Pilot Decontamination in Multi-Cell Massive MIMO Systems via Combining Semi-Blind Channel Estimation With Pilot Assignment

CHENG HU [1], HONG WANG [1], AND RONGFANG SONG [1,2], (Member, IEEE)
[1]College of Communication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China
[2]Key Laboratory of Broadband Wireless Communication and Sensor Network Technology, Nanjing University of Posts and Telecommunications, Ministry of Education, Nanjing 210003, China

Corresponding author: Rongfang Song (songrf@njupt.edu.cn)

ABSTRACT In the multi-cell massive multiple-input multiple-output (MIMO) systems, the mitigation of pilot contamination is an important issue in the semi-blind channel estimations. For the pilot decontamination schemes, the performance of the semi-blind based estimation approaches is constrained by the ratio of the interference signal power to the target signal power. In order to eliminate pilot contamination effectively, the eigenvalues of the target signals should be separated from those of the interference signals. In this paper, we propose a novel uplink channel estimation for mitigating pilot contamination in time division duplex massive MIMO systems, which combines pilot assignments with semi-blind channel estimation methods. In order to reduce the search complexity for the optimal pilot assignment under the condition of a large number of users per cell, we proposed a sector-based pilot assignment method, including inter-sector pilot assignment and intra-sector pilot optimization. Simulation results verify that the joint pilot allocation and semi-blind channel estimation method is capable of improving the system achievable rates and the normalized mean square error performance.

INDEX TERMS Massive MIMO, pilot decontamination, semi-blind channel estimation, pilot assignment.

I. INTRODUCTION

By employing a large number of service antennas at the base stations (BSs), massive multiple-input multiple-output (MIMO) technique has the capability of increasing throughputs and bringing other advantages, such as small latency, energy saving, and low spectrum overhead [1], [2]. In order to make full use of the benefits of massive MIMO, channel state information (CSI) is a necessity for system optimization. As a result, channel estimation is a hot topic in the research on massive MIMO. Currently, a lot of studies on channel estimation focus on single cell massive MIMO system [3]. Given accurate CSI by reverse-link channel estimation [4], forward-link CSI can be obtained through channel reciprocity in time-division duplex(TDD) massive MIMO systems. In

The associate editor coordinating the review of this manuscript and approving it for publication was A. Taufiq Asyhari [ID].

this paper, we focus on a more practical scenario, i.e., multi-cell multi-user (MCMU) massive MIMO systems. However, due to the reuse of pilot sequences across the neighbouring cells, pilot contamination is a vital issue in the channel estimation of massive MIMO systems. Intra-cell pilot contamination is removed by allocating the orthogonal pilot sequences to the users within the same cell, whereas the non-orthogonal pilot sequences are adopted across the different cells, which bring inter-cell interference (named as pilot contamination) in channel estimation. Due to pilot contamination, the performance of a plenty of wireless data services is degraded. For example, pilot contamination caused by an increasing number of smart devices leads to a limited achievable rate and low spectrum efficiency for users, especially at the cell edge [5], [6].

In order to improve the system performance, the mitigation of pilot contamination should be addressed. In the existing

literatures, pilot contamination is eliminated through precoding [7] and power allocation [8] at the BS. It is noted that the literatures [7], [8] are on the focus of pilot-based channel estimation. By leveraging the optimization of pilot assignment, pilot contamination can also be reduced substantially. It is known that semi-blind channel estimation is an effective technique in the massive MIMO system due to the low spectrum overhead. Thus, in this paper, we aim at reducing pilot contamination through the combination of semi-blind channel estimation with pilot assignment. In the following, the literatures on the semi-blind channel estimation and pilot assignment are reviewed carefully.

A. SEMI-BLIND CHANNEL ESTIMATION

To overcome pilot contamination, a couple of schemes are proposed for the pilot-based approaches and subspace-based approaches, respectively [9]. In [3], a linear pilot method is proposed for channel estimation. Chu sequences with perfect auto-correlation property and corresponding optimal pilot assignment schemes are designed in [10], [11] to minimise the effect of pilot contamination. By using a pilot-based scheme, a maximum a-posteriori (MAP) method extracts the interference eigenvectors together with target signals eigenvectors, and then obtains the estimated channel matrix through combinatorial optimization [12], [13]. Superimposed arrangement methods of the pilots in [14] superimpose the pilot and uplink data in original uplink data slot, and perform power allocation on the pilot and uplink data to obtain a better system capacity. Superimposed arrangement methods of the pilots are further extended to multi-path channels in [15]. Besides, a classical greedy pilot assignment is adopted to minimize the total mean squared error (MSE) and the search complexity in [16]–[18]. However, it still has a high computation complexity in the pilot-based schemes.

For the subspace-based schemes, statistics, such as variances and kurtosis [19]–[23], are used to separate the target signals from the noise and interferences. Therefore, the subspace-based approaches are interpreted as the semi-blind methods in general, which can be divide into two categories: singular value decomposition (SVD)-based method and eigenvalue decomposition (EVD)-based method. By using the eigenvalue distribution of the covariance matrix of the received signals, the SVD-based methods mainly depict the separation characteristics of the eigenvalues of the target signals and interference signals [24]–[27]. In the implementation of the real-world channel estimation, SVD-based methods directly estimate the symmetric signal subspace channel matrix, whereas the EVD-based methods are used to calculate the exact CSI.

By extrapolating the symmetric signal subspace channel matrix, the EVD-based channel estimation methods are designed to calculate small-scale fading channel matrix in [28]–[32]. In the ideal case of massive MIMO systems, [28] points out that the small scale fading matrix is the product of target signal eigenvectors matrix and diagonal ambiguity matrix. When the number of received antennas tends to

infinity in massive MIMO systems, the separation characteristics of the eigenvalues of the target signals and interference signals depend on three parameters: the number of interfering cells, the ratio of the user numbers in each cell to the coherent symbol length, and the ratio of the received interference power to the desired signal power [33], [34]. For a given wireless system, the number of interfering cells and the ratio of the user numbers in each cell to the coherent symbol length are determined accordingly. Thus, in order to improve the separation performance, the ratio of the received interference power to the desired signal power should be well designed. In [25], [33], a power-controlled hand-off is developed. However, the advantage of inter-cell coordination is ignored in [25], [33], which plays a vital role in the improvement of the separation performance.

In order to utilize the benefits of the pilot-based and semi-blind channel estimation methods, we investigate the combination of pilot assignment and semi-blind channel estimation methods in this paper.

B. PILOT ASSIGNMENT

In contrast to the shifting of pilot locations in frames dynamically [35], pilot assignment methods make the interfering users far away from the target user with the same pilot sequences in the pilot-based approachs [36]–[39]. Basically, the computation complexity of random pilot assignment (R-PA) is $\mathcal{O}(1)$. However, the effect of pilot contamination is serious. Exhaustive search pilot assignment (ES-PA) is capable of maximizing the minimum uplink signal-to-interference-plus-noise-ratio (SINR) by traversing all the possible pilot sequences, but ES-PA has an enormous high complexity of $\mathcal{O}((K!)^L)$, where K is the number of active users per cell and L is the number of cells.

In addition, the edge-weighted interference graph (EWIG) and graph coloring (GC) based pilot assignments are designed in [37], [38], which make full use of the average value of the mutual interference across the users and sequentially assign pilot sequences to the users in each cell. In [39], a deep learning-based pilot assignment (DL-PA) utilizes a deep multi-layer perceptron (MLP) to train the neural network by learning the relationship between the pilot assignment schemes of ES-PA and the location patterns of the users. However, in practice, DL-PA requires a large amount of ES-PA train data. Thus, it is difficult to pick out the pilot sequences for maximizing the minimum uplink SINR when the number of users in each cell is greater than 4. In order to enhance the minimum uplink SINR, a pilot assignment is developed in [36], in which the pilot sequence with the highest inter-cell interference is allocated to the user with the best channel quality in the target cell. In semi-blind channel estimation, the minimum uplink SINR is constrained by the largest pilot contamination in the massive MIMO system. If separation condition derived in [25], [33] is satisfied, pilot contamination can be eliminated.

On the other hand, a location aware pilot assignment (LA-PA) is considered in heterogeneous cellular

networks (HetNet) in [40]. Macro-users in the same macro-cell are assigned to different sectors, which are assigned the pilot sequences with a high priority. Then, small cell users in the central group are assigned pilot sequences by minimizing inter-tier interference. At last, the rest macro-users and small cell users match the residual pilot sequences exhaustively. Therefore, due to a huge number of macro-users and small cell users in HetNet, partitioning is proposed to reduce complexity via avoiding the global traversal. Motivated by this concept, sector-based pilot assignment is integrated into the semi-blind channel estimation in the MCMU massive MIMO systems in this paper.

C. OUR CONTRIBUTIONS

In this paper, the combination of sector-based smart pilot assignment and semi-blind channel estimation is investigated in the MCMU massive MIMO systems. In order to mitigate the pilot contamination caused by the overlap of target signal eigenvalue bulks and interference signal eigenvalue bulks, we propose a new pilot assignment scheme for semi-blind channel estimation method. The contributions are listed as follows:

- Firstly, the semi-blind channel estimation and pilot decontamination analysis are presented in the MCMU massive MIMO systems. By considering the influence of the location distributions of the users, the relationship between the propagation distance and the ratio of the interference signal power to the target signal power is derived. When the ratio of the interference to the target signal power is larger than a threshold, there will be an overlap between the eigenvalue bulks of the target signal and those of the interference in the EVD-based semi-blind method, which will further lead to pilot contamination.

- Secondly, a sector-based pilot assignment is proposed to reduce the interference on the pilot sequences by maximizing the distance between the users with the same pilot. In the sector-based pilot assignment, each cell is divided into multiple sectors. Then, the implementation of the proposed sector-based scheme includes two steps, i.e., inter-sector pilot assignment is implemented, and then intra-sector pilot assignment is refined by the optimization method. Compared to the exhaustive search method, the computational complexity of the proposed sector-based scheme is reduced substantially.

- Thirdly, the combination of sector-based smart pilot assignment and semi-blind channel estimation is capable of minimizing the upper bound of the ratio of the interference to the desired signal. It is shown by simulations that the combination of the proposed pilot assignment and semi-blind channel estimation method outperforms the existing schemes. To be more specific, the proposed scheme is able to improve average uplink achievable rate and reduce the normalized mean square error (NMSE).

The rest of this paper is organized as follows: in Section II, a MCMU-MIMO system model is introduced. In Section III, a semi-blind channel estimation method and pilot contamination elimination analysis are respectively presented. In Section IV, a new sector-based pilot sequences assignment is proposed for inter-sector pilot assignments and intra-sector pilot optimization. In Section V, we give numerical results of NMSE and the average uplink achievable rate per user under the different conditions. Finally, in Section VI, this paper is concluded.

Throughout the paper, we adopt the following notations: lowercase and uppercase boldface letters denote vectors and matrices, respectively; $\mathbb{E}\{.\}$ is the expectation operator; $\text{diag}\{t_1, \ldots, t_N\}$ denotes a diagonal matrix with the n-th diagonal element t_n; $\mathbf{X}^T$ and $\mathbf{X}^H$ are the transpose and complex-conjugate transpose of matrix $\mathbf{X}$, respectively; $\mathbb{C}^{T_1 \times T_2}$ denotes a complex matrix with T_1 rows and T_2 columns; $\mathbf{I}_K$ is a $K \times K$ identity matrix.

II. MCMU-MIMO SYSTEM MODEL

In this paper, the uplink of a multi-cell massive MIMO scenario is considered, where a hexagonal cell is surrounded by L interfering cells. Hereafter, for the convenience of expressions, the index 0 in the subscript represents the target cell, while the indices for the interfering cells are larger than 0. Each BS equipped with M antennas is located at the center of each hexagonal cell. The radius of each hexagonal cell is R_C. In each cell, K single-antenna users are randomly distributed in the BS coverage and are served by their BSs in TDD mode. To make full use of spectrum resource, the frequency reuse factor is 1, i.e., all the cells share the same frequency band. Besides, a block fading channel is assumed, i.e., the channel keeps unchanged during one transmission block but varies from one block to another. The system architecture is illustrated in Fig. 1.
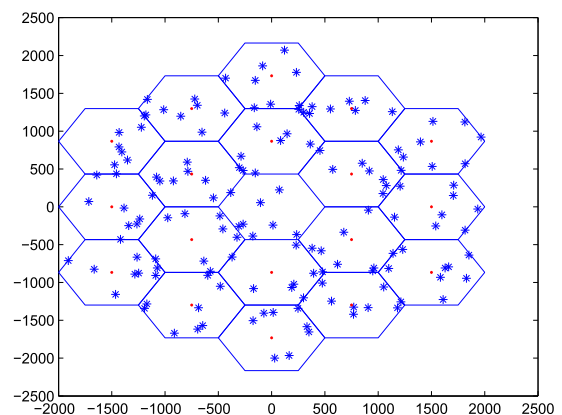

FIGURE 1. An illustration of hexagonal cellular network architecture with $L = 18$ and $K = 8$. Red spots represent the BSs; Blue stars represent the users.

It is assumed that C symbols are transmitted during one transmission block. In each block, both the pilot signals and information data are transmitted. Hereafter, unless otherwise

specified, the superscripts "p" and "d" denote the pilot signals and data signals, respectively. Then, the signals received at the target BS, including pilot and data signals, can be expressed as

$$\mathbf{R}_0 = \sqrt{p_u}\mathbf{H}_{00}\mathbf{S}_0 + \sqrt{p_u}\sum_{i=1}^{L}\mathbf{H}_{0i}\mathbf{S}_i + \mathbf{Z}_0, \qquad (1)$$

where $\mathbf{R}_0 = [\mathbf{R}_0^p, \mathbf{R}_0^d]$, $\mathbf{Z}_0 = [\mathbf{Z}_0^p, \mathbf{Z}_0^d]$, $\mathbf{S}_l = [\mathbf{S}_l^p, \mathbf{S}_l^d]$, $\mathbf{R}_0^p \in \mathbb{C}^{M \times N_p}$ is the matrix of received pilot signals in the desired cell, $\mathbf{Y}_0^d \in \mathbb{C}^{M \times N_d}$ is the matrix of received data symbols with $N_p + N_d \leq C$, $\mathbf{Z}_0^p \in \mathbb{C}^{M \times N_p}$ and $\mathbf{Z}_0^d \in \mathbb{C}^{M \times N_d}$ are the matrices of the additive white noise on the received pilot and data signals, respectively, whose elements following the complex Gaussian distribution with zero mean and unit variance, i.e., $\mathcal{CN}(0, 1)$, $\mathbf{S}_l^d \in \mathbb{C}^{K \times N_d}$ is the matrix of the transmitted data of K users in the lth cell for $l = 0, 1, \dots, L$, $\mathbf{S}_l^p \in \mathbb{C}^{K \times N_p}$ is the matrix of the transmitted pilot sequences of K users in the lth cell, p_u is the average transmit power allocated to each user, and $\mathbf{H}_{li}$ is the $M \times K$ channel matrix between the lth BS and the K users located in the ith cell.

The pilot sequences $\mathbf{S}_l^p$ can be further expressed as

$$\mathbf{S}_l^p = \sqrt{N_p}[\mathbf{s}_{l1}, \mathbf{s}_{l2}, \dots, \mathbf{s}_{lK}]^T = \sqrt{N_p}\Pi\Phi,$$

where $\Phi = [\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_{N_p}]^T \in \mathbb{C}^{N_p \times N_p}$ is the matrix of the total pilot sequence sets with $\Phi\Phi^H = \mathbf{I}_{N_p}$, and $\Pi \in \{0, 1\}^{K \times N_p}$ is the pilot assignment permutation matrix with $K \leq N_p$. To be more specific, if the i-th pilot $\mathbf{q}_i$ is assigned to the k-th user in the cell, the (k, i)-th element of Π is 1, and the rest elements in the k-th row of Π are set to be 0. In order to improve the performance of channel estimation, we will investigate how to assign the pilot sequences in Section IV.

The channel matrices are characterized by the large scale fading matrices and small scale fading matrices, where large scale fading coefficients consists of path loss and shadow fading. The channel matrix between BS l and the users in the i-th cell can be expressed as

$$\begin{cases} \mathbf{H}_{li} = \mathbf{G}_{li} \times \mathbf{D}_{li}, \\ \mathbf{D}_{li} = \text{diag}\{\sqrt{\beta_{li,1}}, \sqrt{\beta_{li,2}}, \dots, \sqrt{\beta_{li,K}}\}, \end{cases} \qquad (2)$$

where $\mathbf{H}_{li}$ is the small scale fading matrix with the (m, k)-th element $h_{li}(m, k)$, and the diagonal matrix $\mathbf{D}_{li}$ represents the large scale fading matrix with the k-th diagonal element $\sqrt{\beta_{li,k}}$, which is constant in C symbol intervals and independent of the receiving antenna index. Accordingly, for the channel matrix $\mathbf{G}_{li}$, the element in mth row and kth column $[\mathbf{G}_{li}]_{m,k} = g_{li}(m, k)$ is the channel coefficient between the m-th antenna of the lth BS and the kth user in the ith cell with $g_{li}(m, k) = h_{li}(m, k)\sqrt{\beta_{lik}}$. The expression of large scale fading coefficient is given as

$$\beta_{li,k} = Z_{li,k}/r_{li,k}^{\gamma},$$

where $Z_{li,k}, r_{li,k}, \gamma$ are the shadowing fading coefficients between BS l and user k in the i-th cell, the propagation distance between BS l and user k in the i-th cell, and the

path loss exponent, respectively. In addition, $10 \log_{10}(Z_{lik}) \sim \mathcal{CN}(0, \sigma_{shadow})$, where σ_{shadow} is usually considered a non-negative constant. Besides, $h_{li}(m, k)$ denotes the small scale fading coefficient from the kth user in the ith cell to the mth antenna of the lth BS, which follows an independent identically distribution (i.i.d) and is a circularly-symmetric Gaussian random variable with mean zero and variance one, i.e., $h_{li}(m, k) \sim \mathcal{CN}(0, 1)$.

In wireless communications, the change of the small scale fading coefficient $h_{li}(m, k)$ is much faster than the larger scale fading coefficient $\beta_{li,k}$. Thus, it is quite challenge to estimate the information of small scale fading matrix, which is also the focus of the existing literatures on this topic. In the following, we focus on the channel estimation by using the semi-blind channel estimation. It aims at estimating the small-scale channel $\mathbf{G}_{li}$ based on the received signals $\mathbf{R}_0$ and the assigned pilot sequences.

III. ANALYSIS OF SEMI-BLIND CHANNEL ESTIMATION
A. SEMI-BLIND CHANNEL ESTIMATION METHODS

In semi-blind channel estimation, the EVD-based channel estimation method is adopted in massive MIMO systems, which requires the pilot sequences and the covariance matrix of the received signals. In this section, the EVD-based semi-blind channel estimation includes two steps: 1) the small scale fading matrix is estimated by using pilot information; 2) the pilot based estimation small scale fading matrix is refined by the covariance matrix of the received data signals. The detailed processes are implemented as follows.

Based on the pilot assignment information and the received pilot signals, the estimated small scale fading channel matrix $\hat{\mathbf{H}}^p$ is calculated as

$$\hat{\mathbf{G}}^p = \frac{1}{\sqrt{N_p}}\mathbf{Y}_0^p\Phi^H\Pi. \qquad (3)$$

The estimated covariance matrix of the received data signals can be expressed as

$$\hat{\mathbf{R}}_y = \frac{1}{N_d}\mathbf{Y}_0^d[\mathbf{Y}_0^d]^H. \qquad (4)$$

Then, the EVD of $\hat{\mathbf{R}}_y$ is carried out as follows:

$$\hat{\mathbf{R}}_y = [\hat{\mathbf{V}}_S, \hat{\mathbf{V}}_I, \hat{\mathbf{V}}_N]\text{diag}\{\Lambda_S, \Lambda_I, \Lambda_N\}[\hat{\mathbf{V}}_S, \hat{\mathbf{V}}_I, \hat{\mathbf{V}}_N]^H, \quad (5)$$

where $\Lambda_S \in \mathbb{C}^{K \times K}, \Lambda_I \in \mathbb{C}^{LK \times LK}$ and $\Lambda_N \in \mathbb{C}^{(M-(L+1)\times K)\times(M-(L+1)\times K)}$ are the eigenvalue matrices corresponding to the desired signals, pilot contamination caused by the interfering cells, and additive noise, respectively, $\hat{\mathbf{V}}_S \in \mathbb{C}^{M \times K}, \hat{\mathbf{V}}_I \in \mathbb{C}^{M \times LK}$, and $\hat{\mathbf{V}}_N \in \mathbb{C}^{M \times (M-(L+1)\times K)}$ are the eigenvector matrices associated with Λ_S, Λ_I, and Λ_N, respectively, and the eigenvectors in $\hat{\mathbf{V}}_S$ associated with the largest K eigenvalues are extracted as the estimated signal subspace matrix.

At last, the channel estimation formula can be further derived as

$$\hat{\mathbf{G}}_{00} = \hat{\mathbf{V}}_S\hat{\mathbf{V}}_S^H\hat{\mathbf{G}}^p. \qquad (6)$$

where $\hat{\mathbf{G}}_{00}$ is the estimated small scale fading matrix in the center cell. The channel estimation in other cells can be obtained in the same way. In (6), $\hat{\mathbf{V}}_S^H$ is leveraged to eliminate pilot contamination and partial noise from $\hat{\mathbf{G}}^p$, and $\hat{\mathbf{V}}_S$ aims to extrapolate the symmetric signal subspace channel matrix into the estimated small scale fading matrix $\hat{\mathbf{G}}_{00}$.

B. PILOT CONTAMINATION ELIMINATION ANALYSIS

Due to the reuse of pilot sequences in the neighbouring cells, pilot contamination is caused in the multi-cell massive MIMO scenario. It is shown that the influence of pilot contamination is dependent on the interference signal strengthes stemming from the neighbouring cells. It is known that the interference power levels are related to the distribution of interfering users.

In this subsection, we discuss two metrics: 1) the relationship between $\alpha = \frac{I}{P}$ and $\frac{K}{C}$; 2) the relationship between $\frac{I}{P}$ and the propagation distance, where I is the interference power caused by the users in the adjacent cells, P is the desired signal power for user in the target cell, $\alpha = \frac{I}{P}$ is defined as the ratio of the interference power to the received desired signal power, and $\frac{K}{C}$ is defined as the ratio of user numbers per cell to the coherent symbol length.

1) RELATIONSHIP BETWEEN $\alpha = \frac{I}{P}$ AND $\frac{K}{C}$

In [25], the influence of the location distributions of the users in the target cell and the interfering users in the neighbouring cells is not considered. In this case, the ratios of the average received interference power to the average received signal power for all the target users in the desired cell are the same. For massive MIMO system, it has been shown in [25] that the relationship between $\alpha = \frac{I}{P}$ and $\frac{K}{C}$ should satisfy the following constraint, i.e.,

$$\frac{K}{C} \leq \frac{(1-\alpha)^2(L\alpha^2+3(L+1)\alpha+1-2(1+\alpha)\sqrt{3L\alpha})}{(L\alpha^2-1)(L\alpha^2+6(L-1)\alpha-1)+(9L^2-2L+9)\alpha^2}. \quad (7)$$

Fig. 2 plots the relationship between $\frac{K}{C}$ and α for $L = 4$ described in (7). For other L's, the behaviors of the curves are similar. It is demonstrated that $\frac{K}{C}$ is a decreasing function
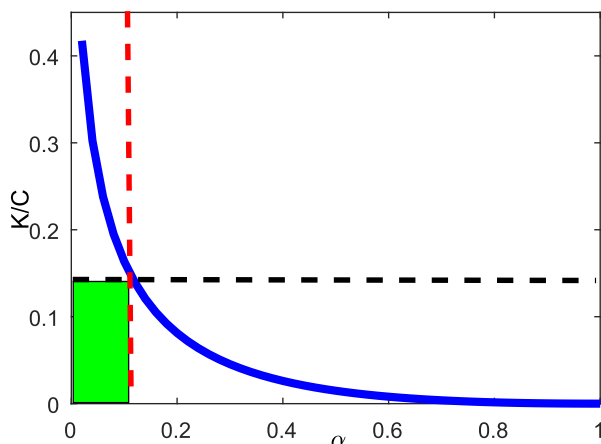

FIGURE 2. The relationship between $\frac{K}{C}$ and α with $L = 4$.

of α. For a given wireless communication, $\frac{K}{C}$ is determined. In this case, in order to satisfy the condition in (7), the ratio α should be less than a threshold. i.e., α should be located in the green area in Fig. 2.

However, when the user locations are taken into consideration, it will lead to a different interference power and received signal power ratio for each target user. For user k, the interference power and received signal power ratio is defined as α_k, where $\alpha_k = \frac{I_k}{P_k}$ for $k \in \{1, \ldots, K\}$. In this situation, for any user in the target cell, α_k should Follow the condition shown in (7). Then, there must have the relationship given as

$$\alpha_{th} \geq \max\{\alpha_1, \ldots, \alpha_K\}, \quad (8)$$

where α_{th} is the threshold to satisfy (7) for a given $\frac{K}{C}$. From (8), we can also obtain that the minimum uplink signal to interference ratio (SIR) $\frac{P_k}{I_k}$ should be larger than a threshold.

2) RELATIONSHIP BETWEEN $\frac{I}{P}$ AND THE PROPAGATION DISTANCE

Assuming that a user in the i-th interfering cell uses the same pilot sequence $\mathbf{q}_k$ as a user in the target cell, the slow fading coefficients of the interfering user and the target user are respectively expressed as

$$\beta_{00,\mathbf{q}_k} = \frac{Z_{00,\mathbf{q}_k}}{r_{00,\mathbf{q}_k}^\gamma}, \quad \beta_{0i,\mathbf{q}_k} = \frac{Z_{0i,\mathbf{q}_k}}{r_{0i,\mathbf{q}_k}^\gamma}, \quad (9)$$

Accordingly, the ratio $\frac{I}{P}$ is given by

$$\frac{I}{P} = \frac{\beta_{0i,\mathbf{q}_k}}{\beta_{00,\mathbf{q}_k}} = (\frac{Z_{0i,\mathbf{q}_k}}{Z_{00,\mathbf{q}_k}})(\frac{r_{00,\mathbf{q}_k}}{r_{0i,\mathbf{q}_k}})^\gamma. \quad (10)$$

Based on the distribution of $Z_{0i,\mathbf{q}_k}$, it can be derived that the logarithm of the shadowing fading ratio $\frac{Z_{0i,\mathbf{q}_k}}{Z_{00,\mathbf{q}_k}}$ follows a normal distribution, i.e., $10\log_{10}(\frac{Z_{0i,,\mathbf{q}_k}}{Z_{00,\mathbf{q}_k}}) \sim \mathcal{CN}(0, 2\sigma_{shadow})$. $\frac{r_{00,\mathbf{q}_k}}{r_{0i,\mathbf{q}_k}}$ is the ratio of the propagation distance between the target user and its serving BS to that between the interfering user using the same pilot sequence and the target BS. Since $\frac{I}{P}$ increases exponentially with $\frac{r_{00,\mathbf{q}_k}}{r_{0i,\mathbf{q}_k}}$ with the exponent γ, $\frac{I}{P}$ is dominant by $\frac{r_{00,\mathbf{q}_k}}{r_{0i,\mathbf{q}_k}}$. Thus, $\frac{I}{P}$ can be mitigated through pilots allocation. It has been shown in [35] that the interference stemming from the second-circle cells is easy to satisfy the condition (7), because of the propagation distances between the remaining outer cell users and the target BS are at least twice as the propagation distances between the users in the target cell and the target BS. As a result, in the following, we only focus on the pilot assignment for the first-circle cells in the semi-blind channel estimation.

IV. SECTOR-BASED PILOT ASSIGNMENT

The polar coordinates $(r_{00,\mathbf{q}_k}, \theta_{00,\mathbf{q}_k})$ are used to denote the location information of user k, where $r_{00,\mathbf{q}_k}$ is the distance between the user and its serving BS, and $\theta_{00,\mathbf{q}_k}$ is the angle relative to the horizontal line. It is known that the interference can be reduced by assigning the different pilot sequences to

the sector with the least interfering distances. Thus, by using the users location information, the distance between users who share same pilot sequences can be increased.

A. INTER-SECTOR PILOT ASSIGNMENT

According to the user location information, the coverage of each cell is divided into K sectors. Then, the pilot sequences are assigned to different sectors by using the criterion expressed as

$$
\eta_{lk}
$$
$$
= \begin{cases} \mathbf{q}_1, & \text{for } r_{ll,\mathbf{q}_1} < R_c/\sqrt{3} \\ \mathbf{q}_i, & \text{for } r_{ll,\mathbf{q}_i} \geq R_c/\sqrt{3}, \\ & \dfrac{2(i-2)\pi}{K-1} \leq \theta_{ll,\mathbf{q}_i} \leq \dfrac{2(i-1)\pi}{K-1}, \quad i = 2,\ldots,K, \end{cases}
$$
$$(11)$$

where η_{lk} is the pilot assignment for the kth sector of the lth cell. An illustration of sector division is shown in Fig. 3.
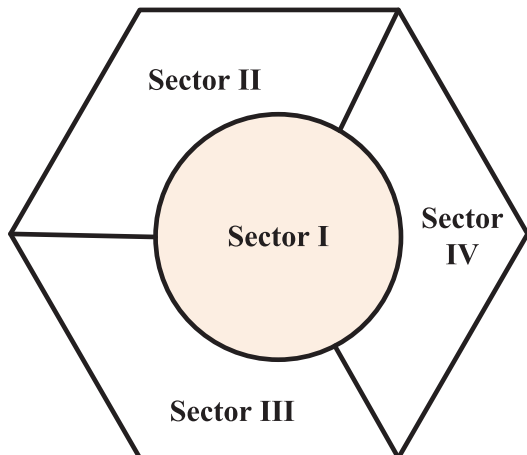
1) DISCUSSION ON INNER SECTOR RADIUS AND INTER-SECTOR PILOT ASSIGNMENT

Firstly, we consider the interference between the users of the inner sector $\mathbf{q}_1$. The distance between the interfering users in the sector with pilot $\mathbf{q}_1$ and the target BS is at least twice that of the desired user located in the target cell. Similar to the interfering users of the second-circle cells, the overlap of the eigenvalue bulks caused by the interfering users in the inner sector can be neglected by leveraging the EVD-based semi-blind channel estimation. For the outer sectors, i.e., Sector $2, \ldots, K$, the sectors of two neighbouring cells is paired and the $\frac{I}{P}$s for these two sectors are calculated accordingly. Based on the calculated information about $\frac{I}{P}$s, the pilot sequences are assigned to the sectors of two neighbouring cells aiming at minimizing $\frac{I}{P}$. Following this methodology, L cells are assigned sequentially until all the sectors of each cell are assigned the enough pilot sequences.

Next, we discuss the selection of inner sector radius. As mentioned above, the propagation distance of the interfering users located in the second-circle cells are at least twice as the radius of the target users. Considering the equivalence of the problem for inner sector, when the radius of the inner sector is about $\frac{1}{\sqrt{3}}$ times of the cell radius, the overlap of the eigenvalue bulks caused by the interfering users in the inner sector can be neglected.

After the inter-sector pilot assignment is complemented, the intra-sector is implemented in sequence. If the sector has only one user, pilot sequences are assigned to the corresponding user in the sector. Then, the pilot sequences assigned to the blank sectors are recycled to the sectors who have multiple users. If there are multiple users in a sector, pilot sequences in the sector are internally assigned so that each user is allocated a pilot sequence. In order to assign the pilot sequence effectively, pilot sequences are optimized in the sector with multiple users by using the smart pilot allocation scheme, which will be presented in the next subsection.

B. INTRA-SECTOR PILOT ASSIGNMENT BY USING PROPOSED PILOT ASSIGNMENT

The pilot sequences of the multi-user sectors are randomly formed in the above subsection. However, this initial assignment method is far from the optimal solution. In this subsection, we focus on the problem: how to optimize the pilot assignment to each user in the multi-user sector.

1) PROBLEM FORMULATION

We consider the pilot assignments for multiple users located in the same sector. The pilot assignment for the target lth cell is specifically dissected from other cells, and thus pilot assignments for all the cells are independently managed by their own BSs. For all K users in lth cell, their assigned pilot sequences η_l can be represented as $[\eta_{l1}, \eta_{l2}, \ldots, \eta_{lK}]$, e.g., for $K = 8$, one assignment scheme may be $\eta_l = [\mathbf{q}_1, \mathbf{q}_7, \mathbf{q}_3, \mathbf{q}_6, \mathbf{q}_5, \mathbf{q}_2, \mathbf{q}_8, \mathbf{q}_4]$. In the conventional exhaustive search method, the search space is $K!$ for each cell. By using sector-based pilot assignment, the number of the candidate solutions is $K_{set} = (K_1!) \times (K_2!) \times (K_3!) \times \ldots \times (K_S!)$ with $\sum_{i=1}^{S} K_i = K$, where S is the number of sectors having the active users. In $\{K_s\}$, there maybe exist $K_s = 1$ for some s. Compared with the conventional exhaustive search method with non-sector, the search space of the proposed sector-based scheme is reduced substantially.

Via pilot assignments, the objective is to maximize the minimum average SIR in each cell. In massive MIMO systems, as the M received base station antennas approaches infinity, the optimization problem in the l-th cell can be formulated as

$$
\mathbf{P} : \underset{\{\mathbf{F}_{ls}\}}{\mathrm{MAX}} : \underset{\forall k}{\min} \left\{ \frac{\beta_{ll,k}^2}{\sum_{i=0, i \neq l}^{L} \beta_{li,k}^2} \right\}, \tag{12}
$$

where $\{\mathbf{F}_{ls}; s = 1, 2, \ldots, K_{set}\}$ denotes all the candidates of the pilot assignments, and $\mathbf{F}_{ls} = [\eta_{l1}, \eta_{l2}, \ldots, \eta_{lK}]$ refers to a particular pilot assignment scheme in lth cell.

2) PROPOSED PILOT ASSIGNMENT

Although the computational complexity of the exhaustive search is smaller than that of non-sector, the number of all pilot assignments, K_{set}, is still huge when the number of users in each cell is large. On the other hand, when the greedy method is directly used, it may fall into the trap of local optimal solution. Therefore, it requires an effective solution that is capable of reducing the search complexity and preventing the obtained solution from falling into the local optimal.

For a specific pilot sequence $\mathbf{q}_k$, the objective function of the problem **P** can be divided into the signal power of the target lth cell and the interference power stemming from the interfering cells, which can be expressed as

$$\alpha_{\mathbf{q}_k} = \beta_{ll,\mathbf{q}_k}^2, \quad \gamma_{\mathbf{q}_k} = \sum_{i=0, i \neq l}^{L} \beta_{li,\mathbf{q}_k}^2. \quad (13)$$

Without loss of generality, there exists a pilot assignment scheme so that $\{\alpha_{\phi_k}\}$s are sorted in descending order. That is,

$$\mathbf{F}_{lq} = [\mathbf{q}_{f_q^1}, \mathbf{q}_{f_q^2}, \ldots, \mathbf{q}_{f_q^K}], \quad (14)$$

$$\alpha_{\mathbf{q}_{f_q^1}} \geq \alpha_{\mathbf{q}_{f_q^2}} \geq \ldots \geq \alpha_{\mathbf{q}_{f_q^K}}. \quad (15)$$

Similarly, there also exists a pilot assignment scheme so that $\{\gamma_{\phi_k}\}$s are sorted in descending order. That is,

$$\mathbf{F}_{lp} = [\mathbf{q}_{f_p^1}, \mathbf{q}_{f_p^2}, \ldots, \mathbf{q}_{f_p^K}], \quad (16)$$

$$\gamma_{\mathbf{q}_{f_p^1}} \geq \gamma_{\mathbf{q}_{f_p^2}} \geq \ldots \geq \gamma_{\mathbf{q}_{f_p^K}}. \quad (17)$$

To maximize the objective function in (12), the pilot sequence $\mathbf{q}_{f_p^1}$ with a large pilot contamination should not be assigned to the target user who has a low channel gain, whereas we should assign $\mathbf{q}_{f_q^1}$ to the user in the sector which has the best channel quality. Therefore, the optimization schemes of pilot assignment can be expressed as

$$\eta_{lf_q^i} = \mathbf{q}_{f_p^i}, \quad f_q^i \leq K_m, \ i \in 1, 2, \ldots, K. \quad (18)$$

where K_m is the number of the users in the sector with multiple users. Since the sectors have already been allocated, only the K_m users in the m-th sector participate in optimization, whereas the pilot sequence of the sector with a single user sectors remains to be fixed.

To apply the proposed method to the whole system with $L + 1$ cells, the above method is used to optimize the pilot assignments of the multi-user sectors for all $L + 1$ cells in sequence. Because the pilot assignments for the $L + 1$ cells are coupled, the proposed optimization process should be iterated for many times until the solution converges. Let $itev_{num}$ denote the maximum number of iteration. The procedures of our proposed pilot assignment method are summarized in Algorithm 1.

It is worth mentioning that the iterations are not required in the semi-blind channel estimation. Besides, in this paper, the pilot assignment scheme is dependent on user location information but is independent of semi-blind channel estimation. Besides, The convergence speed of the pilot assignment

---

Algorithm 1 Proposed Sector-Based Smart Pilot Assignment Scheme

Input: System parameters: K, L, and α_{th};
1: Users location information: $(r_{li,k}, \theta_{li,k})$;
2: Large-scale fading coefficients: $\beta_{li,k}$.
Output:
3: Inter-sector pilot assignment: η_{lk}.
4: Intra-sector pilot allocation,
5: Calculate the initial values:$\eta_{l,\mathbf{q}_k}$.
6: **while** $s < itev_{num}$ **do**
7: $l = 0$,
8: **while** $l \leq L$ **do**
9: Calculate $\alpha_{\mathbf{q}_k}$, $\gamma_{\mathbf{q}_k}$,
10: Sort $\alpha_{\mathbf{q}_k}$, $\gamma_{\mathbf{q}_k}$,
11: $\mathbf{F}_{lq} = [\mathbf{q}_{f_q^1}, \mathbf{q}_{f_q^2}, \ldots, \mathbf{q}_{f_q^K}]$,
12: $\mathbf{F}_{lq} = [\mathbf{q}_{f_q^1}, \mathbf{q}_{f_q^2}, \ldots, \mathbf{q}_{f_q^K}]$,
13: Optimized pilot assignment $\eta_{lf_q^i} = \mathbf{q}_{f_p^i}$.
14: **end while**
15: **end while**
16: Compare the optimized values to obtain the optimal pilot assignments η_{lk}.

---

algorithm is relatively fast. In the proposed scheme, the complexity of eigenvalues decomposition of the semi-blind estimation technique is O(M^3), where the quantity of M in the actual MIMO systems is within a tolerable range. Thus, the computational complexity is not very high.

V. NUMERICAL RESULTS

In this section, we consider a MCMU-MIMO system with $L + 1$ cells, including one center cell and L interfering cells, where each cell has K users equipped with a single antenna and a BS equipped with M antennas. Pilot sequences within the same cell are orthogonal to each other, and all cells reuse the same set of pilot sequences. The average received signal-to-noise ratio (SNR) for the k-th user in the j-th cell is $p_u\beta_{jj,k}$. The detailed parameters are shown in Table 1.

TABLE 1. Simulation parameters.

Number of interfering cells L	3, 6
Number of BS antennas M	$32 \leq M \leq 512$
Number of users in each cell K	$3 \leq K \leq 10$
Coherent symbol length C	100, 300, 1000
Radius of each hexagonal cell R_c	500 m
Radius of the center hole R_i	35 m
Path loss exponents γ	3.5
Shadowing fading coefficient σ_{shadow}	8dB
Signal to noise ratio SNR	25dB
Transmitted power per user	$P_u = SNR/M$.

A. THE STATISTICS OF LARGE SCALE FADING COEFFICIENT AND THE THRESHOLD OF $\frac{I}{P}$

Fig. 4 plots the cumulative distribution function (CDF) of the logarithmic large scale fading coefficients of the signal of
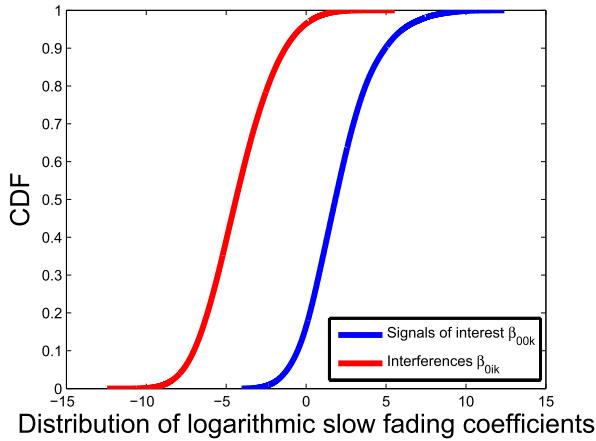
FIGURE 4. CDF of the logarithmic large scale fading coefficients with $\gamma = 3.5$ and $\sigma_{shadow} = 8$dB.

FIGURE 5. The threshold of $\frac{I}{P}$ to separate the target signal eigenvalue bulks and interference eigenvalue bulks with $\gamma = 3.5$ and $\sigma_{shadow} = 8$dB.

interest and the interference links. It is shown that the CDF of the interfering links is located at the left of the desired links. Via the simulation results, it can be obtained that the mean value of large scale fading coefficients of signals of interest is 309.385 and the mean value of large scale fading coefficients of interference link is 0.1925. The worst case is that the large scale fading coefficient of signals of interest is the minimum value, whereas the large scale fading coefficient of interference link achieves the maximum value. In simulations, the minimum value of the large scale fading coefficients of signals of interest is 0.0053, and the maximum value of the large scale fading coefficients of interference link is 323.138. However, when comparing the large scale fading coefficients between the target signals and interference links, it can be observed that the probability of the large scale fading coefficients of interference links is larger than the large scale fading coefficients of target signals is about 7.15%, which inevitably makes the eigenvectors of the desired signals be replaced by the eigenvectors of the interference signals in the semi-blind channel estimation.

In addition, according to eq.(7), when the number of the interfering cells L, the number of users K in each cell and the coherent symbol length C are given, the threshold α_{th} can be calculated accordingly. When the maximum value of $\{\frac{I_k}{P_k}\}$ is greater than α_{th}, the eigenvalue bulks of the target signal will overlap with those of the interference signals, and thus the situation of the eigenvector substitution described above will occur. Fig. 5 shows the threshold of $\frac{I}{P}$ in order to separate the target signal and interference eigenvalue bulks with $\gamma = 3.5$ and $\sigma_{shadow} = 8$ dB. It is shown that more interfering cells have a greater effect on the upper bound of $\frac{I}{P}$. Besides, the upper bound of $\frac{I}{P}$ is decreased with the number of users in each cell. When the coherent symbol length C is small, such as in the high-speed mobile application scenarios, it can only tolerate a small interference to desired power ratio. Therefore, by using an appropriate allocation method of pilot sequences, an increasing number of the target users meet the constraint condition on $\frac{I}{P}$ given in Subsection III-B. This
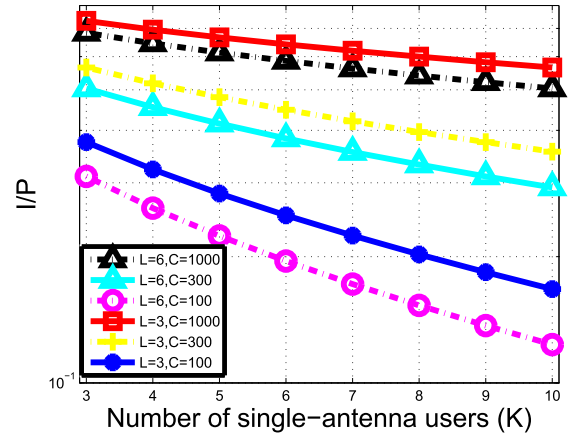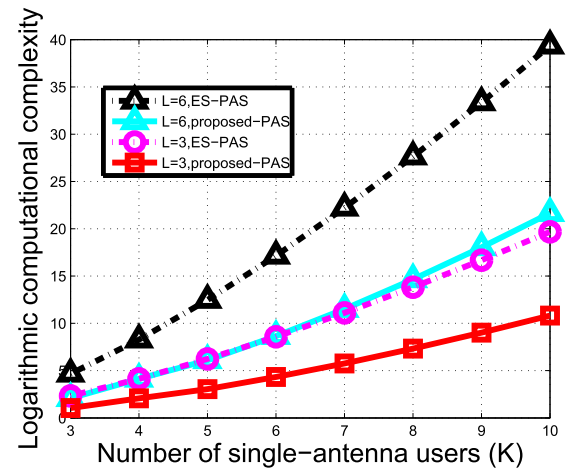

FIGURE 6. The logarithmic computational complexity against the number of single antenna users K. ES-PAS: exhaustive search pilot assignment scheme; Proposed-PAS: proposed pilot assignment scheme.

operation will reduce the overlap of eigenvalues, and thus the system capacity will be increased.

B. COMPLEXITY ANALYSIS

Fig. 6 plots the logarithmic computational complexity of the exhaustive search pilot assignment and the proposed pilot assignment. As the number of cells and single antenna users in each cell increases, the complexity of both the exhaustive search pilot assignment and the proposed pilot assignments increases exponentially. It is also shown that the computational complexity of the exhaustive search scheme is always larger than that of the proposed pilot assignment scheme. By using sector division, the maximum received interference power and target signal power ratio $\frac{I}{P}$ is reduced. In addition, the search complexity of traversing all feasible solutions is also reduced.

C. CAPACITY AND NMSE PERFORMANCE

Fig. 7 plots the average uplink achievable rate per user versus the number of BS antennas M. The average uplink achievable
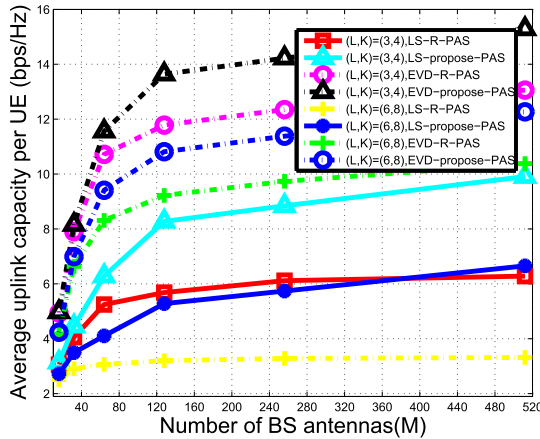
FIGURE 7. The average uplink achievable rate per user against the number of BS antennas *M*. LS-R-PAS: least square based with random pilot assignment scheme; LS-propose-PAS: least square based with the proposed pilot assignment scheme; EVD-R-PAS: EVD based with random pilot assignment scheme; EVD-propose-PAS: EVD based with the proposed pilot assignment scheme.

rate can be represented as

$$C_{ave}^{UL} = (1 - \mu_0)\mathbb{E}\left\{\log_2(1 + SINR_{ave}^{UL})\right\}$$

$$= (1 - \mu_0)\mathbb{E}\left\{\log_2(1 + \frac{\sum_{i=1}^{itev}\sum_{l=1}^{L}\sum_{k=1}^{K} SINR_{l,k}^{(i)}}{itev*L*K})\right\} \quad (19)$$

with

$$SINR_{l,k}^{(i)} = \frac{||\hat{\mathbf{H}}_{ll,k}^{(i)}||^2}{||\Delta\hat{\mathbf{H}}_{lm,k}^{(i)}||^2 + ||\Delta N_0||^2}, \quad (20)$$

where μ_0 evaluates the uplink pilot spectrum overheads, *itev* is the total iterative number in the simulation, $\hat{\mathbf{H}}_{ll,k}^{(i)}$, $\Delta\hat{\mathbf{H}}_{lm,k}^{(i)}$ and ΔN_0 are the estimated target signals, leaked interference signals, and noise signals in the *i*-th iteration, respectively. In practice, the parameter should be carefully designed to make a trade-off between pilot overhead and the error of channel estimation. In this paper, the design of pilot overhead is not our focus. Actually, we focus on the pilot assignment in the semi-blind channel estimation. For fairness, in the comparison of the channel capacity for different channel estimation methods, we set the pilot overhead to be a fixed number.

In Fig. 7, it is shown that, in the absence of the eigenvectors substitution described above, the average uplink achievable rates increase with the antenna numbers in the massive MIMO systems for all methods. When the number of interference cells *L* and the number of users *K* in each cell are considered, the average uplink achievable rate is reduced by 1.5-5 bps/Hz. By using the EVD-based semi-blind channel estimation method, the average uplink achievable rate is increased by 1.88-6.77 bps/Hz because the influence of $\Delta\hat{\mathbf{H}}_{lm,k}^{(i)}$ is reduced. Under all different conditions, the average uplink achievable rate of the proposed pilot assignment scheme is larger than that of the random pilot assignment scheme. It is because the pilot sequences are initialized by sectors to eliminate interference from other cells, and then

they are iteratively optimized in our proposed pilot assignment scheme. In addition, EVD-based semi-blind methods are used to reduce and eliminate interference and partial noise. Therefore, EVD-based channel estimation method achieves a larger average rate than that of least square (LS) based channel estimation method.
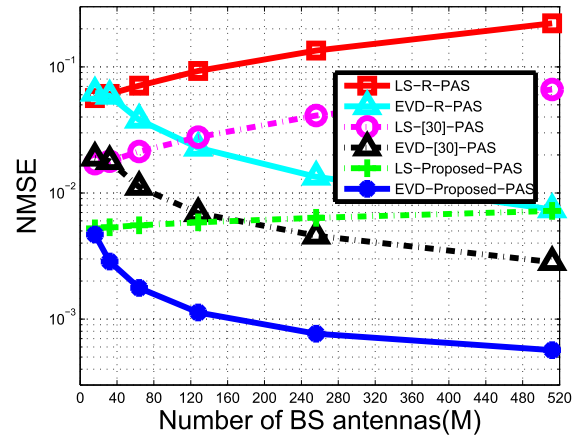

FIGURE 8. The NMSE versus the number of antennas *M* with signal perturbation with *K* = 4 and N_d = 300.

Fig. 8 describes the NMSE peformance versus the number of received antennas *M*. We consider the pilot-based methods and EVD-based semi-blind channel estimation methods with random pilot assignment schemes(R-PAS), pilot assignment schemes in [36] ([30]-PAS), and proposed pilot assignment schemes. Here, NMSE of the least square method can be derived as

$$NMSE_{LS} = \frac{\mathbb{E}\{||\hat{\mathbf{H}}^p - \mathbf{H}_{00}||^2\}}{\mathbb{E}\{||\mathbf{H}_{00}||^2\}}$$

$$= \frac{\mathbb{E}\{||\Delta_0||^2\}}{\mathbb{E}\{||\mathbf{H}_{00}||^2\}}$$

$$= \frac{M\sum_{k=1}^{K}\sum_{i=1}^{L}\beta_{0i,k} + \frac{MK}{p_u N_p}}{M\sum_{k=1}^{K}\beta_{00,k}}$$

$$= \frac{\sum_{k=1}^{K}\sum_{i=1}^{L}\beta_{0i,k} + \frac{K}{p_u N_p}}{\sum_{k=1}^{K}\beta_{00,k}}. \quad (21)$$

It can be observed that, with the increase of the received antenna numbers *M*, the NMSEs of LS methods are increased accordingly due to the decrease of p_u, whereas the NMSEs of EVD-based semi-blind channel estimation methods are gradually decreased. On one hand, the orthogonality of the eigenvectors increases with the increase of the received antenna numbers *M*. On the other hand, the increase of the distance between the eigenvalue distributions leads to the decrease of the error caused by the decrease of the probability of eigenvalues substitution. The terms $\{\beta_{0i,k}\}$ are reshaped by the different pilot assignment schemes, and thus the NMSEs of the two channel estimation methods are reduced at the same time. Furthermore, since the overlap probability of the target signal eigenvalue bulk and interference bulk in the proposed pilot assignment schemes is less than that in smart pilot assignment schemes, the substitution error is reduced.
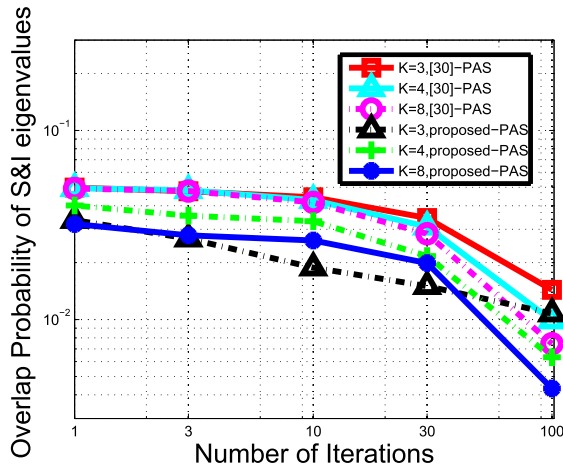
FIGURE 9. Overlap probability of the target signal eigenvalue bulks and interference signal eigenvalue bulks with $L = 7$ and $N_d = 1000$.

Fig. 9 shows the overlap probability of the target signal eigenvalue bulks and interference signal eigenvalue bulks. It can be seen that in the case of poor channel conditions, it takes several iterations in the optimization of pilot sequences to reduce the overlap probability of the target signal eigenvalue bulks and interference bulks. When the number of users K is large, the proposed pilot assignment schemes make the overlap probability of the target signal eigenvalue bulk and interference bulk become smaller after multiple iterations. In sector-based pilot sequence assignment, the initial value of proposed pilot assignment schemes is small and avoids the obtained solution is far from the optimal value. By the way, when the number of users K in each cell is 3, the overfitting will lead to a limited system performance improvement.

VI. CONCLUSION

In this paper, the semi-blind channel estimation in MCMU massive MIMO system is investigated. Considering the effects of the different large-scale fading coefficients for each user, the relationship between the ratio of the interference power to the received desired signal power and the propagation distance is analyzed. Based on the derived relationship, we propose a new joint signal estimation scheme in order to eliminate pilot contamination. In the proposed method, the sector-based pilot assignment scheme is integrated into EVD-based semi-blind method. Simulation experiments show that the combination of proposed pilot assignment with EVD-based semi-blind channel estimation method can reduce the overlap probability. As a result, by using the proposed scheme, the average uplink achievable rate is increased, and the system NMSE is reduced. Besides, the system performance of our proposed scheme is superior to that of other benchmark schemes.

REFERENCES

[1] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[2] H. Wang, R. Song, and S.-H. Leung, "Throughput analysis of interference alignment for a general centralized limited feedback model," *IEEE Trans. Veh. Technol.*, vol. 65, no. 10, pp. 8775–8781, Oct. 2016.

[3] M. Biguesh and A. B. Gershman, "MIMO channel estimation: Optimal training and tradeoffs between estimation techniques," in *Proc. IEEE Int. Conf. Commun.*, vol. 5, Jun. 2004, pp. 2658–2662.

[4] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[5] B. Gopalakrishnan and N. Jindal, "An analysis of pilot contamination on multi-user MIMO cellular systems with many antennas," in *Proc. IEEE 12th Int. Workshop Signal Process. Adv. Wireless Commun.*, Jun. 2011, pp. 381–385.

[6] X. Jia, M. Xie, M. Zhou, H. Zhu, and L. Yang, "Achievable uplink rate analysis for distributed massive MIMO systems with interference from adjacent cells," *China Commun.*, vol. 14, no. 5, pp. 112–123, May 2017.

[7] A. Ashikhmin and T. Marzetta, "Pilot contamination precoding in multi-cell large scale antenna systems," in *Proc. IEEE Int. Symp. Inf. Theory Proc.*, Jul. 2012, pp. 1137–1141.

[8] O. Saatlou, M. O. Ahmad, and M. N. S. Swamy, "Joint data and pilot power allocation for massive MU-MIMO downlink TDD systems," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 66, no. 3, pp. 512–516, Mar. 2019.

[9] O. Elijah, C. Y. Leow, T. A. Rahman, S. Nunoo, and S. Z. Iliya, "A comprehensive survey of pilot contamination in massive MIMO—5G system," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 905–923, 2nd Quart., 2016.

[10] J. Zhao, S. Ni, L. Yang, Z. Zhang, Y. Gong, and X. You, "Multiband cooperation for 5G HetNets: A promising network paradigm," *IEEE Veh. Technol. Mag.*, vol. 14, no. 4, pp. 85–93, Dec. 2019.

[11] J. Zhao, S. Ni, Y. Gong, and Q. Zhang, "Pilot contamination reduction in TDD-based massive MIMO systems," *IET Commun.*, vol. 13, no. 10, pp. 1425–1432, Jun. 2019.

[12] D. Neumann, A. Gruendinger, M. Joham, and W. Utschick, "Pilot coordination for large-scale multi-cell TDD systems," in *Proc. 18th Int. ITG Workshop Smart Antennas (WSA)*, Mar. 2014, pp. 1–6.

[13] D. Neumann, M. Joham, and W. Utschick, "Covariance matrix estimation in massive MIMO," *IEEE Signal Process. Lett.*, vol. 25, no. 6, pp. 863–867, Jun. 2018.

[14] K. Upadhya, S. A. Vorobyov, and M. Vehkapera, "Superimposed pilots are superior for mitigating pilot contamination in massive MIMO," *IEEE Trans. Signal Process.*, vol. 65, no. 11, pp. 2917–2932, Jun. 2017.

[15] M. M. Babar, N. Syed, and G. Sardar, "Massive-MIMO sparse uplink channel estimation using implicit training and compressed sensing," *Appl. Sci.*, vol. 7, no. 1, pp. 63–78, 2017.

[16] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 264–273, Feb. 2013.

[17] H. Yin, L. Cottatellucci, D. Gesbert, R. R. Muller, and G. He, "Robust pilot decontamination based on joint angle and power domain discrimination," *IEEE Trans. Signal Process.*, vol. 64, no. 11, pp. 2990–3003, Jun. 2016.

[18] H. Yin, D. Gesbert, M. C. Filippou, and Y. Liu, "Decontaminating pilots in massive MIMO systems," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2013, pp. 3170–3175.

[19] D. Hu, L. He, and X. Wang, "Semi-blind pilot decontamination for massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 1, pp. 525–536, Jan. 2016.

[20] N. Fatema, Y. Xiang, and I. Natgunanathan, "Analysis of a semi blind pilot decontamination method in massive MIMO," in *Proc. 27th Int. Telecommun. Netw. Appl. Conf. (ITNAC)*, Nov. 2017, pp. 1–6.

[21] M. Hosseinzadeh, H. Aghaeinia, and M. Kazemi, "Channel estimation improvement in massive MIMO systems based on pilot decontamination," in *Proc. IEEE Int. Conf. Adv. Netw. Telecommun. Syst. (ANTS)*, Dec. 2017, pp. 1–6.

[22] V. Zarzoso and P. Comon, "Robust independent component analysis by iterative maximization of the kurtosis contrast with algebraic optimal step size," *IEEE Trans. Neural Netw.*, vol. 21, no. 2, pp. 248–261, Feb. 2010.

[23] M. A. Sedaghat, A. Bereyhi, and R. R. Muller, "Least square error precoders for massive MIMO with signal constraints: Fundamental limits," *IEEE Trans. Wireless Commun.*, vol. 17, no. 1, pp. 667–679, Jan. 2018.

[24] R. R. Muller, "A random matrix model of communication via antenna arrays," *IEEE Trans. Inf. Theory*, vol. 48, no. 9, pp. 2495–2506, Sep. 2002.

[25] R. R. Muller, L. Cottatellucci, and M. Vehkapera, "Blind pilot decontamination," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 773–786, Oct. 2014.

[26] S. Asaad, A. Bereyhi, A. M. Rabiei, R. R. Muller, and R. F. Schaefer, "Optimal transmit antenna selection for massive MIMO wiretap channels," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 4, pp. 817–828, Apr. 2018.

[27] S. Asaad, A. M. Rabiei, and R. R. Muller, "Massive MIMO with antenna selection: Fundamental limits and applications," *IEEE Trans. Wireless Commun.*, vol. 17, no. 12, pp. 8502–8516, Dec. 2018.

[28] H. Q. Ngo and E. G. Larsson, "EVD-based channel estimation in multi-cell multiuser MIMO systems with very large antenna arrays," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2012, pp. 3249–3252.

[29] H. Q. Ngo and E. G. Larsson, "No downlink pilots are needed in TDD massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2921–2935, May 2017.

[30] B. Muquet, M. de Courville, and P. Duhamel, "Subspace-based blind and semi-blind channel estimation for OFDM systems," *IEEE Trans. Signal Process.*, vol. 50, no. 7, pp. 1699–1712, Jul. 2002.

[31] A. Hu, T. Lv, and Y. Lu, "Subspace-based semi-blind channel estimation for large-scale multi-cell multiuser MIMO systems," in *Proc. IEEE 77th Veh. Technol. Conf. (VTC Spring)*, Jun. 2013, pp. 1–5.

[32] T. Lv, S. Yang, and H. Gao, "Semi-blind channel estimation relying on optimum pilots designed for multi-cell large-scale MIMO systems," *IEEE Access*, vol. 4, pp. 1190–1204, 2016.

[33] R. R. Mueller, M. Vehkaperae, and L. Cottatellucci, "Blind pilot decontamination," in *Proc. 17th Int. ITG Workshop Smart Antennas (WSA)*, pp. 1–6, Mar. 2013.

[34] L. Cottatellucci, R. R. Muller, and M. Vehkapera, "Analysis of pilot decontamination based on power control," in *Proc. IEEE 77th Veh. Technol. Conf. (VTC Spring)*, Jun. 2013, pp. 1–5.

[35] Y. Xu and Y. Wu, "An approach of pilot contamination reduction based on power control and orthogonal identification," in *Proc. IEEE 3rd Int. Conf. Cloud Comput. Big Data Anal. (ICCCBDA)*, Apr. 2018, pp. 476–480.

[36] X. Zhu, Z. Wang, L. Dai, and C. Qian, "Smart pilot assignment for massive MIMO," *IEEE Commun. Lett.*, vol. 19, no. 9, pp. 1644–1647, Sep. 2015.

[37] X. Zhu, L. Dai, and Z. Wang, "Graph coloring based pilot allocation to mitigate pilot contamination for multi-cell massive MIMO systems," *IEEE Commun. Lett.*, vol. 19, no. 10, pp. 1842–1845, Oct. 2015.

[38] X. Zhu, L. Dai, Z. Wang, and X. Wang, "Weighted-graph-coloring-based pilot decontamination for multicell massive MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 66, no. 3, pp. 2829–2834, Mar. 2017.

[39] K. Kim, J. Lee, and J. Choi, "Deep learning based pilot allocation scheme (DL-PAS) for 5G massive MIMO system," *IEEE Commun. Lett.*, vol. 22, no. 4, pp. 828–831, Apr. 2018.

[40] P. Zhao, Z. Wang, C. Qian, L. Dai, and S. Chen, "Location-aware pilot assignment for massive MIMO systems in heterogeneous networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, pp. 6815–6821, Aug. 2016.

CHENG HU received the B.S. degree in network engineering from the Nanjing University of Posts and Telecommunications, Nanjing, China, in 2013, where he is currently pursuing the Ph.D. degree.

His research interests include broadband wireless communications, particularly in pilot contamination and interference management in massive multiple-input multiple-output.
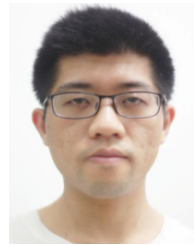
HONG WANG received the B.S. degree in telecommunications engineering from Jiangsu University, Zhenjiang, China, in 2011, and the Ph.D. degree from the Department of Telecommunications Engineering, Nanjing University of Posts and Telecommunications (NUPT), Nanjing, China, in 2016.

From 2014 to 2015, he was a Research Assistant with the Department of Electronic Engineering, City University of Hong Kong, Kowloon, Hong Kong, where he has been a Senior Research Associate with the State Key Laboratory of Millimeter Waves, Department of Electronic Engineering, since 2016. Since July 2019, he has also been an Associate Professor with the Department of Telecommunication Engineering, NUPT. His research interests include broadband wireless communications, particularly in interference analysis and management in HetNets.

RONGFANG SONG (Member, IEEE) received the B.S. and M.S. degrees in telecommunications engineering from the Nanjing University of Posts and Telecommunications (NUPT), Nanjing, China, in 1984 and 1989, respectively, and the Ph.D. degree in telecommunications engineering from Southeast University, Nanjing, in 2001.

From 2002 to 2003, he was a Research Associate with the Department of Electronic Engineering, City University of Hong Kong, Kowloon, Hong Kong. Since 2002, he has been a Professor with the School of Telecommunication and Information Engineering, NUPT. His research interest includes broadband wireless communications with current focus on 5G (B5G) technology.

• • •