, more than one atom are allowed to be selected in iterations to update the support set. The size of the subspace can be flexibly adjusted in order to improve the reconstruction speed to a certain controlled extent. It can be seen from the calculation flow shown in Table 3 and Table 4: the algorithm first multiplies the conjugate transposition of the sensing matrix by the observation vector  $y$  to obtain a proxy of the signal, and then finds the largest  $S$  number of atoms in the support set.

#### 4.6 Robust Sparsity Adaptive Matching Pursuit Algorithm

SAMP algorithm is a greedy one designed for the recovery of the blind signals when the prior knowledge of the sparsity is unavailable [15]. It employs a backtracking iterative mechanism to add the support set with a certain size. The halting condition of the general SAMP relies on the residual's norm that is smaller than a certain threshold. Since the sensing matrix is constructed by partial DFT matrix as well as the measurements are polluted by the noise, it is difficult to properly predetermine the threshold. An extremely small threshold is an obstacle for the greedy algorithm to achieve good performance as the overestimation may occur and resulting in an excessive approximation. On the contrary, a large threshold will lead to underestimate such that the estimation would suffer from less calculation and thus deteriorate the accuracy.

**Table 3.** SP algorithm with SMV

---

Input: measurement matrix  $\mathbf{A}$ , measurements  $\mathbf{y}$ , sparsity  $S$

---

Output: sparse vector  $\mathbf{h}$

---

Initialization:  
 sparse vector  $\mathbf{h} = 0$ ; iteration index  $k = 1$ ; select  $S$  number of atoms according to  $S_0 = \max\{\|\mathbf{A}^H \mathbf{y}\|\}$ ;  $E_0 = S_0$ ; calculate the resider  $\mathbf{r} = \mathbf{y} - \mathbf{A}_E (\mathbf{A}_E \mathbf{A}_E^H)^{-1} \mathbf{A}_E^H \mathbf{y}$

Repeat

- (1) select  $S$  number of atoms according to  $S_k = \max\{\|\mathbf{A}^H, \mathbf{r}_{k-1}\|\}$  add the atoms into the support set  $C_k = E_{k-1} \cup S_k$
- (2) calculate the LS solution  $\mathbf{h}_t = (\mathbf{A}_{C_k} \mathbf{A}_{C_k}^H)^{-1} \mathbf{A}_{C_k}^H \mathbf{y}$
- (3) select  $S$  number of the largest value to reconstruct  $C'_k$
- (4) calculate  $\mathbf{r}_{new} = \mathbf{y} - \mathbf{A}_{C'_k} (\mathbf{A}_{C'_k} \mathbf{A}_{C'_k}^H)^{-1} \mathbf{A}_{C'_k}^H \mathbf{y}$
- (5) if  $\|\mathbf{r}_{new}\|_2 > \|\mathbf{r}\|_2$ ,  $E_k = E_{k-1}$ , go to (1)
- (6) output  $\hat{\mathbf{h}}_t = (\mathbf{A}_E \mathbf{A}_E^H)^{-1} \mathbf{A}_E^H \mathbf{y}$

---

**Table 4.** SP algorithm with MMV

---

Input: measurement matrix  $\mathbf{A}$ , measurements  $\mathbf{Y}$ , sparsity  $S$

---

Output: sparse vector  $\mathbf{h}$

---

Initialization:  
 sparse vector  $\mathbf{h} = 0$ ; iteration index  $k = 1$ ; select  $S$  number of atoms according to  $S_0 = \max \prod_{j=1}^n \{\|\mathbf{A}^H \mathbf{Y}\|\}$ ;  $E_0 = S_0$ ; calculate the resider  $\mathbf{R} = \mathbf{Y} - \mathbf{A}_E (\mathbf{A}_E \mathbf{A}_E^H)^{-1} \mathbf{A}_E^H \mathbf{Y}$

Repeat

- (1) select  $S$  number of atoms according to  $S_k = \max \prod_{j=1}^n \{\|\mathbf{A}^H \mathbf{Y}\|\}$  add the atoms into the support set  $C_k = E_{k-1} \cup S_k$
- (2) calculate the LS solution  $\mathbf{h} = (\mathbf{A}_{C_k} \mathbf{A}_{C_k}^H)^{-1} \mathbf{A}_{C_k}^H \mathbf{Y}$
- (3) select  $S$  number of the largest value to reconstruct  $C'_k$
- (4) calculate  $\mathbf{R}_{new} = \mathbf{Y} - \mathbf{A}_{C'_k} (\mathbf{A}_{C'_k} \mathbf{A}_{C'_k}^H)^{-1} \mathbf{A}_{C'_k}^H \mathbf{Y}$
- (5) if  $\|\mathbf{R}_{new}\|_2 > \|\mathbf{R}\|_2$ ,  $E_k = E_{k-1}$ , go to (1)
- (6) output  $\hat{\mathbf{h}}_t = (\mathbf{A}_E \mathbf{A}_E^H)^{-1} \mathbf{A}_E^H \mathbf{y}$

---

A basic SAMP algorithm selects a certain length of atoms from the columns of the sensing matrix using the criterion of the maximum inner product and then calculates the estimates

$$\hat{\mathbf{h}}_{t_k} = [\mathbf{A}_{t_k}^H \mathbf{A}_{t_k}]^{-1} \mathbf{A}_{t_k}^H \mathbf{r}_{k-1}. \quad (24)$$

The corresponding residual is

$$\mathbf{r}_k = \mathbf{y} - \mathbf{A}_{t_k} \hat{\mathbf{h}}_t = \mathbf{y} - \mathbf{A}_{t_k} [\mathbf{A}_{t_k}^H \mathbf{A}_{t_k}]^{-1} \mathbf{A}_{t_k}^H \mathbf{r}_{k-1}, \quad (25)$$

where  $\mathbf{A}_{t_k} [\mathbf{A}_{t_k}^H \mathbf{A}_{t_k}]^{-1} \mathbf{A}_{t_k}^H \mathbf{r}_{k-1}$  is the orthogonal projection of  $\mathbf{r}$  on the column space of  $\mathbf{A}$ .

Since the matching pursuit algorithm projects the residual in a stage by stage manner, and the residual gradually decreases from its initial value to a relatively small and smaller value over the stages, the residual variation is still too large to determine a suitable threshold for the termination of the algorithm. Moreover, the backtracking mechanism validates the atoms from the candidate set by comparing two adjacent residuals. However, this comparison leads to a point that the residuals do not always decrease in some adjacent iterative steps when the measurement is polluted by noise. The fluctuation of residuals disturbs the setting of a correct threshold and thus the evolution direction. It can be observed from these residuals that the values of these residuals sometimes keep constant when no new candidates are suitable

to be added to the finalist. Thus the smoothing of just two or three residuals are not enough, and it is necessary to design a relatively large smooth length, i.e. 5~10.

Intended to approximate the sparsity adaptively with a relatively large range of step size, an improved robust algorithm was presented suitable for the recovery when the measurement is polluted by noise. By the use of the greedy strategy of the SAMP, it can be ensured that all the residuals move towards the reduction trend direction after the establishment of the size of initial support set. The convergence speed of the residuals starts to become slow when the finalist set is close to the true sparsity. Since the noise is related to SNR and contained within the residuals, the threshold is thus set as  $\sigma^2$ . The halting condition of the algorithm is then like

$$\frac{1}{\beta \cdot \zeta} \sum_{w=0}^{\zeta-1} (r_{\zeta-w} - r_{\zeta-w-1}) \leq \sigma^2, \quad (26)$$

where  $\beta$  is the initial size of a support set,  $\zeta$  is the smoothing length.

The whole algorithm with SMV is shown in Table 5, and the corresponding MMV algorithm is shown in Table 6.

**Table 5.** Robust SAMP algorithm with SMV

---

Input: measurement matrix  $\mathbf{A}$ , measurements  $\mathbf{y}$ , size of the candidate set in the first stage

---

Output: sparse vector  $\mathbf{h}$

---

Initialization:

sparse vector  $\mathbf{h} = \mathbf{0}$ ; residue  $\mathbf{r}_0 = \mathbf{y}$ ; support set  $E_0 = \emptyset$ ; size of support set  $\delta = \beta$ ;  
iteration index  $k = 1$ ; stage index  $w = 1$ ; threshold =  $\sigma^2$ ; residue number  $\xi$

Repeat

- (1) select  $L$  number of atoms according to  $S_k = \max\{\|\mathbf{A}\mathbf{r}_{k-1}\|, \delta\}$  add the atoms
- (2) add the atoms into the support set  $C_k = E_{k-1} \cup S_k$
- (3) select  $L$  number of atoms according to  $E = \max\{(\mathbf{A}_{C_k} \mathbf{A}_{C_k}^H)^{-1} \mathbf{A}_{C_k}^H \mathbf{y}, \delta\}$
- (4) calculate the residue  $\mathbf{r} = \mathbf{y} - \mathbf{A}_E (\mathbf{A}_E \mathbf{A}_E^H)^{-1} \mathbf{A}_E^H \mathbf{y}$
- (5) if  $\frac{1}{\beta \cdot \xi} \sum_{w=0}^{\xi-1} \|\mathbf{r}_{\xi-w} - \mathbf{r}_{\xi-w-1}\|_2 \geq \text{threshold}$ ,  $E_k = E$ , break  
else if  $\|\mathbf{r}\|_2 \geq \|\mathbf{r}_{k-1}\|_2$ , update  $w = w + 1$ ,  $\delta = w \times \beta$ , go to (1)  
else  $E_k = E$ ,  $\mathbf{r}_k = \mathbf{r}$ ,  $k = k + 1$ , go to (1)
- (6) output  $\hat{\mathbf{h}}_k = (\mathbf{A}_E \mathbf{A}_E^H)^{-1} \mathbf{A}_E^H \mathbf{y}$

---

**Table 6.** Robust SAMP algorithm with MMV

---

Input: measurement matrix  $\mathbf{A}$ , measurements  $\mathbf{Y}$ , size of the candidate set in the first stage

---

Output: sparse vector  $\mathbf{h}$

---

Initialization:

sparse vector  $\mathbf{h} = \mathbf{0}$ ; residue  $\mathbf{R}_0 = \mathbf{Y}$ ; support set  $E_0 = \emptyset$ ; size of support set  $\delta = \beta$ ;  
iteration index  $k = 1$ ; stage index  $w = 1$ ; threshold =  $\sigma^2$ ; residue number  $\xi$

Repeat

- (1) select  $L$  number of atoms according to  $S_k = \max \prod_{j=1}^n \{\|\mathbf{A}\mathbf{R}_{k-1}\|, \delta\}$  add the atoms
- (2) add the atoms into the support set  $C_k = E_{k-1} \cup S_k$
- (3) select  $L$  number of atoms according to  $E = \max\{(\mathbf{A}_{C_k} \mathbf{A}_{C_k}^H)^{-1} \mathbf{A}_{C_k}^H \mathbf{Y}, \delta\}$
- (4) calculate the residue  $\mathbf{R} = \mathbf{y} - \mathbf{A}_E (\mathbf{A}_E \mathbf{A}_E^H)^{-1} \mathbf{A}_E^H \mathbf{Y}$
- (5) if  $\frac{1}{\beta \cdot \xi} \sum_{w=0}^{\xi-1} \|\mathbf{R}_{\xi-w} - \mathbf{R}_{\xi-w-1}\|_2 \geq \text{threshold}$ ,  $E_k = E$ , break  
else if  $\|\mathbf{R}\|_2 \geq \|\mathbf{R}_{k-1}\|_2$ , update  $w = w + 1$ ,  $\delta = w \times \beta$ , go to (1)  
else  $E_k = E$ ,  $\mathbf{R}_k = \mathbf{R}$ ,  $k = k + 1$ , go to (1)
- (6) output  $\hat{\mathbf{h}}_k = (\mathbf{A}_E \mathbf{A}_E^H)^{-1} \mathbf{A}_E^H \mathbf{Y}$

---

To evaluate its recovery performance, we define normalized MSE (NMSE) as

$$\text{NMSE} = \frac{1}{N_s} \sum_{w=1}^{N_s} \frac{\|\hat{\mathbf{h}} - \mathbf{h}\|_2^2}{\|\mathbf{h}\|_2^2}, \quad (27)$$

where  $N_s$  is the number of Monte Carlo simulation.

It is worth to note that the MMSE estimation is to evaluate the performance of algorithms as lower bound, represented by

$$\hat{\mathbf{h}} = \mathbf{R}_h \mathbf{A}^H(n) [\mathbf{A}(n) \mathbf{R}_h \mathbf{A}(n) + \sigma^2 \mathbf{I}]^{-1} \mathbf{Y}(n), \quad (28)$$

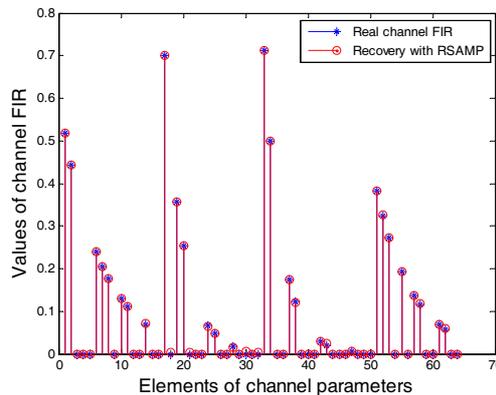
where  $\mathbf{R}_h$  is the correlation matrix of channels.

## 5 Simulations

During the setup of the simulations, a massive MIMO-OFDM is considered firstly with multiple transmit antennas and 4096 subcarriers. Then these parameters are pre-defined and set up respectively. The bandwidth is set to 1GHz and the tolerable time delay is assumed as 200ns. The guard interval is 800ns occupied by CP. The symbol period is 4.8 $\mu$ s, and 4 $\mu$ s is for the division of sub-channels. The subcarrier interval is 250kHz. The modulation is 16-QAM, and the OSTBC code rate is 1/2. The time domain channel parameters refer to the Cost-207 multipath channel, and the exponential power delay spectrum is utilized. The response of each tap is calculated by  $h(l) = p \exp(-u \cdot l)$ , where the power profile of the first tap  $p$  is set to random within the range of 1 to 100; the spectral fading factors  $u$  are uniform distributed; and  $l$  here denotes the FIR tap. Every sub-channel response is normalized to unitary energy. In the  $L$  taps of FIR, the sparsity is arbitrary. Comb phase-shifting pilots are allocated over equal subcarrier intervals. Monte Carlo simulation is applied in each case.

### 5.1 Simulation Results

Fig. 2 shows the recovery results compared between the real channel of 16 taps of FIR with 4 transmit antennas and the RSAMP algorithm with MMV. The step size is 1 and SNR is set to 6dB. The number of measurement vectors is 20. It is clearly seen that the recovery of the proposed algorithm is performing well, although it has small values on some zero taps. It is also known that the estimated sparsity is larger than that of the real channel. These small responses, however, contribute much less to the NMSE.



**Fig. 2.** Channel FIR and Recovery

Fig. 3 shows the NMSE performances of OMP, SP and RSAMP with phase-shift orthogonal pilot denoted by MMV-P as well as random pilot denoted by MMV-R through the corresponding MMV algorithms. It can be observed that the NMSE gets smaller when the pilot is phase-shift orthogonal. It can also be observed that when the number of measurement vectors increases, the NMSE tends to become smaller.

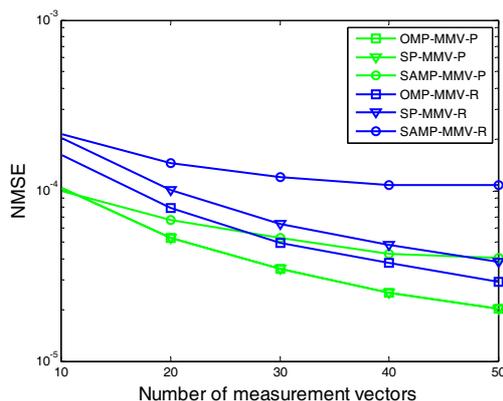


Fig. 3. NMSE of recovery algorithm using MMV

Fig. 4 shows the NMSE performances of OMP, SP and RSAMP algorithms versus SNR when the number of transmit antennas is 16 and the length of sub-carrier FIR is 16. The pilot is phase-shift orthogonal. The number of measurement vectors in MMV is also set to 20. It can be seen that the upper bound of the estimations is oracle LS, while the lower bound is MMSE. In the SMV scenario, the NMSE performance of the RSAMP algorithm is the best over OMP and SP algorithms. Moreover, firstly the NMSE performance of the proposed algorithm is comparable with those of OMP and SP algorithms in the MMV scenario at low SNR values; and secondly the proposed algorithm achieves the accuracy of OMP and SP but with higher SNR. In the MMV scenario, the NMSE performances of the three algorithms are much lower than those in the SMV scenario. The performances of these algorithms cannot approach close to the lower bound of MMSE. The initial step size is not much accounted for NMSE in terms of SNR. However, its calculation efficiency is much higher when the initial candidate set is set to 4 or 8. Fig. 4 demonstrates that the NMSE of RSAMP approaches to other algorithms in the greedy frame, which indicates that the accurate recovery can be accomplished with limited computational cost when the sparsity is unknown. Fig. 5 shows the NMSE performances of these algorithms with unit-energy random pilot. It can be observed that the NMSE performances of these algorithms are worse than the ones with the phase-shift orthogonal pilot.

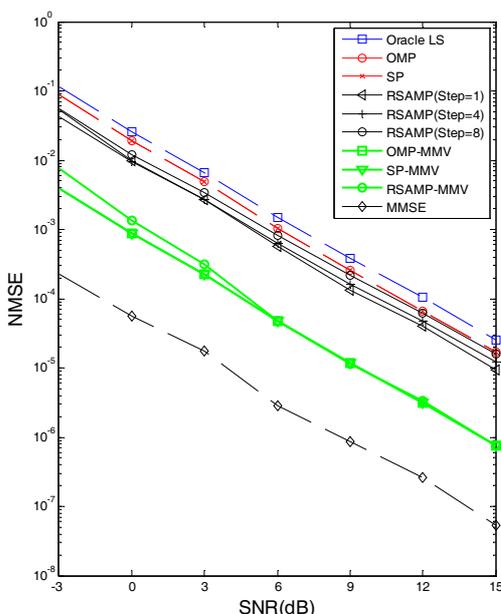


Fig. 4. NMSE versus SNR with phase-shift pilot

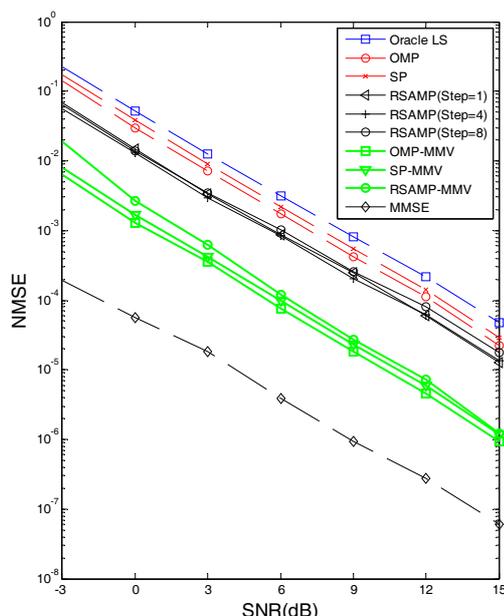
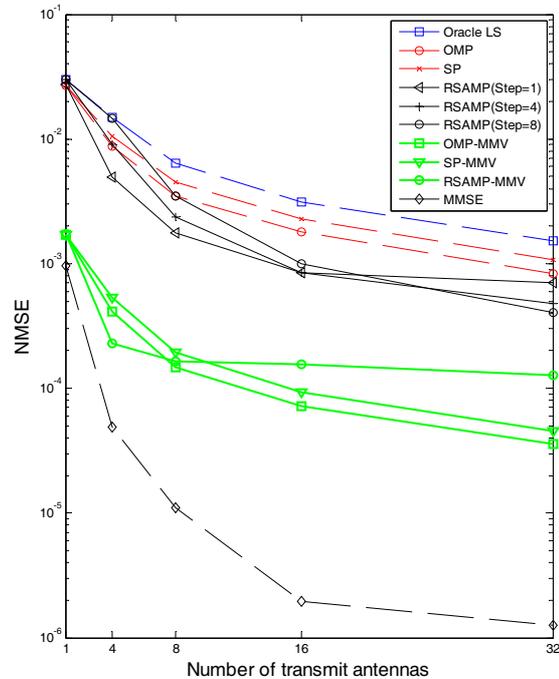
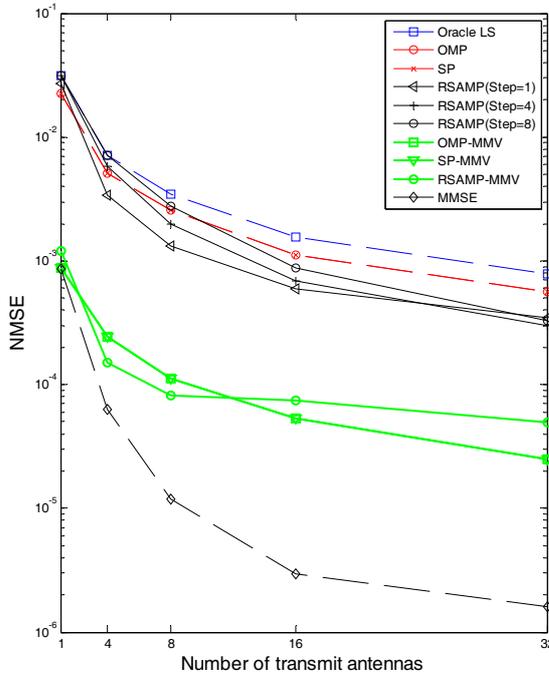


Fig. 5. NMSE versus SNR with random pilot

Fig. 6 presents the NMSE curves versus the number of transmit antennas when the SNR is set to 6dB and the length of subcarrier FIR is 16. The number of MMV is still set to 20. It is known that the twice measurements of the channel dimension is sufficient for the recoveries by OMP, SP and RSAMP. More

performed measurements can help improve the reconstruction precision. The slope of the NMSE reduction, however, is getting small when the number of transmit antennas is getting larger. The results suggest that much less pilot overhead can be used to achieve the same accurate estimation. The calculation time used by RSAMP is the longest among those by OMP and SP. This is due to the fact that RSAMP makes the candidate set adaptive using the backtracking mechanism along with the LS inversion calculation. Fig. 7 shows the NMSE performances with random pilot, and illustrates that the NMSE performances of these algorithms are worse than those using phase-shift orthogonal pilot.



**Fig. 6.** NMSE versus antennas with phase-shift pilot **Fig. 7.** NMSE versus antennas with random pilot

## 5.2 Analysis and Discussion

The simulations reported in the Subsection 5.1 show that the recovery of sparse channels depends largely on the design of pilot signals. It has been observed that when the pilot subcarriers are shared in a MIMO-OFDM system, the introduction of a phase-shifting orthogonal pilot can drive the sensing matrix towards a fixed matrix with the minimum column correlation. And the estimation performance is only affected by the observation noise. Therefore, the support vectors are always set as orthogonal so as to ensure the numerical stability by using an LS method. It is difficult, however, to control the correlation of columns by using just random pilots, especially under the situations of using the accelerated algorithms in which selecting multiple candidate atoms to join the support set occurs during one iteration. Even though the scheme of using random pilot can be optimized, it still cannot guarantee the minimum correlation between each column.

It is also found out from the simulations that the observation noise should be taken into account in the adaptive iterative algorithm. In a noiseless model with a high SNR application, the recovery trend is generally approaching the real sparsity as well as getting the small MSE. Nevertheless, in a transmission system with low SNR, it is always difficult to distinguish whether the residuals are caused by inaccurate sparsity or by the observation noise. With either a fixed step size or a variable step size that is used in the iterative scheme, the resultant residual determines whether adding candidate atoms to the support set is needed or the iteration should continue. It happens sometimes that the candidate set fails to meet the residual requirements. It has also been observed that when only the reduction of the residuals of the two adjacent iterations is used as the condition to determine whether the search should stop or not, the general SAMP algorithm will not always stop searching atoms until all the columns of the measurement matrix are searched. Therefore, in order to reduce the interference by the observation noise, the algorithm proposed in this paper smooths the residuals, and stops once the smoothed residuals are smaller than the

SNR of the receiver. This smoothening approach with pre-set stop mechanism effectively avoids the over estimation of sparsity. Since this proposed algorithm uses the SNR as the threshold during each iterative updating step, it demonstrates superior adaptability by being able to stop the iteration near the real sparsity. The algorithm has been proven to be of good robustness under different observation conditions. Moreover, the phase-shift orthogonal pilot makes the atoms uncorrelated, and thus the algorithm can get better NMSE with high reliability.

In majority cases, to achieve the best estimation performance by the accurate approximation of sparsity is required. But it is difficult to estimate the real sparsity with the needed accuracy when the inevitable observation noise exists. Although it is difficult for this method to obtain accurate sparsity values, this method does have little influence on the NMSE because the values of many parameters to be estimated are quite small and thus the contributions to the error are negligible.

## 6 Conclusions

The estimation issue of fading parameters of multipath channels exists in massive MIMO-OFDM systems. The recovery of these fading parameters is able to be achieved by a proposed robust sparsity adaptive matching pursuit (RSAMP) algorithm in both single and multiple measurement vector scenarios with unknown sparsity. The halting condition is designed and built in the algorithm and works once the smoothing residual is smaller than the noise variance. The pilot pattern is selected to be phase-shift orthogonal comb pilot to ensure that the columns of the sensing matrix are mutually uncorrelated. Extensive simulations and comparisons have been conducted under various scenarios, including the normalized mean square error versus SNR and the number of transmit antennas. The analysis of these results demonstrates that, performance wise, the proposed RSAMP algorithm is comparable with OMP and SP algorithms.

Specifically in this paper, the recovery of sparse channels in a massive MIMO-OFDM system is described from two aspects - pilot design and reconstruction method. The relatively smaller MSE performance is achieved via a robust algorithm design. Although this method is difficult to obtain accurate sparsity values, it has little influence on the normalized mean square error because the estimated values of many parameters are quite small. Move forward, the next research step is about of the real sparsity approximation via a well designed adaptive algorithm in the presence of observation noise. And the future research focus will be in these areas - sparse filtering in multiple measurement scenarios for joint estimation of channel fading, time delay and frequency offset.

## Acknowledgements

This work is supported by the Foundation of Shanghai Normal University (KF202820).

## References

- [1] E. Hossain, M. Hasan, 5G cellular: key enabling technologies and research challenges, *IEEE Instrumentation & Measurement Magazine* 18(3)(2015) 11-21.
- [2] E. Olfat, M. Bengtsson, A general framework for joint estimation-detection of channel, nonlinearity parameters and symbols for OFDM in IoT-based 5G networks, *Signal Processing* 167(2020) 1-11.
- [3] F. Rosário, F.A. Monteiro, A. Rodrigues, Fast matrix inversion updates for massive MIMO detection and precoding, *IEEE Signal Processing Letters* 23(1)(2016) 75-79.
- [4] E. Nayebi, B.D. Rao, Semi-blind channel estimation for multiuser massive MIMO systems, *IEEE Trans. Signal Processing* 66(2)(2018) 540-553.
- [5] X. He, R. Song, W. Zhu, Pilot allocation for distributed-compressed-sensing-based sparse channel estimation in MIMO-OFDM systems, *IEEE Trans. Vehicular Technology* 65(5)(2016) 2990-3003.

- [6] R. Mohammadian, A. Amini, B. Hossein Khalaj, Deterministic pilot design for sparse channel estimation in MISO/multi-user OFDM systems, *IEEE Trans. Wireless Communications*, 16(1)(2017) 129-140.
- [7] S. Somashekar, N.K.D. Venkategowda, A.K. Jagannatham, Bandwidth efficient optimal superimposed pilot design for channel estimation in OSTBC based MIMO-OFDM systems, *Physical Communication* 26(2018) 185-195.
- [8] Y. Nan, X. Sun, L. Zhang, Joint channel estimation algorithm via weighted homotopy for massive MIMO OFDM system, *Digital Signal Processing* 50(2016) 34-42.
- [9] R. Jeya, B. Amutha, Optimized semiblind sparse channel estimation algorithm for MU-MIMO OFDM system, *Computer Communications* 146(2019) 103-109.
- [10] A.N. Uwaechia, N.M. Mahyuddin, M.F. Ain, N.M.A. Latiff, N.F. Za'bah, Compressed channel estimation for massive MIMO-OFDM systems over doubly selective channels, *Physical Communication* 36(2019) 1-16.
- [11] Q. Qin, L. Gui, B. Gong, S. Luo, Sparse channel estimation for massive MIMO-OFDM systems over time-varying channels, *IEEE Access* 6(2018) 33740-33751.
- [12] W. Zhang, X. Gao, Z. Li, Y. Shi, Pilot-assisted MIMO-V-OFDM systems: compressed sensing and deep learning approaches, *IEEE Access* 8(2020) 7142-7159.
- [13] A. Mishra, A.K. Jagannatham, L. Hanzo, Sparse Bayesian learning-aided joint sparse channel estimation and ML Sequence detection in space-time trellis coded MIMO-OFDM systems, *IEEE Trans. Communications* 68(2)(2020) 1132-1145.
- [14] D.C. Araújo, A.L.F.D. Almeida, J.P.C.L.D. Costa, R.T.D. Sousa, Tensor-based channel estimation for massive MIMO-OFDM systems, *IEEE Access* 7(2019) 42133-42147.
- [15] T. Do, L. Gan, N. Nguyen, T.D. Tran, Sparsity adaptive matching pursuit algorithm for practical compressed sensing, in: *Proc. the 42nd Asilomar Signals, Systems and Computers*, 2008.
- [16] G. Li, T. Li, An improved SAMP scheme for sparse OFDM channel estimation, in *Proc. IEEE 8th Joint International Information Technology and Artificial Intelligence*, 2019.
- [17] X. Wang, Y. Zhang, Z. Huang, Regularized backtracking adaptive pursuit algorithm based variable step size, *Acta Electronica Sinica* 46(8)(2018) 1829-1834.
- [18] C.-C. Tseng, J.-Y. Wu, T.-S. Lee, Enhanced compressive downlink CSI recovery for FDD massive MIMO systems using weighted block-minimization, *IEEE Trans. Communications* 64(3)(2016) 1055-1067.
- [19] W. Shen, L. Dai, Y. Shi, B. Shim, Z. Wang, Joint channel training and feedback for FDD massive MIMO systems, *IEEE Trans. Vehicular Technology* 65(10)(2016) 8762-8767.
- [20] V.K.N. Lau, S. Cai, A. Liu, Closed-loop compressive CSIT estimation in FDD massive MIMO systems with 1 bit feedback, *IEEE Trans. Signal Processing* 64(8)(2016) 2146-2155.
- [21] A. Liu, F. Zhu, V.K.N. Lau, Closed-loop autonomous pilot and compressive CSIT feedback resource adaptation in multi-user FDD massive MIMO systems, *IEEE Trans. Signal Processing* 65(1)(2017) 173-183.
- [22] A. Liu, V.K.N. Lau, W. Dai, Exploiting burst-sparsity in massive MIMO with partial channel support information, *IEEE Trans. Wireless Communications* 15(11)(2016) 7820-7830.
- [23] J.-C. Shen, J. Zhang, E. Alsusa, K.B. Letaief, Compressed CSI acquisition in FDD massive MIMO: how much training is needed?, *IEEE Trans. Wireless Communications* 15(6)(2016) 4145-4156.
- [24] M. Khumalo, W.-T. Shi, C.-K. Wen. Fixed-point implementation of approximate message passing (AMP) algorithm in massive MIMO systems, *Digital Communications and Networks* 2(4)(2016) 218-224.
- [25] X. Zhang, *Matrix Analysis and Applications*, second ed., Tsinghua University Press, Beijing, (2013) (Chapter 6).