

RESEARCH

Open Access



# Improving MIMO relay compressed sensing-based channel estimation and pilot allocation

Abbas Akbarpour-Kasgari<sup>1</sup> and Mehrdad Ardebilipour<sup>1\*</sup>

## Abstract

Increasing the data rate of communication systems together with the performance are main goals of the communication networks in the future. Multiple-input multiple-output (MIMO)-orthogonal frequency division multiplexing (OFDM) relay networks as a key technology is one the main techniques to maintain the performance and data rate. Being influenced by fading channels, the MIMO-OFDM relay networks need some reliable approaches to estimate the channel response. Compressed sensing (CS) is one of the critical approaches to maintain the accuracy together with the spectral efficiency. In this paper, we have utilized CS-based approaches to estimate the MIMO-OFDM relay channel. Specifically, forward backward prediction (FBP) is used to estimate the fading channels. This approach benefits from backward correction which distinguishes it from other iterative approaches and helps interestingly in channel estimation accuracy. Moreover, in order to improve the channel estimation, pilot allocation approaches are proposed based on the system model and probability function. Furthermore, a cross entropy-based approach is utilized to propose two different approaches called sequential cross entropy self-coherence (SCE-SC) and parallel cross entropy self-coherence. Actually, mutual coherence is divided into two parts namely cross coherence and self coherence. It is demonstrated that in FBP-based approach self-coherence is important in mutual coherence. Consequently, the computations are interestingly decreased. The superiority of the method is represented using comparing simulations with other well-known approaches.

**Keywords:** Cross entropy, Compressed sensing, Forward-backward prediction, MIMO-OFDM relay channel estimation, Pilot allocation

## 1 Introduction

By increasing demand of data rate and coverage area in modern communication systems, utilizing of relay communication networks is necessitated. In order to combat with the frequency selectivity of the channel and long-distance impairments, multiple-input multiple-output (MIMO)-orthogonal frequency division multiplexing (OFDM) is used as the pioneer technology [1–3]. Supplying some advanced features in MIMO-OFDM relays such as beamforming, relay selection, and power minimization required the active nodes in the network to

be aware of channel state information (CSI) [4, 5]. Pilot-aided channel estimation suffers from bandwidth efficiency which is very essential for the high-rate communication. To provide bandwidth-efficient pilot-aided channel estimation, compressed sensing (CS) is emerged, recently [6]. Wireless channels could be modeled by sparse signals, since there are considerable diffusion sources in the wireless environments. Hence, CS is very precious in sparse channel estimation together with the increase of accuracy and bandwidth efficiency. Furthermore, block sparse behavior of MIMO communication channels which is resulted by the existence of common scatterers, is the critical characteristic which is utilized by researchers in recent MIMO-OFDM compressed channel estimation [7]. As a consequence of joint sparsity, the support of the different channel ensembles between transmit-receive antenna pairs in MIMO nodes is identically distributed.

\*Correspondence: [mehrdad@eed.kntu.ac.ir](mailto:mehrdad@eed.kntu.ac.ir)

<sup>1</sup>Department of Electrical Engineering, K. N. Toosi University of Technology, Shariati St., Tehran, Iran

Thus, channel estimation could be extended to block-sparse signal processing in MIMO communications [8, 9]. At first, distributed CS (DCS)-simultaneous orthogonal matching pursuit (SOMP) was developed for channel estimation in single-input single-output (SISO)-OFDM channel estimation in [10]. Then, the authors of [11] utilized DCS-SOMP-based channel estimation for MIMO channel identification. Subsequently, a joint-OMP algorithm has been proposed to estimate the CSI of massive MIMO (mMIMO) in [12]. Recently, MIMO-OFDM relay channel estimation has been developed using DCS-based approach and a novel algorithm has been proposed using compressive sampling matching pursuit (CoSaMP) called block-verified CoSaMP (B-vCoSaMP) [13]. Furthermore, CS-based channel estimation is getting more attractive in massive MIMO communications [14, 15] which utilize angular domain sparsity and time domain sparsity.

Designing appropriate pilot sequences to improve channel estimation performance is the other key obstacle in front of researchers. In DCS-based channel estimation, the trivial pilot pattern is the random pilot allocation; while utilizing restricted isometry property (RIP), the measurement matrix could be designed to develop channel estimation. Of course, mutual coherence is optimized instead of RIP since there is no polynomial time [16]. In DCS-based channel estimation, the pilot sequences are translated to the measurement matrix identification and one may optimize the mutual coherence to design appropriate pilot sequence. It is shown that designing pilot sequences in terms of mutual coherence optimization is a combinatorial optimization. Hence, evolutionary algorithms are used to design pilot sequences. Specifically, the authors in [17] proposed a genetic algorithm (GA)-based pilot allocation algorithm in CS-based channel estimation. Furthermore, in [18] a suboptimal pilot allocation algorithm is designed based on GA for DCS-based channel estimation. Moreover, Qi et al. in [19] utilize the estimation of distribution algorithm (EDA) to define optimized pilot positions in SISO-OFDM systems. Moreover, He et al. [18] and Akbarpour-Kasgari and Ardebilipour [13] generate optimized pilot sequences for MIMO-OFDM compressed channel estimation using a GA and cross entropy (CE)-based approach.

In this paper, we have proposed to utilize a forward backward prediction (FBP) algorithm in order to increase the estimation accuracy. We proposed a FBP-based channel estimation algorithm where the algorithm further to forward collecting steps consists of backward elimination steps. In forward steps, the best atom is gathered. Further, exploiting backward steps makes it possible to omit bad atoms gathered in current and previous iterations. Hence, it would increase the accuracy of estimation rather than OMP and CoSaMP. In OMP and CoSaMP, the backward elimination steps are absent; thus, the previously

unsuitable atoms would increase the error in channel estimation and decrease the accuracy as well. To exploit the proposed FBP, we design the channel matrix and measurement matrix to exploit the common sparsity using  $\zeta$ -norm which is introduced. The system model is formulated using matrix representation to exploit the common sparsity of the channel. Then an FBP-based channel estimation is developed to estimate the channel coefficients in a time domain. By simulation results, it is validated that the proposed method is superior than the other existing methods. The superiority is caused by the backward stages which omit the evil atoms for accuracy amplification.

Moreover, we have introduced two probability-based pilot allocation called sequential cross entropy self-coherence (SCE-SC) and parallel cross entropy self-coherence (PCE-SC). As mentioned, in order to exploit the FBP-based channel estimation and  $\zeta$ -norm, we have designed appropriate measurement matrix using pilot symbols. It is represented in [13] that mutual coherence is related to the self-coherence of the sub-matrices and cross coherence of the sub-matrices related to each of the transmitter antenna, simultaneously. Here, we have changed the measurement matrix and we have shown that mutual coherence is only related to the self-coherence of the transmitter's sub-matrices. Consequently, the computation burden is considerably reduced rather than the previous works in [9, 13]. Using self-coherence phenomena, we have designed more simpler mutual coherence fitness function to be optimized. To optimize the introduced fitness function, we have proposed the probability-based pilot allocation algorithms, specifically SCE-SC and PCE-SC, where we have tried to optimize the new fitness function using a probability approach. SCE-SC performs optimization based on sampling the probability density function (pdf) of the available subcarriers. The pdf is updated in each iteration of the algorithm and then the fitness function is calculated. Afterward, the updated pdf is sampled for the next iteration. The iterations are continued till the steady state is resulted in the fitness function. In SCE-SC, in each iteration, only one pdf is followed. It will reduce the speed of convergence. In order to increase the speed of convergence, PCE-SC is developed where the number of pdfs are followed, simultaneously.

The contributions of the paper are summarized as follows:

- At first, we have proposed a novel approach to estimate the channel impulse response. In the proposed method which is based on FBP, the extracted appropriate atoms are reconsidered in the backward stages to be omitted if they are not suitable enough for estimation.
- In order to exploit the joint sparsity in a FBP approach, the channel measurement matrix is

developed utilizing the introduced norm. Hence, the Gram matrix is in diagonal matrix form, and as a consequence, mutual coherence could be defined using self-coherence metrics.

- A novel approach called SCE-SC is designed using a cross entropy-based method to optimize the self-coherence of the developed measurement matrix.
- To improve the pilot sequences and considering multiple probability density functions, PCE-SC is introduced which is more rapid and accurate in convergence rather than SCE-SC.

The remainder of the paper is as follows. The methods are demonstrated in Section 2. The system model is represented in Section 3. Section 4 covers the channel estimation approach using the proposed FBP. The pilot allocation scheme for SISO and MIMO systems is described in Section 5. Moreover, the proposed pilot allocation methods are represented in Section 6. Eventually, numerical results are expressed in Section 7 and concluding remarks are demonstrated in Section 8.

*Notations:* Matrices and vectors are denoted by uppercase and lowercase boldfaced letters, respectively.  $|\cdot|$  and  $(\cdot)^*$  denote the complex modulus and the conjugate of a complex number. For a given matrix  $\mathbf{A}$ ,  $\mathbf{A}^T$ ,  $\mathbf{A}^H$ , and  $\text{Trace}(\mathbf{A})$  denote its transpose, conjugate transpose, and trace, respectively, and  $\mathbf{A}_{i,j}$  denotes the  $(i,j)$ th element of  $\mathbf{A}$ . For a given vector  $\mathbf{x}$  with its element denoted by  $x_l$ ,  $\|\mathbf{x}\|_2 = \sqrt{\mathbf{x}^H \mathbf{x}}$  represents the Euclidean norm,  $\|\mathbf{x}\|_1 = \sum_l |x_l|$  is the  $l_1$ -norm, and  $\text{diag}(\mathbf{x})$  denotes a diagonal matrix with  $\mathbf{x}$  on its main diagonal. For two vectors  $\mathbf{x}$  and  $\mathbf{y}$ ,  $\langle \mathbf{x}, \mathbf{y} \rangle$  denotes their inner product. For a given set  $\Lambda$ ,  $n(\Lambda)$  is the number of elements in  $\Lambda$ .  $\mathbb{C}^{m \times n}$  stands for the set of all complex-valued  $m \times n$  matrices, and  $\emptyset$  denotes the null set.

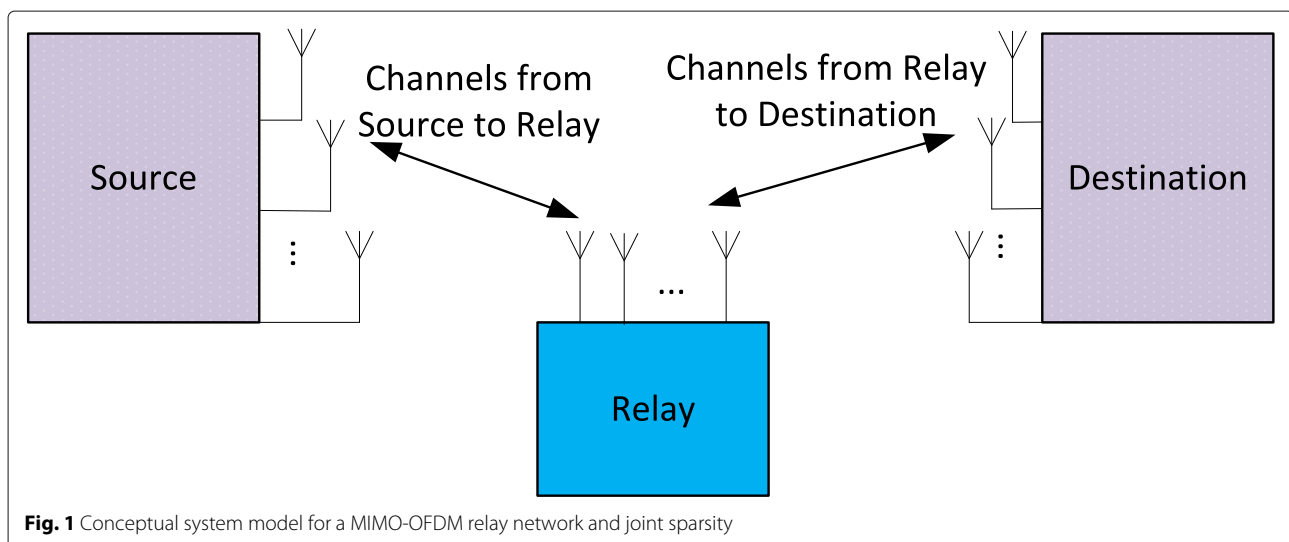
## 2 Methods

In this paper, the channel estimation is addressed using CS perspective. In order to increase the channel estimation accuracy together with decreasing the spectrum utilization, we have proposed the FBP-based channel estimation and SCE-SC and PCE-SC pilot allocation algorithms. To compare the proposed method with the existing ones, we have considered a GA-based approach and B-vCoSaMP-based channel estimation.

## 3 System model

Consider an amplify-and-forward (AF) relay network consists of MIMO terminals (Fig. 1). The network consists of a source node ( $\mathbb{S}$ ), a relay node ( $\mathbb{R}$ ), and a destination node ( $\mathbb{D}$ ). Besides, each terminal is equipped with  $N_{\mathbb{S}}$ ,  $N_{\mathbb{R}}$ , and  $N_{\mathbb{D}}$  transceiver antennas, respectively. Without loss of generality, we consider OFDM transmission in MIMO terminals. Consider an OFDM system with  $N$  subcarriers where  $N_p$  of them are selected as pilot subcarriers. Besides, to omit the interference of other antenna's pilots on each other, we consider to utilize the orthogonal pilot allocation, i.e., not only  $N_p$  subcarriers are allocated to pilot subcarriers, but also pilot subcarriers assigned to other  $N_{\mathbb{S}} - 1$  transmit antennas are reserved to be zero [20]. Thus, the number of data subcarriers on each of the transmit antenna is equal to  $N - N_{\mathbb{S}}N_p$ . Assuming  $\mathbf{x}_m \in \mathbb{C}^{N \times 1}$  as the OFDM symbol before cyclic prefix (CP) adding to be transmitted on the  $m$ th transmit antenna of  $\mathbb{S}$ . In the first time slot of the time-division-duplexing (TDD), the received signal  $\mathbf{r}_q$  for  $q = 1, 2, \dots, N_{\mathbb{R}}$  is formulated as

$$\mathbf{r}_q = \sum_{m=1}^{N_{\mathbb{S}}} \mathbf{X}_m \mathbf{F}_m^L \mathbf{f}_{mq} + \mathbf{v}_q \quad (1)$$



where  $\mathbf{X}_m$  is the diagonal matrix with  $\mathbf{x}_m$  as its main diagonal,  $\mathbf{F}_m^L$  is the partial discrete Fourier transform (DFT) matrix with  $N_p$  rows corresponding to the  $N_p$  pilot subcarriers of  $m$ th transmit antenna and first  $L$  columns of  $N \times N$  DFT matrix,  $\mathbf{f}_{mq} = [f_{mq}(0), f_{mq}(1), \dots, f_{mq}(L-1)]^T$  is the channel vector between source node  $\mathbb{S}$  and relay node  $\mathbb{R}$  with length  $L$  which  $K$  of them are non-zero to represent the channel sparsity. Besides,  $\mathbf{v}_q$  is the additive white Gaussian noise (AWGN) vector in the  $q$ th antenna of relay  $\mathbb{R}$ . In the second time slot, relay  $\mathbb{R}$  amplifies and retransmits the received signal to the destination node  $\mathbb{D}$ . As a consequence, the received pilots in the  $n$ th antenna of destination  $\mathbb{D}$  could be formulated as

$$\mathbf{y}_n = \sum_{q=1}^{N_{\mathbb{R}}} \mathbf{X}_m \mathbf{F}_m^{2L-1} \mathbf{h}_{mn} + \mathbf{z}_n \quad (2)$$

where  $\mathbf{h}_{mn} = \sum_{m=1}^{N_{\mathbb{S}}} \beta \mathbf{f}_{mq} * \mathbf{g}_{qn}$  where  $*$  stands for convolution, is the overall channel between  $\mathbb{S}$  and  $\mathbb{D}$  passing by  $\mathbb{R}$  and  $\mathbf{g}_{qn}$  is the channel vector between  $q$ th antenna of  $\mathbb{R}$  and  $n$ th antenna of  $\mathbb{D}$  and  $\mathbf{F}_m^{2L-1}$  is the partial DFT matrix with  $N_p$  rows corresponding to the  $N_p$  pilot subcarriers of  $m$ th transmit antenna and first  $2L-1$  columns of  $N \times N$  DFT matrix. Collecting all the received pilot sequences, we can represent the  $N_p \times N_{\mathbb{D}}$  received pilot matrix as  $\mathbf{Y} = [\mathbf{y}_{00}, \mathbf{y}_{01}, \dots, \mathbf{y}_{0N_{\mathbb{D}}-1}]$  and collecting all the channel ensembles, we can represent the  $2L-1 \times N_{\mathbb{D}}$  as  $\mathbf{H} = [\mathbf{h}_{00}, \mathbf{h}_{01}, \dots, \mathbf{h}_{0N_{\mathbb{D}}-1}]$ . Hence, the extension of (2) could be represented in matrix form for single-input multiple-output (SIMO) case as

$$\mathbf{Y} = \Phi \mathbf{H} + \mathbf{Z} \quad (3)$$

where  $\mathbf{Z}$  is the AWGN matrix with corresponding columns according to the  $\mathbf{z}_{0v}$ .  $\Phi$  is the measurement matrix with size of  $N_p \times L$ . Defining  $\zeta$  norm for matrices as  $\zeta(\mathbf{H}) = \text{card}\{\|\mathbf{H}_v\|_2 \neq 0\}$  where  $\mathbf{H}_v$  is the  $v$ th column of  $\mathbf{H}$ , estimating the channel could be accomplished by following optimization criterion.

$$\begin{aligned} \min_{\mathbf{H}} \quad & \|\mathbf{Y} - \Phi \mathbf{H}\|_2^2 \\ \text{s.t.} \quad & \zeta(\mathbf{H}) \leq K \end{aligned} \quad (4)$$

Obviously,  $K$  is the maximum sparsity of the columns of  $\mathbf{H}$ . We called the objective function as  $\mathbf{F}(\mathbf{H}) = \|\mathbf{Y} - \Phi \mathbf{H}\|_2^2$ . Utilizing  $\zeta(\mathbf{H})$ , we can exploit the joint sparsity of the channel ensembles in the system. The objective function in (4) represents the error of channel estimation method and the constraint controls the sparsity order of the channel ensembles in  $\mathbf{H}$ . Moreover, using  $\zeta$ -norm definition, the block sparsity of the channels are exploited.

As mentioned, the channel between transmit-receive pair which is denoted by  $\mathbf{h}_{uv}$  consists of  $L$  resolvable paths. These resolvable paths are resulted from  $L$  scatterers which are encountered by the signal conveying from the transmitter antenna  $u$  to the receiver antenna

$v$ .  $K$  of these  $L$  scatterers are significant scatterers where  $K \ll L$ . Consequently, the channel could be modeled using sparse vectors. Moreover, the signal conveyed distance is large relative to the transmit-receive antenna spacing in each terminal. Hence, the encountered scatterers in each chip period is identical between different antennas. In other words, the delays of different paths are the same in all the channel ensembles between two terminals. Thus, the sparsity pattern of different channel pairs could be assumed to be the same while the channel attenuation is different. Each path consists of different subpaths which are scattered from different scatterers which are zero-mean and identically independent distributed (i.i.d.). Thus, each path's attenuation is assumed to be  $\mathcal{CN}(0, \sigma^2)$ . Hence, the channel coefficient could be represented as

$$h_{uv}(l) = \sum_{i=0}^{I-1} \alpha_{uv}(i) g(lT - \tau(i)) \quad (5)$$

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

#### 4 Forward-backward pursuit channel estimation

In order to handle the optimization in (4), we have proposed a forward-backward pursuit (FBP) based on [21] where  $\ell_0$  norm was used. The algorithm is represented in details in Algorithm 1. Equation 4 could be solved using three different methods as convex relaxation, greedy methods, and message-passing (MP) algorithms. FBP which is based on MP algorithm is used here because of its forward selection and backward fixing. Specifically, OMP which is a greedy algorithm is a special case of FBP which the forward selection is present but the backward fixing is absent; consequently, it cannot fix its own mistakes in the previous steps. Moreover, the FBP algorithm constructs the new subspace by adding just one atom to the previous subspace, and in the backward steps, it reconstructs the subspace by omitting bad atoms. As a consequence, the proposed FBP algorithm could be compared with its greedy one called OMP, where OMP is the special case of FBP without backward steps to increase the estimation accuracy.

At the end of forward stages, a metric called  $\delta_F^{(t)}$  is defined which represents the difference by considering the new added atom. In the backward stages, we first consider

**Algorithm 1** FBP-based channel estimation

---

1: **Input:** the measurement matrix  $\Phi$ , the received pilot matrix  $\mathbf{Y}$  and sparsity order  $K$

2: **Output:** the estimation of channel impulse responses  $\mathbf{H}$

3:  $\mathbf{R}^{(0)} = \mathbf{Y}$ ,  $\lambda^0 = \phi$  and  $t = 0$

4: **while** stop criterion not met **do**

5:  $i^{(t)} = \arg \max_{j \notin \lambda^{(t-1)}} \|(\mathbf{R}^{(t)})^H \phi_j\|_2$

6: Merging supports of the pervious iteration and the present one  $\lambda^{(t)} = \lambda^{(t-1)} \cup i^{(t)}$

7: Define  $\mathbf{H}^{(t)} = \Phi_{\lambda^{(t)}}^\dagger \mathbf{Y}$

8:  $\delta_F^{(t)} = \mathbf{F}(\mathbf{H}^{(t-1)}) - \mathbf{F}(\mathbf{H}^{(t)})$

9: **while** 1 **do**

10:  $j^{(t)} = \arg \min_{j \in \lambda^{(t)}} \mathbf{F}(\mathbf{H}^{(t)} - \mathbf{H}_{j'}^{(t)})$

11:  $\delta_B^{(t)} = \mathbf{F}(\mathbf{H}_{j'}^{(t)}) - \mathbf{F}(\mathbf{H}^{(t)})$

12: **if**  $\delta_B^{(t)} \geq 0.5\delta_F^{(t)}$  **then**

13: Update residual  $\mathbf{R}^{(t)} = \mathbf{Y} - \Phi_{\lambda^{(t)}} \mathbf{H}^{(t)}$

14: **break**

15: **end if**

16: Omit bad atom  $\lambda^{(t)} = \lambda^{(t)} - j^{(t)}$

17: Update  $\mathbf{H}^{(t)} = \Phi_{\lambda^{(t)}}^\dagger \mathbf{Y}$

18: Update residual  $\mathbf{R}^{(t)} = \mathbf{Y} - \Phi_{\lambda^{(t)}} \mathbf{H}^{(t)}$

19: **end while**

20: **end while**

---

all the collected atoms individually and their effect on the residual is considered by calculating  $\arg \min_{j \in \lambda^{(t)}} \mathbf{F}(\mathbf{H}^{(t)} - \mathbf{H}_{j'}^{(t)})$  and subsequently  $\delta_B^{(t)}$  is computed to determine the worst collected atom. These calculation is absent in other known algorithms which makes this algorithm an efficient one in the CS-based channel estimation approaches.

In Algorithm 1,  $|i^{(t)}|$  denotes the number of selected atoms in forward step, and  $\mathbf{H}_{j'}^{(t)}$  represents the  $\mathbf{H}^{(t)}$  while  $j^{(t)}$ th column is omitted.

In order to extend the proposed method to the MIMO case, one can extend the measurement matrix and channel matrix as follows. In order to extend the channel matrix, we add the other transmitting antenna caused channels row-wise to the each other and represent channel matrix  $\mathbf{H} \in \mathbb{C}^{(2L-1)N_S \times N_D}$  as

$$\mathbf{H} = \begin{pmatrix} \mathbf{h}_{00} & \mathbf{h}_{01} & \dots & \mathbf{h}_{0(N_D-1)} \\ \mathbf{h}_{10} & \mathbf{h}_{11} & \dots & \mathbf{h}_{1(N_D-1)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{h}_{(N_S-1)0} & \mathbf{h}_{(N_S-1)1} & \dots & \mathbf{h}_{(N_S-1)(N_D-1)} \end{pmatrix}. \quad (6)$$

Moreover, the measurement matrix  $\Phi \in \mathbb{C}^{N_S N_p \times (2L-1)N_S}$  is extended as

$$\Phi = \begin{pmatrix} \Phi_0 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \Phi_1 & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \Phi_{N_S-1} \end{pmatrix} \quad (7)$$

where  $\mathbf{0}$  is  $N_p \times (2L-1)$  zeros matrix. Furthermore, the received pilots are gathered in  $\mathbf{Y} \in \mathbb{C}^{N_S N_p \times N_D}$  as

$$\mathbf{Y} = \begin{pmatrix} \mathbf{Y}_{00} & \mathbf{Y}_{01} & \dots & \mathbf{Y}_{0(N_D-1)} \\ \mathbf{Y}_{10} & \mathbf{Y}_{11} & \dots & \mathbf{Y}_{1(N_D-1)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{Y}_{(N_S-1)0} & \mathbf{Y}_{(N_S-1)1} & \dots & \mathbf{Y}_{(N_S-1)(N_D-1)} \end{pmatrix}. \quad (8)$$

Using the introduced matrices, one can utilize the proposed FBP algorithm to estimate MIMO-OFDM relay system channels. One of the main advantages of the proposed approach is the advantage of the measurement matrix in designing optimal pilot subcarriers to improve the channel estimation accuracy which will be discussed in the following sections.

## 5 Pilot allocation for compressed channel estimation

In FBP, jointly sparse channels are estimated altogether. The space existing between adjacent antennas in MIMO nodes is close to each other where the sparsity pattern between transmit-receive pairs is the same. Moreover, the channel coefficients in each of the non-zero paths are not the same and are rayleigh random variable since they are a consequence of the number of normal distributed paths. Utilizing the FBP, the jointly sparse channels could be estimated altogether. The measurement matrix which is represented in (7) could be generated using random pilot subcarriers and optimized pilot subcarriers. In order to improve the accuracy of estimated channels, it is mandatory to select pilot subcarriers to optimize the estimation metric. In CS, RIP is used as the key metric in designing appropriate measurement matrices. But, there is no polynomial time approach to calculate RIP; thus, we have used mutual coherence to design optimal measurement matrix.

Mutual coherence is defined as

$$\mu\{\Phi\} = \max_{\lambda_i, \lambda_j; \lambda_i \neq \lambda_j} \left| \sum_{l=0}^{2L-2} e^{-j2\pi(\lambda_i - \lambda_j)l/N} \right| \quad (9)$$

where  $\lambda_i$  and  $\lambda_j$  are pilot subcarriers among  $N$  available subcarriers. Consequently, mutual coherence is defined as maximum off-diagonal entries of Gram matrix  $\mathbf{G}\{\Phi\} = \Phi^H \Phi$  if  $\Phi$  is orthonormal [13, 22]. Accordingly,  $\mu\{\Phi\}$  is related to the positions of pilots  $\lambda_i$  and  $\lambda_j$ . Since, they are positions, the problem can be deduced which is defined as

$$\min_{\lambda_i, \lambda_j} \mu\{\Phi\} \quad (10)$$

$$\lambda_i \neq \lambda_j$$

$$\lambda_i, \lambda_j \in \mathcal{A}$$

where  $\mathcal{A}$  is the set of all the available subcarriers. Obviously, the selection of subcarriers is a combinatorial optimization. In order to generate optimal pilot sequences, we have used probability-based approaches which will be discussed in the following sections. Here, we consider the measurement matrix  $\Phi$ .

**Theorem 1** Assuming  $\Phi$  as an orthonormal measurement matrix in Eq. (7), then  $\mu\{\Phi\}$  could be defined by

$$\max_{i \in \{1, 2, \dots, N_S\}} \mu\{\Phi_i, \Phi_i\} \quad (11)$$

Considering Gram matrix

$$\mathbf{G}\{\Phi\} = \begin{pmatrix} \Phi_1^H \Phi_1 & 0 & \dots & 0 \\ 0 & \Phi_2^H \Phi_2 & \dots & 0 \\ \dots & \dots & \ddots & \dots \\ 0 & 0 & \dots & \Phi_{N_S}^H \Phi_{N_S} \end{pmatrix} \quad (12)$$

where  $\mathbf{G}\{\Phi\} \in \mathbb{C}^{N_S(2L-1) \times N_S(2L-1)}$ , we define two concepts. The first one is self-coherence and the other one is the cross-coherence.  $\Phi$  is consist of different submatrices from different antennas called  $\Phi_i$  for  $i = 0, 1, \dots, N_S - 1$ . We define self-coherence as  $\mu\{\Phi_i, \Phi_i\}$  and cross-coherence as  $\mu\{\Phi_i, \Phi_j\}$  where  $i \neq j$  where

$$\mu\{\Phi_i, \Phi_i\} = \max_{k \neq t} \phi_{ik}^H \phi_{it} \quad (13)$$

$$\mu\{\Phi_i, \Phi_j\} = \max_{k \neq t} \phi_{ik}^H \phi_{jt} \quad (14)$$

and  $\phi_{ik}$  is the  $k$ th column of  $\Phi_i$ . Obviously, according to (12), the mutual coherence could be defined as the maximum of self-coherence between all the antennas. ■

According to the above, among other jointly sparse estimation, our proposed method and formulation leads to self-coherence while others lead to self-coherence and cross-coherence [11, 13]. Consequently, it could decrease the number of computations since the number of matrix multiplications is  $N_S$ , while in other approaches the number of matrix multiplications is  $N_S N_D$ . Hence, the number of matrix multiplications is outstandingly minimized.

In order to omit the interference of antennas on each other, we have considered orthogonal pilot sequences which are demonstrated in [13].

## 6 Proposed pilot allocation algorithms

Here, we will demonstrate two numerical algorithms to optimize the mutual coherence of measurement matrix and defining optimized pilot sequences. The optimization is performed over the search space  $\mathcal{S} \in \binom{N}{N_p}$ . Hence, the exhaustive search is intractable and computationally inefficient by increased number of  $N$  and  $N_p$ . Consequently, the combinatorial optimization problem in Eq. (10) is solved using pdf sampling by the proposed algorithms.

The sampling is continued until convergence of the pdf. The steady state is achieved while the pdf consists of zero or  $1/N_p N_S$  values. The pdf in each of the iteration is updated utilizing appropriate population in each generation. In order to detect the appropriate individuals, we use the fitness function

$$\max_{i \in \{1, 2, \dots, N_S\}} \mu\{\Phi_i, \Phi_i\} \quad (15)$$

To decrease the computational complexity in fitness function, we used self-coherence of different measurement matrices. Evidently, there are lots of zero elements in matrix  $\Phi$  which are unnecessary to be multiplied. Moreover, the cross-coherence is not used in mutual coherence. Since, in Eq. (12) only self-coherence is included. Hence, we used self-coherence in Eq. (13) to optimize the computations in fitness function evaluation. In each generation, the updated pdf is utilized to generate the new individuals.

### 6.1 Sequential cross entropy self-coherence (SCE-SC) pilot allocation approach

In this approach, there are some definitions which are very important to understand the approach. *Generation* ( $\mathcal{G}$ ) is the set of pilot sequences which are sampled from the pdf. Each generation is constructed from ( $I$ ) *individuals* which are referred to each of the sample pilot sequence. Furthermore, each individual consists of  $N$  *elements*. The probability of each element is represented in a *pdf* called  $\mathcal{P}$  which demonstrates the probability of the element to be selected as an appropriate selection for the pilot sequence. Moreover, *elite individuals* in each generation are  $J$  individuals with best fitness function. Hence, *fitness function* is the metric for detecting elite individuals in each generation which is demonstrated in Eq. (15). The initial state of the pdf  $\mathcal{P}$  is the uniform pdf with elements equal to  $1/N$ . Gradually, during different generations, the pdf is updated using  $J$  elite individuals and its elements converge to the steady state. Since the generation of individuals consists of sampling from the probability function, this method is more robust than other mutation-based approaches in trapping to local minima.

The step-by-step representation of the SCE-SC algorithm is demonstrated in Algorithm 2. Considering  $p_g(\mathcal{G})$  as the pdf vector of  $g$ th iteration, we represent the pdf vector by

$$p_g(\mathcal{G}) = [p_g(\mathcal{G}_1), p_g(\mathcal{G}_2), \dots, p_g(\mathcal{G}_N)]^T \quad (16)$$

where  $p_g(\mathcal{G}_i)$  refers to the probability of obtaining a value of 1 in the  $i$ th element of  $\mathcal{G}$ . In each generation, the pdf vector is utilized to generate  $I$  individuals. Among these individuals,  $J$  elite individuals are selected. Then, pdf is

**Algorithm 2** Sequential cross entropy self-coherence (SCE-SC) pilot allocation

---

```

1: Initialize the remaining subcarrier set  $\mathcal{A} = [1 : N]$ 
2: for  $n = 1$  to  $N_s$  do
3:   Initialize the population  $\mathcal{G}^0$  using  $I$  random individuals with length  $N_p$  from remaining subcarrier set  $\mathcal{A}$ 
4:   while stop criteria not met do
5:     Sort the individuals in  $\mathcal{G}^g$  in ascending order according to corresponding self-coherence
6:     Select  $J$  of the individuals in  $\mathcal{G}^g$  with the best fitness function and establish subset  $\mathcal{G}_{j:I}^g$  for  $j = 1, 2, \dots, J$ 
7:     Construct pdf  $p_{g+1}(\mathcal{G})$  using  $\mathcal{G}_{j:I}^g$  and Eq. (17)
8:     Sample the probability function and generate the new individual without repeat and generate the new generation  $\mathcal{G}_{g+1}$ 
9:   end while
10:  Update the remaining subcarrier set  $\mathcal{A}$ 
11: end for

```

---

updated using these elite individuals according to

$$p_{g+1}(\mathcal{G}) = (1 - \tau)p_g(\mathcal{G}) + \frac{\tau}{J} \sum_{j=1}^J \mathcal{G}_{j:M}^g \quad (17)$$

where  $\mathcal{G}_{j:I}^g$  is the  $j$ th elite individual in  $g$ th generation and  $\tau$  is the regulation parameter of the algorithm which demonstrate the dependency of the update procedure to the current generation. Furthermore,  $\tau$  is non-zero positive lower than 1. According to the previous discussions, the algorithm is performed till the steady state is met. In the steady state, the pdf vector is only included for zero or  $1/N_s N_p$  elements.

**6.2 Parallel cross entropy self-coherence (PCE-SC) pilot allocation algorithm**

In the SCE-SC algorithm, the initial condition is set just one time. Consequently, iteratively the pdf is updated based on the initial generation. Since the initial generation is randomly selected, it would be helpful to revise the algorithm by generating some generations as the initial generation. Consequently, by selecting different initial conditions, we are pursuing the optimal point in some parallel avenues. Parallelism is helpful as it would decrease the probability of local minima trapping and the resulting sub-optimal point would be more robust and closer to the optimal point. Hence, in PCE-SC algorithm which is represented stage-by-stage in Algorithm 3,  $\alpha$  initial generation from the search space  $\mathcal{S}$  are considered and iteratively the pdf of these generations are converged to the steady state, simultaneously. Considering multiple pdfs at the same time will increase the computation burden which

**Algorithm 3** Parallel cross entropy self-coherence (PCE-SC) pilot allocation

---

```

1: Initialize the remaining subcarrier set  $\mathcal{A} = [1 : N]$ 
2: for  $n = 1$  to  $N_s$  do
3:   Initialize the population  $\mathcal{G}^0$  using  $I$  random individuals with length  $N_p$  from remaining subcarrier set  $\mathcal{A}$ 
4:   while stop criteria not met do
5:     for  $\alpha = 1$  to  $M$  do
6:       Sort the individuals in  $\mathcal{G}_\alpha^g$  in ascending order
7:       Select  $J$  of the individuals in  $\mathcal{G}_\alpha^g$  with the best objective value and establish subset  $\mathcal{G}_{\alpha,j:I}^g$  for  $j = 1, 2, \dots, J$ 
8:       Construct pdf  $p_g^\alpha(\mathcal{G})$  using  $\mathcal{G}_{\alpha,j:I}^g$  and Eq. (17)
9:       Sample the probability function and generate the new individual without repeat and generate the new generation  $\mathcal{G}_{g+1}^\alpha$ 
10:    end for
11:   end while
12:  Update the remaining subcarrier set  $\mathcal{A}$ 
13: end for

```

---

can be handled by the parallel processing units. However, the proposed method will decrease the local minima trapping more than the SCE-SC approach and increase the accuracy of the resulted steady state point. Consequently, utilizing PCE-SC will increase two substantial perspective of pilot allocation algorithm in channel estimation.

**7 Numerical results**

Here, we demonstrate the performance of the proposed algorithms in channel estimation and pilot allocations called FBP, SCE-SC, and PCE-SC by simulation results. The simulation parameters are listed in Table 1. Sparse Rayleigh channel is modeled using finite impulse response (FIR) filters where non-zero taps (significant taps) are independently and identically distributed (i.i.d.) utilizing zero mean and unit variance complex Gaussian variables. Moreover, Monte Carlo simulations are performed over 1000 independent runs and the results are averaged over them.

**Table 1** Simulation parameters

Simulation parameter	Value
Modulation	16-QAM
Number of subcarriers (N)	512
Number of transmit antenna ( $N_s$ )	4
Number of receive antenna ( $N_D$ )	4
Number of relay antenna ( $N_R$ )	4
Channel taps (L)	30
Channel significant taps (K)	3

Here, we will consider the applicability of the proposed algorithm and the obtained pilot sequences utilizing two main characteristics of channel estimator called NMSE and BER. NMSE is calculated using

$$\text{NMSE} = \frac{1}{N_{MC}} \sum_{i=1}^{N_{MC}} \frac{\|\mathbf{h}_i - \hat{\mathbf{h}}\|_2^2}{\|\mathbf{h}_i\|_2^2} \quad (18)$$

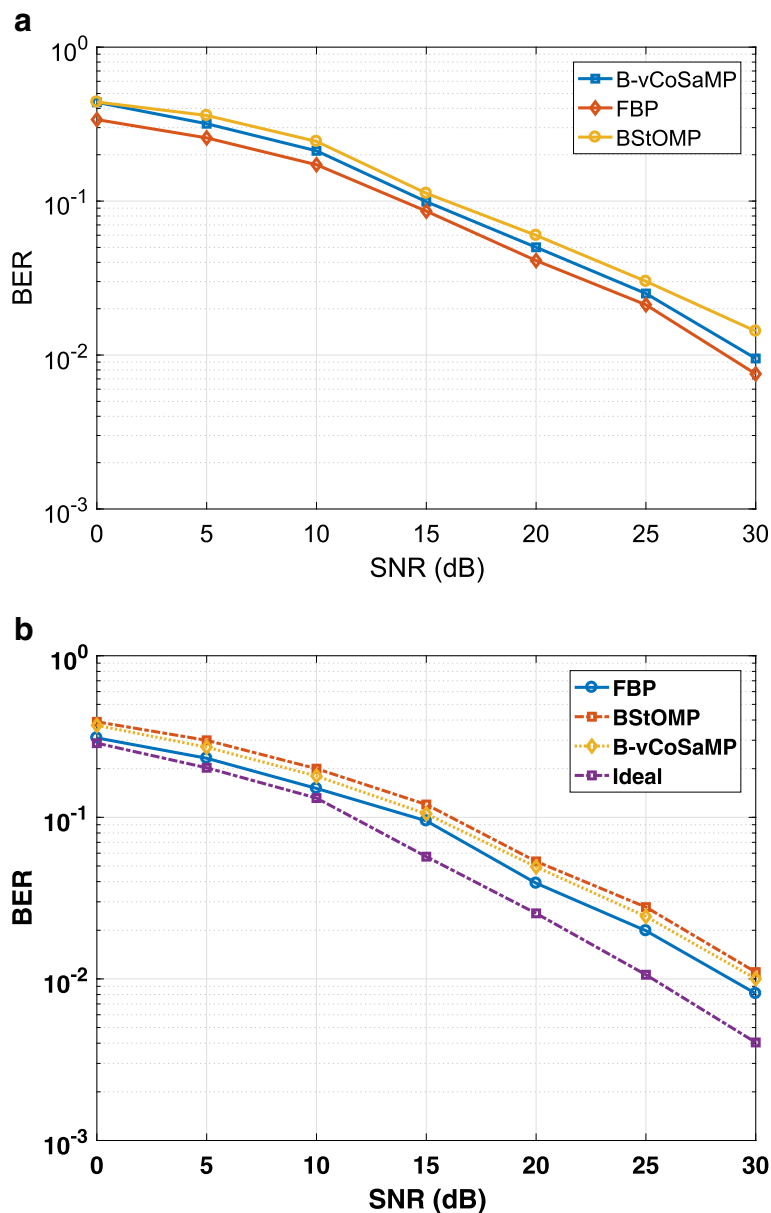
where  $\mathbf{h}$  is the complete channel vector and  $\hat{\mathbf{h}}$  is its estimation. Moreover,  $N_{MC}$  is the Monte Carlo iterations which is 1000. BER is evaluated using Monte Carlo simulation using  $N_{MC}$  individual simulation according to the following equation.

$$\text{BER} = \sum_{i=1}^{N_{MC}} \frac{N_b(i)}{N_t(i)} \quad (19)$$

where  $N_b(i)$  is the number of the correct received bits in  $i$ th Monte Carlo iteration and  $N_t(i)$  is the number of transmitted bits in each Monte Carlo iteration.

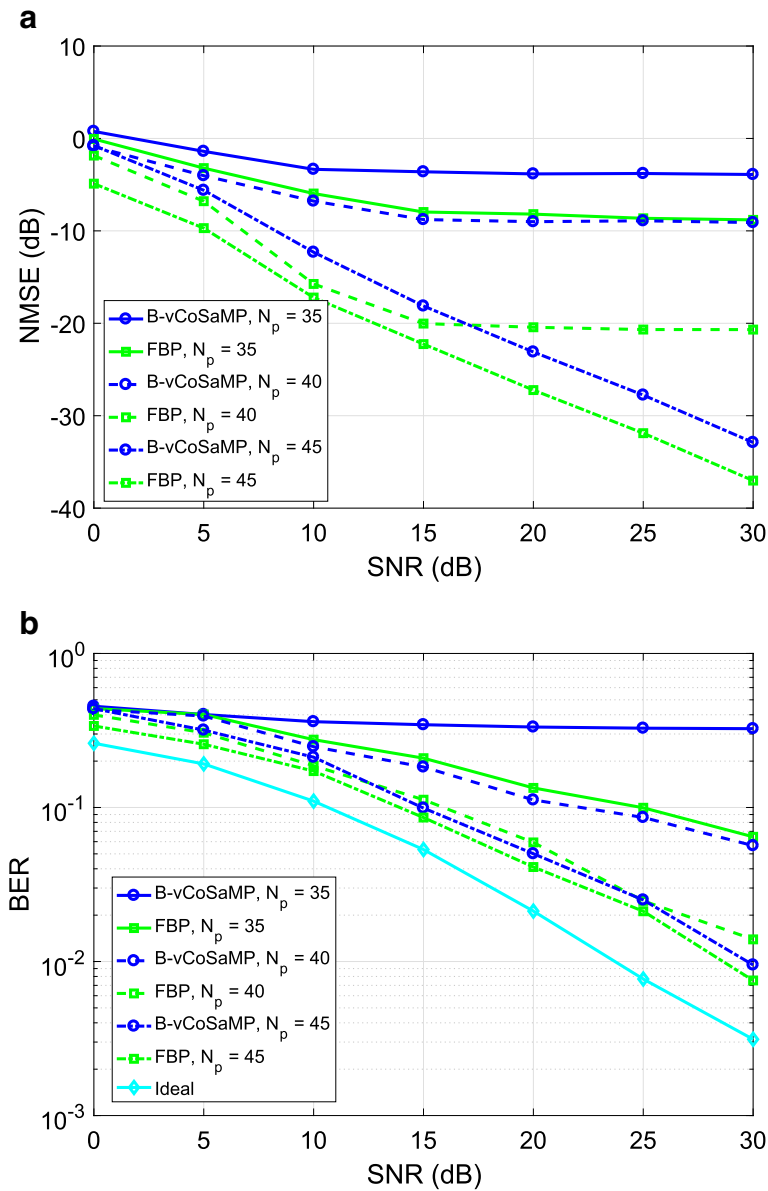
### 7.1 Comparison of channel estimation performance

The results of the proposed channel estimation method are represented in Fig. 2 considering MSE and BER, respectively. For comparison, we have utilized BStOMP algorithm developed in [23] and B-vCoSaMP developed in [13]. Here, in Fig. 3a, from the NMSE point of view, the



**Fig. 2** Comparison of the proposed FBP-based method, B-vCoSaMP, and BStOMP: **a** MSE and **b** BER





**Fig. 3** The effect of pilot numbers in channel estimation: **a** MSE and **b** BER

proposed approach is more efficient than others. Moreover, the comparison of BER is demonstrated and, as it is shown, the BER of the proposed method is superior almost 2 dB better than BStOMP approach and 1.8 dB better than B-vCoSaMP. Moreover, using only  $N_p = 45$  subcarriers as pilot, we are almost 3 dB away from the ideal case. This means that more than 65% spectral efficiency is maintained.

## 7.2 Comparison of the number of pilots

In order to compare the effect of the number of pilots, we have considered the proposed method and B-vCoSaMP

developed in [13], since these two methods were performed better than BStOMP. Moreover, the number of pilots is changed and selected to be 35, 40, and 45. Illustratively, the NMSE and BER comparisons are represented in Fig. 3. Obviously, in case of 35 pilots, both methods are irreducible while the proposed FBP-based approach is irreducible in lower NMSE. By increasing the number of pilots to 40 and 45, the estimation accuracy is getting more applicable. Furthermore, in all the cases of pilots, the proposed FBP-based approach is superior than the B-vCoSaMP. This superiority is obvious in BER, too.

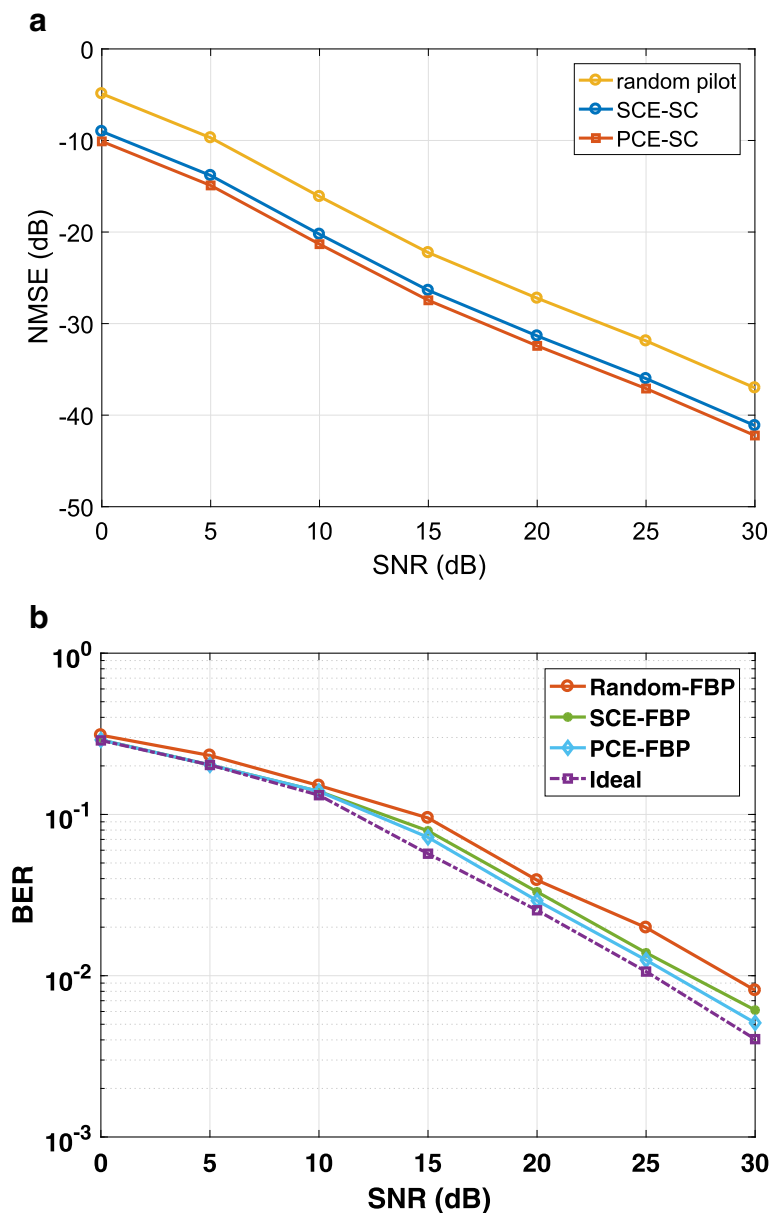
### 7.3 Comparison of proposed pilot allocation algorithms

Here, we consider pilot allocation algorithms proposed in this paper called PCE-SC and SCE-SC. In order to compare these proposed algorithms, we have considered NMSE and BER, separately. These figures are represented in Fig. 4. NMSE is compared in Fig. 4a and BER is compared in Fig. 4b. Obviously, using optimized pilot placement creates approximately 6 dB and 7 dB superiority than random placement utilizing SCE-based and PCE-based algorithms, respectively. Moreover, this superiority is encountered in BER represented in Fig. 4b almost 2 dB and 3 dB in higher SNRs. The BER gap between PCE-SC pilot placement using 45 pilots and the ideal case is

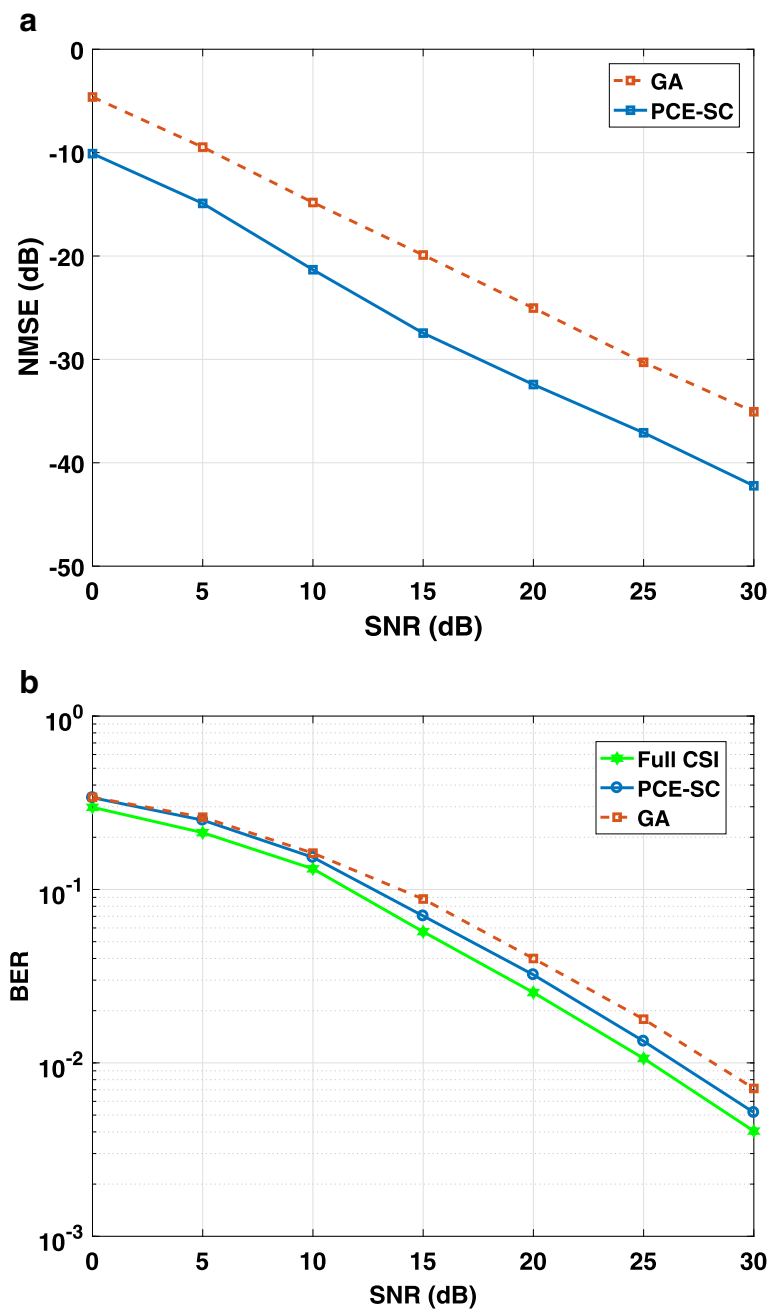
approximately 1 dB which represented the superiority of the proposed approach in terms of bandwidth efficiency. Among 512 subcarriers, only 100 of them is utilized as the pilots and others are used as data subcarriers. Hence, more than 65% of the subcarriers are used as the data subcarriers which can be used by transmitting antennas.

### 7.4 Comparison of the proposed pilot allocation algorithm with the existing one

In Fig. 5, two different algorithms for pilot allocation are compared. These two algorithms are defined by proposed PCE-SC and GA algorithm developed in [18]. To estimate the channel, we have utilized the proposed FBP-based



**Fig. 4** Comparison of different proposed pilot allocation algorithms utilizing (a) NMSE and (b) BER performance metric

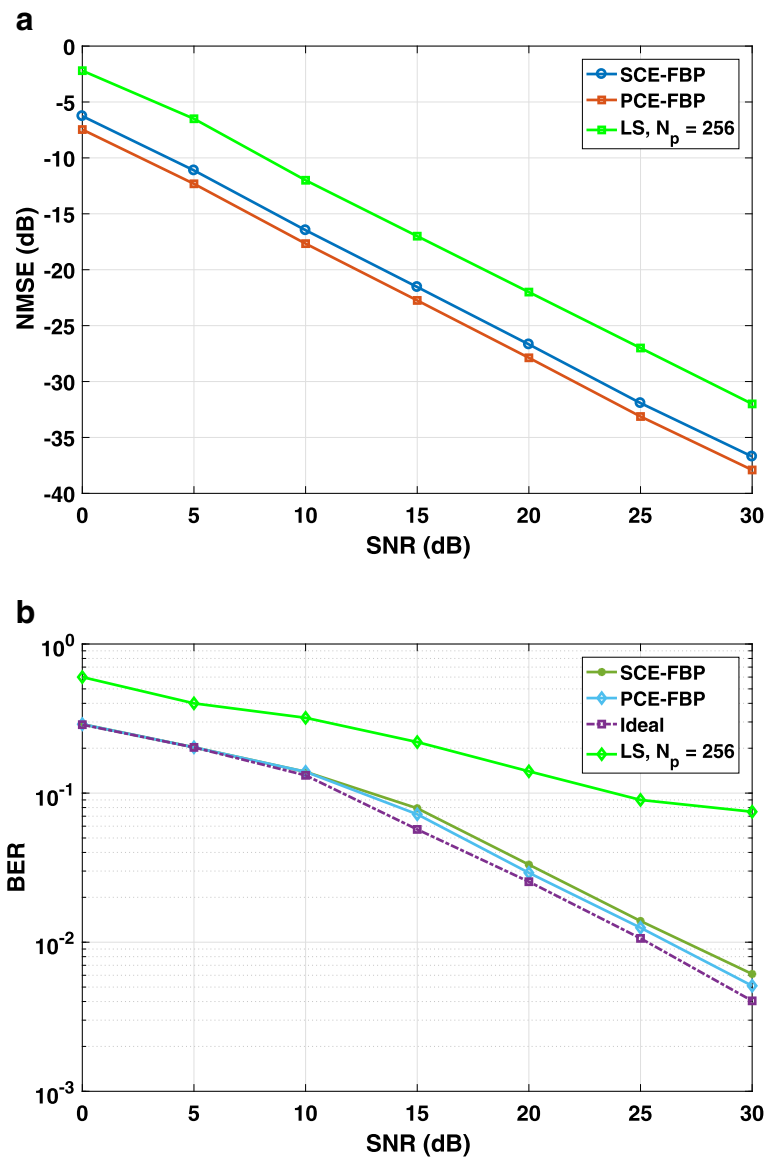


**Fig. 5** Performance comparison of proposed method and GA-based pilot allocation algorithm: **a** NMSE performance metric and **b** BER performance metric

channel estimation approach. The number of the pilots for each of the transmitting antenna is  $N_p = 45$ . Illustratively, the results of the NMSE and BER are represented in Fig. 5. Obviously, NMSE of the proposed method is better than a GA-based approach approximately 3 dB. Moreover, concerning BER, this superiority is almost 1.5 dB. Evidently, by utilizing  $N_p = 45$  optimized pilots on each antenna, the performance is almost 1 dB away from the optimal performance which could be compensated using other facilities.

Actually, using optimized pilots will lead to 65% spectral efficiency which is one of the most critical characteristics of the system.

In Fig. 6, the proposed pilot allocation algorithms are compared with the optimized least squares (LS)-based channel estimation as the state-of-the-art approach in channel estimation of MIMO-OFDM systems in [24]. Obviously, the LS-based channel estimation is performed using 256 number of pilots where the spectral efficiency



**Fig. 6** Performance comparison of proposed method and optimized LS-based pilot allocation algorithm: **(a)** NMSE performance metric, **(b)** BER performance metric

is extremely decreased. The simulation parameters for the proposed pilot allocation is as before.

## 8 Results and discussion

In this paper, channel estimation of AF MIMO relay is considered and CS-based approaches is utilized due to their spectral efficiency and accuracy improvement. Here, we have proposed FBP-based channel estimation algorithm for forward selection and backward elimination. The proposed method benefits from backward elimination to improve the accuracy of estimation. Moreover, the measurement matrix is introduced and its Gram matrix is developed to minimize the mutual coherence. As discussed, mutual coherence is related to the

self-coherence of sub-matrices related to each of the antennas. Hence, two pilot allocation algorithms are proposed based on the cross entropy where the number of multiplications are decreased rather than other approaches since they utilize cross-coherence and self-coherence, simultaneously. Accordingly, PCE-SC and SCE-SC algorithms are proposed for pilot generation and channel estimation accuracy improvement. They are compared with other approaches and their superiority validation is represented using simulation results.

### Abbreviations

BER: Bit error rate; B-vCoSaMP: Block-verified compressive sampling; CS: Compressed sensing; CSI: Channel state information; DCS: Distributed compressed sensing; EDA: Estimating of distribution algorithm; FBP: Forward

backward prediction; MIMO: Multiple-input multiple-output; NMSE: Normalized mean square error; OFDM: Orthogonal frequency division multiplexing; OMP: Orthogonal matching pursuit; PCE-SC: Parallel cross entropy self-coherence; RIP: Restricted isometry property; SCE-SC: Sequential cross entropy self-coherence

#### Funding

There is no source of funding for this paper.

#### Availability of data and materials

Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

#### Authors' contributions

Both authors carried out the mathematical proofs and developed the algorithms. AA participated in simulation part. Both authors read and approved the final manuscript.

#### Competing interests

This paper is written in order to fulfill the Ph.D. of telecommunication engineering and there are no other competing interests.

#### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 22 December 2017 Accepted: 27 February 2019

Published online: 21 March 2019

#### References

1. D. Hu, X. Wang, L. He, A new sparse channel estimation and tracking method for time-varying ofdm systems. *IEEE Trans. Veh. Technol.* **62**(9), 4648–4653 (2013)
2. V. R. Asghari, M. Ardebilipour, Spread spectrum code estimation by genetic algorithm. *Int. J. Signal Proc.* **1**(4), 301–304 (2004)
3. M. A. Mohammadi, M. Ardabilipour, B. Moussakhani, Z. Mobini, in *2008 5th IFIP International Conference on Wireless and Optical Communications Networks (WOCON'08)*. Performance comparison of RLS and LMS channel estimation techniques with optimum training sequences for MIMO-OFDM systems (IEEE, 2008), pp. 1–5
4. C. Qi, G. Yue, L. Wu, N. Arumugam, Pilot design for sparse channel estimation in OFDM-based cognitive radio systems. *IEEE Trans. Veh. Technol.* **63**(2), 982–987 (2014)
5. F. Azami, S. M. Safavi Hemami, A. Akbarpour-Kasgari, Joint power allocation and beamforming in amplify-and-forward relay networks under per-node power constraint. *Wirel. Commun. Mob. Comput.* **2017**(5681236), 9 (2017)
6. M. F. Duarte, Y. Eldar, Structured compressed sensing: From theory to applications. *IEEE Trans. Signal Process.* **59**(9), 4053–4085 (2011)
7. Z. Gao, L. Dai, W. Dai, S. B. Z. Wang, Structured compressive sensing-based spatio-temporal joint channel estimation for FDD massive MIMO. *IEEE Trans. Commun.* **64**(2), 601–617 (2016)
8. S. Sarvotham, D. Baron, M. Wakin, M. F. Duarte, R. G. Baraniuk, in *Asilomar conference on signals, systems, and computers*. Distributed compressed sensing of jointly sparse signals, (2005), pp. 1537–1541
9. A. Akbarpour-Kasgari, M. Ardebilipour, in *2017 Iranian Conference on Electrical Engineering (ICEE)*. MIMO relay compressed channel estimation using optimized pilot placements (IEEE, 2017), pp. 2028–2031
10. P. Cheng, Z. Chen, Y. Rui, Y. J. Guo, L. Gui, M. Tao, Q. T. Zhang, Channel estimation for OFDM systems over doubly selective channels: A distributed compressive sensing based approach. *IEEE Trans. Commun.* **61**(10), 4173–4185 (2013)
11. L. Xu, K. Niu, Z. He, W. Xu, Z. Zheng, in *2013 9th International Conference on Information, Communications & Signal Processing*. MIMO channel estimation based on distributed compressed sensing for LTE-advanced (IEEE, 2013), pp. 1–5
12. R. X, V. K. N. Lau, Distributed compressive CSIT estimation and feedback for FDD multi-user massive MIMO systems. *IEEE Trans. Signal Process.* **62**(12), 3261–3271 (2014)
13. A. Akbarpour-Kasgari, M. Ardebilipour, Probability-based pilot allocation for MIMO relay distributed compressed sensing based channel estimation. *Eurasip J. on Advances in Signal Process.* **2018**(1), 18 (2018)
14. Y. Han, J. Lee, D. J. Love, Compressed sensing-aided downlink channel training for FDD massive MIMO systems. *IEEE Trans. Commun.* **65**(7), 2852–2862 (2017)
15. L. Chen, A. Liu, X. Yuan, Structured turbo compressed sensing for massive MIMO channel estimation using a Markov prior. *IEEE Trans. Veh. Technol.* **67**(5), 4635–4639 (2018)
16. J. Tropp, Greed is good: Algorithmic results for sparse approximation. *IEEE Trans. Inf. Theory.* **50**(10), 2231–2242 (2004)
17. X. He, R. Song, W. P. Zhu, Pilot allocation for sparse channel estimation in MIMO-OFDM systems. *IEEE Trans. Circuits Syst. II, Exp. Briefs.* **60**(9), 612–616 (2013)
18. X. He, R. Song, W. P. Zhu, Pilot allocation for distributed-compressed-sensing-based sparse channel estimation in MIMO-OFDM systems. *IEEE Trans. Veh. Technol.* **65**(5), 2990–3004 (2016)
19. C. Qi, G. Yue, L. Wu, Y. Huang, A. Nallathan, Pilot design schemes for sparse channel estimation in OFDM systems. *IEEE Trans. Veh. Technol.* **64**(4), 1493–1505 (2015)
20. A. Akbarpour-Kasgari, M. Ardebilipour, F. Afghah, in *2015 49th Annual Conference on Information Sciences and Systems (CISS)*. MIMO-OFDM non-regenerative relay channel estimation using compressed sensing, (2015), pp. 1–5. IEEE
21. T. Zhang, Adaptive forward-backward greedy algorithm for learning sparse representations. *IEEE Trans. Inf. Theory.* **57**(7), 4689–4708 (2011)
22. J. M. Duarte-Carvajalino, G. Sapiro, Learning to sense sparse signals: Simultaneous sensing matrix and sparsifying dictionary optimization. *IEEE Trans. Image Proc.* **18**(7), 1395–1408 (2009)
23. D. Lee, MIMO-OFDM channel estimation via block stagewise orthogonal matching pursuit. *IEEE Commun. Lett.* **20**(10), 2115–2118 (2016)
24. I. Barhumi, G. Leus, M. Moonen, Optimal training design for MIMO OFDM systems in mobile wireless channels. *IEEE Trans. Signal Proc.* **51**(6), 1615–1624 (2003)

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)