Fluid Antenna Aided Intra-Cell Pilot Reuse for MIMO Wireless Networks

Yijia Li, Shuaixin Yang, Yue Xiao, Saviour Zammit, and Kai-Kit Wong

***Abstract*—This paper addresses the problem of pilot contamination in single-cell networks, especially in dense-user areas where intra-cell pilot reuse is inevitable. We invoke fluid antennas (FAs) at the base station (BS) to improve the uplink (UL) channel estimation accuracy in terms of normalized mean square error (NMSE). Specifically, inspired by the fact that the users with low spatial correlation experience less interference, we propose to mitigate pilot contamination by optimizing the FA positions with an innovative objective of minimizing the channel spatial correlation among the pilot-sharing users. To simplify the intractable fractional programming problem involved, we further derive the upper and lower bounds for the objective function. Subsequently, the reformulated problem is effectively solved by a double-loop based algorithm where the outer loop conducts alternating optimization (AO) with respect to the position of each FA while the inner loop handles the subproblems based on the successive convex approximation (SCA) method. Finally, simulation results validate the effectiveness of the proposed algorithm and demonstrate the remarkable NMSE performance provided by FA.**

***Index Terms*—Fluid antenna (FA), pilot reuse, pilot contamination, alternating optimization, successive convex approximation.**

I. Introduction

THE explosive increase of mobile users and the rapid development of internet-of-things (IoT) applications have put forward demanding requirements for future communication systems to support the access of an unprecedentedly large number of users [1]–[3]. To achieve efficient multiuser communication, the multiple-input multiple-output (MIMO) technology has been widely investigated and recognized as a key component of the fifth-generation (5G) networks [4] due to its enormous potential in enhancing system capacity [5], improving energy efficiency [6], and suppressing interference [7]. However, the performance of MIMO systems is fundamentally affected by the estimation accuracy of channel state information (CSI) at the base station (BS). In general, to reduce the overhead of CSI acquisition, MIMO systems can operate in time division duplexing (TDD) mode [8], where the BS receives uplink (UL) pilots sent from the users, conducts channel estimation, and obtains both uplink and downlink CSI by exploiting channel reciprocity [9]. As an ideal situation for TDD UL training, orthogonal training [10] assigns a unique orthogonal pilot sequence to each user, therefore enabling simultaneous pilot transmissions without interference. Unfortunately, for the hotspot areas with high user density and insufficient pilot sequences, reusing pilots among the users within a single cell becomes unavoidable, leading to the problem of pilot contamination [11], which can significantly degrade the accuracy of channel estimation and the overall performance of MIMO systems [9].

To address the issue of pilot contamination, a variety of methods have been developed in [12]–[16] by considering the concept of channel spatial correlation [17]. The motivation stems from that, when the BS employs large-scale uniform linear array (ULA), users with strictly non-overlapping angle of arrival (AoA) intervals have orthogonal channel covariances, thus can completely get rid of pilot contamination [12]. Inspired by this fact, the authors of [13] proposed the statistic greedy pilot scheduling (SGPS) algorithm to minimize the mean square error (MSE) of channel estimation by assigning orthogonal pilots to users with similar channel covariance matrices. In [14]–[16], the authors utilized channel charting to extract spatial information embedded in statistical CSI for the purpose of maximizing the AoA distances between the pilot-sharing users. Apart from only depending on statistical CSI, other strategies to mitigate pilot contamination include allocating pilots based on location information [18]–[20], AoA information [21]–[23], and interference graph coloring [24]–[27]. Nevertheless, even with appropriate pilot assignment, users may still suffer severe pilot contamination in practice, as the random behavior of real propagation environments can destroy the channel orthogonality [28]. On the other hand, the performance of conventional MIMO systems that equipped with fixed-position antennas (FPAs) is intrinsically limited due to the inflexibility of antenna positions and the inadequate utilization of spatial degree of freedom (DoF) [29].

Meanwhile, as a technology to introduce additional DoFs in the spatial domain, the recently proposed fluid antenna (FA) [30] could further enhance system performance by dynamically adjusting the positions of antennas. For instance, it was shown in [31] that deploying FAs at both transmitter and receiver sides in a point-to-point MIMO system could yield tremendous diversity gain over the traditional FPA-based system. Apart from point-to-point communications, the FA-aided multiuser MIMO systems have also attracted widespread

Y. Li, S. Yang, and Y. Xiao are with the National Key Laboratory of Wireless Communications, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: nikoeric@foxmail.com; shuaixin.yang@foxmail.com; xiaoyue@uestc.edu.cn).

S. Zammit is with the Department of Communications and Computer Engineering, University of Malta, Msida MSD 2080, Malta (e-mail: saviour.zammit@um.edu.mt).

K. K. Wong is with the Department of Electronic and Electrical Engineering, University College London, WC1E 7JE London, U.K., and also with the Department of Electronic Engineering, Kyung Hee University, Yongin-si, Gyeonggi-do 17104, South Korea (e-mail: kai-kit.wong@ucl.ac.uk).

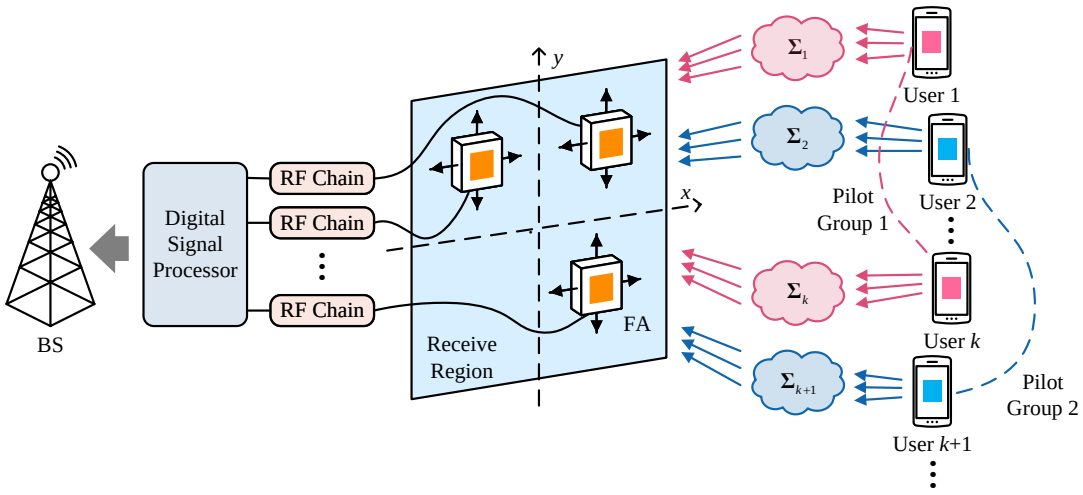


Fig. 1. Illustration of the uplink training between K single-FPA users and the BS equipped with M FAs. Only two groups of pilot-sharing users are depicted in this example, with user 1 and user k reusing the same pilot sequence and user 2 and user $k+1$ sharing another.

research interest [32]–[37]. Specifically, the authors of [32] explored the feasibility of FA-aided multiuser communications and revealed the superiority of FAs in mitigating inter-user interference. Considering the uplink transmission in cellular networks, the authors of [33] minimized the total transmit power of users subject to a minimum rate requirement by expressing the receive signal-to-interference-plus-noise ratio (SINR) as a function of the FA positions. With similar methodology, the authors of [34] maximized the minimum achievable rate among the users within a limited uplink power budget. For the downlink scenario, the joint antenna position and beamforming design was investigated in [35], and for the similar problem, the authors of [36] further emphasized the finite precision of electromechanical devices and established a discrete optimization problem. Moreover, the sum-rate maximization problem for FA-aided MIMO downlink was studied in [37].

In a nutshell, this contribution introduces the FA technology into the realm of single-cell pilot reuse, aiming at mitigating pilot contamination via exploiting the extra spatial DoFs provided by FAs. Specifically, we adjust the positions of FAs at the BS to minimize the channel spatial correlation between the users reusing pilots, thereby enhancing channel estimation quality. In general, the main contributions of this paper are summarized as follows:

- Firstly, we formulate a channel spatial correlation minimization problem as a novel FA optimization objective, so as to mitigate pilot contamination among the pilot-sharing users.
- Secondly, an upper bound along with a lower bound for the objective function is derived to simplify the problem.
- Thirdly, we obtain at least a suboptimal solution for the transformed problem by adopting the alternating optimization (AO) framework as well as the SCA method while proving the convergence of the proposed algorithm.
- Moreover, we investigate the pilot assignment algorithm under the guidance of our proposed intra-group correlation minimization criterion and show that the desired algorithm is essentially synonymous with the Max-τ-Cut algorithm [38] and the SGPS algorithm [13].
- Finally, numerical results are presented to demonstrate the interference mitigation gain of the proposed FA-enhanced scheme over the conventional FPA scheme.

The remainder of the paper is organized as follows. In Section II, we present the system model and formulate the optimization problem of minimizing intra-group correlation among pilot-sharing users. Additionally, we reveal the relationship between the SGPS algorithm and our optimization objective. In Section III, we develop an algorithm based on the AO and SCA techniques to solve the formulated problem. Section IV provides simulation results and discussions. Finally, the study is concluded in Section V.

Notation: Boldface uppercase (lowercase) letters are used to denote matrices (column vectors). The symbol $\mathbf{I}$ denotes the identity matrix. The notations $(\cdot)^T$, $(\cdot)^H$, $\|\cdot\|_p$, and $\|\cdot\|_F$ denote the transpose operation, Hermitian transpose operation, ℓ_p-norm, and Frobenius norm, respectively. The notations $\mathrm{Re}\{\cdot\}$, $\mathbb{E}\{\cdot\}$, $\mathrm{Tr}\{\cdot\}$, and $\mathrm{rank}\{\cdot\}$ represent the real part, the expectation, the trace, and the rank operations, respectively. $\mathrm{diag}\{\mathbf{x}\}$ denotes the diagonal matrix with $\mathbf{x}$ along its main diagonal. $\mathbf{A} \succeq \mathbf{B}$ ($\mathbf{A} \preceq \mathbf{B}$) indicates that the matrix $\mathbf{A}-\mathbf{B}$ is Hermitian positive (negative) semi-definite. Additionally, $\mathcal{CN}(\mathbf{0}, \mathbf{\Sigma})$ denotes the circularly symmetric complex Gaussian (CSCG) distribution with mean zero and covariance matrix $\mathbf{\Sigma}$.

II. System Model and Problem Formulation

As shown in Fig. 1, we consider a multiuser communication scenario in a single cell, where K single-FPA users are served by a base station equipped with M ($\geq K$) FAs. For the UL training process that we are concerned about, a total number of τ available orthogonal pilots are allocated to the K users. We consider the scenario in which the number of available orthogonal pilots is insufficient, i.e., $\tau < K$, thus pilot reuse is supposed to be conducted. More specifically, K users are

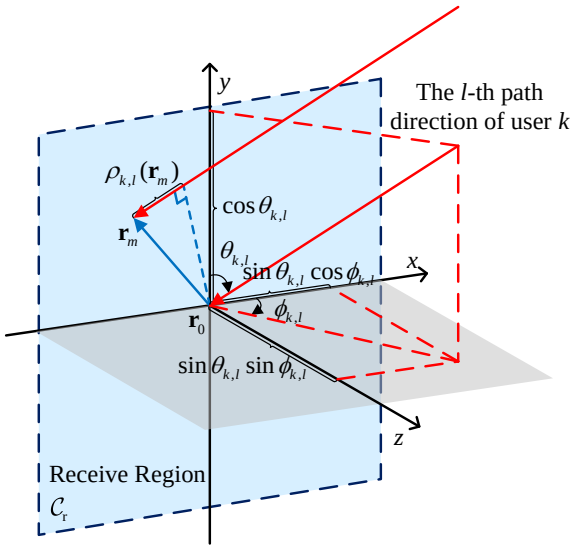


Fig. 2. Illustration of the local coordinate system and spatial angles for the l-th path of user k, $1 \le l \le L_k^{\mathrm{r}}$, $1 \le k \le K$.

divided into τ pilot groups, and the users within the same group share the same pilot, leading to a pilot reuse factor of $\eta = K/\tau$, i.e., each orthogonal pilot is shared by η users on average.

A. Channel Model

The receive FAs at the BS are connected to radio frequency (RF) chains via flexible cables so that they are able to move freely in the given two-dimensional (2D) region $\mathcal{C}_{\mathrm{r}}$. We denote the coordinate of the m-th ($1 \le m \le M$) FA by $\mathbf{r}_m = [x_m, y_m]^T \in \mathcal{C}_{\mathrm{r}}$, and use $\mathbf{r} = [\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_M] \in \mathbb{R}^{2\times M}$ to represent the collection of all receive FAs' positions. Without loss of generality, we set $\mathcal{C}_{\mathrm{r}}$ as a square region with a side length of A, i.e., $\mathcal{C}_{\mathrm{r}} = [-A/2, A/2] \times [-A/2, A/2]$.

Assuming that the size of the moving region $\mathcal{C}_{\mathrm{r}}$ is significantly smaller than the distance between the transmitter and the receiver, it is reasonable to employ the far-field channel model [39]. For each channel path component, the angle of departure (AoD), the AoA, and the amplitude of the complex path coefficient remain invariant regardless of the movement of the FAs, while the phase of the complex path coefficient varies as a function of FA positions.

Let L_k^{t} and L_k^{r} denote the total numbers of transmit and receive channel paths from user k ($1 \le k \le K$) to the BS, respectively. As depicted in Fig. 2, for the l-th ($1 \le l \le L_k^{\mathrm{r}}$) receive path from the k-th user, the elevation and azimuth AoAs are denoted as $\theta_{k,l} \in [0, \pi]$ and $\phi_{k,l} \in [0, \pi]$, respectively. Additionally, the elevation AoAs of all paths of user k are modeled as independent and identically distributed (i.i.d.) random variables with uniform distribution $U\left(\theta_k^{\min}, \theta_k^{\max}\right)$, where $\theta_k^{\min} = \bar{\theta}_k - \sqrt{3}\sigma_\theta$ and $\theta_k^{\max} = \bar{\theta}_k + \sqrt{3}\sigma_\theta$, with $\bar{\theta}_k$ denoting the elevation incident angle of user k and σ_θ denoting the elevation angle standard deviation (ASD). The azimuth AoAs are modeled in a similar way, i.e., $\phi_{k,l} \sim U\left(\phi_k^{\min}, \phi_k^{\max}\right), \forall l$, where $\phi_k^{\min} = \bar{\phi}_k - \sqrt{3}\sigma_\phi$ and $\phi_k^{\max} = \bar{\phi}_k + \sqrt{3}\sigma_\phi$, with $\bar{\phi}_k$ denoting the azimuth incident angle of user k and σ_ϕ denoting the azimuth ASD.

At the BS, the signal propagation distance difference between the position of the m-th FA and the reference point $\mathbf{r}_0 = [0, 0]^T$ can be obtained by

$$\rho_{k,l}(\mathbf{r}_m) = x_m \sin\theta_{k,l}\cos\phi_{k,l} + y_m \cos\theta_{k,l}. \tag{1}$$

Correspondingly, the signal phase difference between the positions $\mathbf{r}_m$ and $\mathbf{r}_0$ is $2\pi\rho_{k,l}(\mathbf{r}_m)/\lambda$, where λ is the wavelength. Then, for the signal sent form user k, the field-response vector (FRV) of the m-th FA at the BS is written as

$$\mathbf{f}_k(\mathbf{r}_m) = \left[e^{j\frac{2\pi}{\lambda}\rho_{k,1}(\mathbf{r}_m)}, e^{j\frac{2\pi}{\lambda}\rho_{k,2}(\mathbf{r}_m)}, \ldots, e^{j\frac{2\pi}{\lambda}\rho_{k,L_k^{\mathrm{r}}}(\mathbf{r}_m)}\right]^T. \tag{2}$$

At the user side, by setting the position of the single FPA at the reference point, the FRV of each user becomes an all-ones vector, i.e., $\mathbf{g}_k = \mathbf{1} \in \mathbb{R}^{L_k^{\mathrm{t}}\times 1}, \forall k$. The resulting channel vector between the BS and user k can be expressed as

$$\mathbf{h}_k(\mathbf{r}) = \mathbf{F}_k^H(\mathbf{r})\,\mathbf{\Sigma}_k\mathbf{g}_k \in \mathbb{C}^{M\times 1}, \tag{3}$$

where $\mathbf{F}_k(\mathbf{r}) = [\mathbf{f}_k(\mathbf{r}_1), \mathbf{f}_k(\mathbf{r}_2), \ldots, \mathbf{f}_k(\mathbf{r}_M)] \in \mathbb{C}^{L_k^{\mathrm{r}}\times M}$ is the field-response matrix (FRM) between the BS and user k, and $\mathbf{\Sigma}_k \in \mathbb{C}^{L_k^{\mathrm{r}}\times L_k^{\mathrm{t}}}$ is the path-response matrix (PRM) between the BS and user k, with its (q,p)-th element characterizing the channel response between the p-th transmit path and the q-th receive path.

Given the FA position variable $\mathbf{r}$, we now present the expression of the channel covariance matrix in the following. Suppose that $L_k^{\mathrm{r}} = L_k^{\mathrm{t}}, \forall k$ and that $\mathbf{\Sigma}_k = \mathrm{diag}\left\{\Sigma_{k,1}, \ldots, \Sigma_{k,L_k^{\mathrm{r}}}\right\}$ is a diagonal matrix with its diagonal elements being i.i.d. CSCG random variables, i.e., $\Sigma_{k,l} \sim \mathcal{CN}\left(0, \sigma_{h,k}^2\right)$ ($1 \le l \le L_k^{\mathrm{r}}$). By noting that $\mathbf{\Sigma}_k\mathbf{g}_k$ is a random vector following $\mathcal{CN}\left(\mathbf{0}, \sigma_{h,k}^2\mathbf{I}\right)$, the channel covariance martrix $\mathbf{R}_k \in \mathbb{C}^{M\times M}$ can be expressed as

$$\begin{aligned}\mathbf{R}_k(\mathbf{r}) =& \mathbb{E}\left\{\mathbf{h}_k(\mathbf{r})\,\mathbf{h}_k^H(\mathbf{r})\right\}\\ =& \mathbf{F}_k^H(\mathbf{r})\,\mathbb{E}\left\{\mathbf{\Sigma}_k\mathbf{g}_k\mathbf{g}_k^H\mathbf{\Sigma}_k^H\right\}\mathbf{F}_k(\mathbf{r})\\ =& \sigma_{h,k}^2\mathbf{F}_k^H(\mathbf{r})\,\mathbf{F}_k(\mathbf{r}).\end{aligned} \tag{4}$$

It merits attention that we aim at mitigating pilot contamination for the UL training phase, at which time instantaneous CSI is not available yet at the BS. For this reason, we are limited to solving the problem depending solely on the knowledge of relatively slow-varying statistical CSI, i.e., covariance matrices, which are mainly determined by AoAs and FA positions.

B. Uplink Channel Estimation

We consider a block-fading channel that is time-invariant within a coherence block consisting of T_c symbols. The pilot length is assumed to be exactly equal to the number of pilots, τ, to guarantee the orthogonality of the pilot sequences. The resulting pilot codebook is a scaled unitary matrix $\mathbf{\Psi} = [\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \ldots, \boldsymbol{\psi}_\tau] \in \mathbb{C}^{\tau\times\tau}$, satisfying that $\mathbf{\Psi}^H\mathbf{\Psi} = \sigma_\psi^2\mathbf{I}$, where σ_ψ^2 is the transmit power of the pilot signal. Additionally, the pilot length τ must satisfy $\tau < T_c$ so that enough symbols are reserved for data transmission. Let $\mathcal{K} = \{1, 2, \ldots, K\}$ and $\mathcal{T} = \{1, 2, \ldots, \tau\}$ be the set of

users and the set of available pilot indices, respectively. The pilot assignment patterns are represented as $\{\pi_k\}_{k=1}^K$, where $\pi_k \in \mathcal{T}$ indicates that the π_k-th pilot sequence $\boldsymbol{\psi}_{\pi_k}$ is allocated to user k.

With a given pilot assignment pattern $\{\pi_k\}_{k=1}^K$, users transmit their pilots simultaneously during the training phase made up of τ symbols, and the received signals $\mathbf{Y} \in \mathbb{C}^{M\times\tau}$ at the BS can be written as

$$\begin{aligned}\mathbf{Y} &= \mathbf{H}(\mathbf{r})\,\mathbf{X} + \mathbf{N} \\ &= \sum_{k=1}^{K} \mathbf{h}_k(\mathbf{r})\,\boldsymbol{\psi}_{\pi_k}^T + \mathbf{N},\end{aligned} \tag{5}$$

where $\mathbf{H}(\mathbf{r}) = [\mathbf{h}_1(\mathbf{r}), \mathbf{h}_2(\mathbf{r}), \ldots, \mathbf{h}_K(\mathbf{r})] \in \mathbb{C}^{M\times K}$ is the channel matrix, $\mathbf{X} = \left[\boldsymbol{\psi}_{\pi_1}, \boldsymbol{\psi}_{\pi_2}, \ldots, \boldsymbol{\psi}_{\pi_K}\right]^T \in \mathbb{C}^{K\times\tau}$ is the pilot signal matrix, and $\mathbf{N} \in \mathbb{C}^{M\times\tau}$ is the Gaussian noise matrix whose elements are i.i.d. random variables following $\mathcal{CN}\left(0, \sigma_z^2\right)$. Note that although the channels (along with some other correlation matrices) are determined by the FA positions $\mathbf{r}$, their dependence on $\mathbf{r}$ is suppressed for notational simplicity when it does not cause ambiguity.

To estimate the channel of user k, the processed received signal is obtained as

$$\begin{aligned}\mathbf{y}_k^{\mathrm{p}} &= \frac{1}{\sigma_\psi^2}\mathbf{Y}\boldsymbol{\psi}_{\pi_k}^* \\ &= \mathbf{h}_k + \sum_{j\in\mathcal{I}_k} \mathbf{h}_j + \mathbf{n}_k^{\mathrm{p}},\end{aligned} \tag{6}$$

where $\mathcal{I}_k = \{j \mid j \in \mathcal{K}, j \neq k, \pi_j = \pi_k\}$ denotes the set of users sharing the same pilot sequence with user k, and $\mathbf{n}_k^{\mathrm{p}} \triangleq \mathbf{N}\boldsymbol{\psi}_{\pi_k}^*/\sigma_\psi^2$ is the effective noise following $\mathcal{CN}\left(\mathbf{0}, \frac{1}{\rho}\mathbf{I}\right)$, with $\rho \triangleq \sigma_\psi^2/\sigma_z^2$ being the signal-to-noise ratio (SNR).

It can be observed from (6) that severe interference in channel estimation (the so-called pilot contamination) is induced by pilot reuse. In order to mitigate this issue, the minimum mean square error (MMSE) estimator is employed, and the consequent channel estimation of user k is derived as

$$\hat{\mathbf{h}}_k = \mathbf{R}_k\mathbf{Q}_k^{-1}\mathbf{y}_k^{\mathrm{p}}, \tag{7}$$

where $\mathbf{R}_k = \mathbb{E}\left\{\mathbf{h}_k\mathbf{h}_k^H\right\}$ is the channel covariance matrix, and $\mathbf{Q}_k = \mathbb{E}\left\{\mathbf{y}_k^{\mathrm{p}}\left(\mathbf{y}_k^{\mathrm{p}}\right)^H\right\}$ is the correlation matrix of $\mathbf{y}_k^{\mathrm{p}}$, which can be alternatively written as

$$\mathbf{Q}_k = \mathbf{R}_k + \sum_{j\in\mathcal{I}_k}\mathbf{R}_j + \frac{1}{\rho}\mathbf{I}. \tag{8}$$

The channel estimation error of the above MMSE estimator is denoted by $\tilde{\mathbf{h}}_k \triangleq \hat{\mathbf{h}}_k - \mathbf{h}_k$ and is distributed according to $\mathcal{CN}\left(\mathbf{0}, \mathbf{R}_{\tilde{\mathbf{h}}_k}\right)$, where

$$\mathbf{R}_{\tilde{\mathbf{h}}_k} = \mathbf{R}_k - \mathbf{R}_k\mathbf{Q}_k^{-1}\mathbf{R}_k. \tag{9}$$

Correspondingly, the mean square error of channel estimation (MSE-CE) of user k is defined as

$$\epsilon_k \triangleq \mathbb{E}\left\{\left\|\hat{\mathbf{h}}_k - \mathbf{h}_k\right\|_2^2\right\} = \mathrm{Tr}\left\{\mathbf{R}_{\tilde{\mathbf{h}}_k}\right\}. \tag{10}$$

MSE-CE is a pivotal metric to quantify the accuracy of channel estimation, and the channel estimation quality can inherently affect the performance of data transmission. However, as is shown in (6), pilot contamination causes inaccuracies in channel estimation, leading to larger MSE-CE.

C. Problem Formulation

It was proved in [13] that the minimum of MSE-CE is achieved under the condition that every group of pilot-sharing users exhibits channel orthogonality, i.e., for $\forall i \in \mathcal{K}$,

$$\delta(\mathbf{R}_i, \mathbf{R}_j) = 0, \quad \forall j \in \mathcal{I}_i, \tag{11}$$

where $\delta(\mathbf{R}_i, \mathbf{R}_j) \in [0, 1]$ is a correlation metric between two channel covariance matrices $\mathbf{R}_i$ and $\mathbf{R}_j$. Essentially, $\delta(\mathbf{R}_i, \mathbf{R}_j)$ is defined by the normalized inner product on $\mathbb{C}^{M\times M}$, expressed as

$$\delta(\mathbf{R}_i, \mathbf{R}_j) \triangleq \frac{\mathrm{Tr}\left\{\mathbf{R}_i^H\mathbf{R}_j\right\}}{\|\mathbf{R}_i\|_F\|\mathbf{R}_j\|_F}. \tag{12}$$

Note that $1 - \delta(\mathbf{R}_i, \mathbf{R}_j)$ is known as the correlation matrix distance (CMD) [40], while $\delta(\mathbf{R}_i, \mathbf{R}_j)$ itself is called "channel spatial correlation" in this paper, as it can effectively measure how similar the spatial properties of users are in spatially correlated channels [17]. Obviously, $\delta(\mathbf{R}_i, \mathbf{R}_j) = 0$ indicates that the users i and j have covariance matrices with orthogonal support, i.e., they are spatially orthogonal.

To further give a theoretical interpretation of the minimum MSE-CE condition in (11), we resort to the user-averaged normalized MSE of channel estimation (NMSE-CE), which takes the form of

$$\begin{aligned}\bar{\epsilon} &\triangleq \frac{1}{K}\sum_{k=1}^{K}\frac{\mathbb{E}\left\{\left\|\hat{\mathbf{h}}_k - \mathbf{h}_k\right\|_2^2\right\}}{\mathbb{E}\left\{\|\mathbf{h}_k\|_2^2\right\}} \\ &= \frac{1}{K}\sum_{k=1}^{K}\frac{\mathrm{Tr}\left\{\mathbf{R}_{\tilde{\mathbf{h}}_k}\right\}}{\mathrm{Tr}\left\{\mathbf{R}_k\right\}} \\ &= \frac{1}{K}\sum_{k=1}^{K}\left(1 - \frac{\mathrm{Tr}\left\{\mathbf{R}_k\mathbf{Q}_k^{-1}\mathbf{R}_k\right\}}{\mathrm{Tr}\left\{\mathbf{R}_k\right\}}\right).\end{aligned} \tag{13}$$

In the context of ideal channel estimation, we have $\mathbf{Q}_k = \mathbf{R}_k, \forall k$, resulting in zero NMSE-CE. However, recalling (8), the terms $\mathbf{R}_j$ and $(1/\rho)\mathbf{I}$ entering into $\mathbf{Q}_k$ embody the impact of interference from pilot-sharing users and the noise, respectively. When the optimal condition in (11) is satisfied, it can be proved that the NMSE-CE is minimized as

$$\bar{\epsilon}_{\min} = 1 - \frac{1}{K}\sum_{k=1}^{K}\frac{\mathrm{Tr}\left\{\mathbf{R}_k\left(\mathbf{R}_k + \frac{1}{\rho}\mathbf{I}\right)^{-1}\mathbf{R}_k\right\}}{\mathrm{Tr}\left\{\mathbf{R}_k\right\}}. \tag{14}$$

By examining the difference between (13) and (14), it becomes evident that the minimum MSE-CE condition described in (11) corresponds to the situation where pilot contamination among the users completely vanishes.

Unfortunately, the optimal condition given in (11), i.e., $\delta(\mathbf{R}_i, \mathbf{R}_j) = 0$, or equivalently $\mathbf{R}_i\mathbf{R}_j = 0$, for $j \in \mathcal{I}_i, \forall i \in \mathcal{K}$, is unlikely to be satisfied in practice [5]. In fact, the channel spatial orthogonality between the pilot-sharing users can be

destroyed by several commonly seen factors, including the insufficient number of BS antennas, the excessive number of users, and the propagation environments with spatial variations [28].

In this paper, we aim to optimize the NMSE-CE performance and mitigate pilot contamination by exploiting the spatial DoFs provided by FA. Instead of directly minimizing the NMSE-CE, which involves intractable matrix inversion, we resort to another interference suppression criterion, i.e., reducing the spatial correlation $\delta(\mathbf{R}_i, \mathbf{R}_j)$ among the pilot-sharing users as much as possible [17]. Hence we introduce an objective function referred to as "intra-group correlation" throughout this paper and written as

$$\begin{aligned} J(\mathbf{r}) &= \sum_{i\in\mathcal{K}} \sum_{\substack{j>i \\ j\in\mathcal{I}_i}} \delta\left(\mathbf{R}_i(\mathbf{r}), \mathbf{R}_j(\mathbf{r})\right) \\ &= \sum_{i\in\mathcal{K}} \sum_{\substack{j>i \\ j\in\mathcal{I}_i}} \frac{\mathrm{Tr}\left\{\mathbf{R}_i^H(\mathbf{r})\mathbf{R}_j(\mathbf{r})\right\}}{\|\mathbf{R}_i(\mathbf{r})\|_F \|\mathbf{R}_j(\mathbf{r})\|_F}. \end{aligned} \tag{15}$$

As can be seen, the intra-group correlation is obtained by summing up the channel spatial correlations among all pilot-sharing users.

In the sequel, we focus on minimizing the intra-group correlation in (15), which is consistent with the intuition that users with low spatial correlation are almost non-interfering with each other. The corresponding optimization problem is formulated as

$$\text{(P1)} \quad \min_{\mathbf{r}} \quad \sum_{i\in\mathcal{K}} \sum_{j\in\mathcal{I}_i'} \frac{\mathrm{Tr}\left\{\mathbf{R}_i^H(\mathbf{r})\mathbf{R}_j(\mathbf{r})\right\}}{\|\mathbf{R}_i(\mathbf{r})\|_F \|\mathbf{R}_j(\mathbf{r})\|_F} \tag{16a}$$

$$s.t. \quad \mathbf{r}_m \in \mathcal{C}_\mathrm{r}, \quad m = 1, 2, \ldots, M \tag{16b}$$

$$\|\mathbf{r}_m - \mathbf{r}_n\|_2 \geq D, \quad m, n = 1, 2, \ldots, M, m \neq n, \tag{16c}$$

where $\mathcal{I}_k' = \{j \mid j \in \mathcal{I}_k, j > i\}$, and D is a minimum required distance between FAs to avoid the coupling effect. Problem (P1) is an intractable fractional programming problem, thus we propose an algorithm to handle this problem in Section III.

D. Supplementary Notes for Pilot Assignment

Before proceeding further, it is important to clarify that our proposed algorithm operates under a predefined pilot assignment pattern $\{\pi_k \in \mathcal{T}\}_{k=1}^K$ which itself can substantially affect the degree of pilot contamination. In this subsection, we focus on developing an intelligent pilot assignment algorithm that aligns with the aforementioned idea of minimizing intra-group correlation. Coincidentally, the desired algorithm eventually turns out to be essentially synonymous with the Max-τ-Cut algorithm [38] and the SGPS algorithm [13].

Under the intra-group correlation minimization criterion, we first formulate the pilot assignment problem as

$$\min_{\{\pi_k\}_{k=1}^K} \quad \sum_{i\in\mathcal{K}} \sum_{\substack{j>i \\ j\in\mathcal{I}_i}} \delta\left(\mathbf{R}_i, \mathbf{R}_j\right). \tag{17}$$

Minimizing (17) is about strategically allocating identical pilot sequences to the spatially low-correlated users, thereby suppressing the mutual interference. By constructing additional constant summation terms, (17) is equivalently rewritten as

$$\max_{\{\pi_k\}_{k=1}^K} \quad \sum_{i\in\mathcal{K}} \left(\sum_{j>i} \delta\left(\mathbf{R}_i, \mathbf{R}_j\right) - \sum_{\substack{j>i \\ j\in\mathcal{I}_i}} \delta\left(\mathbf{R}_i, \mathbf{R}_j\right) \right). \tag{18}$$

Thus, problem (17) is equivalent to

$$\max_{\{\pi_k\}_{k=1}^K} \quad \sum_{i\in\mathcal{K}} \sum_{\substack{j>i \\ j\notin\mathcal{I}_i}} \delta\left(\mathbf{R}_i, \mathbf{R}_j\right). \tag{19}$$

This NP-hard problem is also known as the Max-τ-Cut problem in graph theory [26], [27]. Fortunately, a heuristic algorithm proposed in [38] is capable of obtaining a suboptimal solution that achieves at least $(1 - 1/\tau)$ of the optimal solution for the Max-τ-Cut problem. According to these previous works, the resulting pilot assignment algorithm is given in Algorithm 1 and referred to as the "Max-τ-Cut algorithm."

Algorithm 1 Pilot Assignment Algorithm for a Suboptimal Solution to the Max-τ-Cut Problem

1: Randomly assign τ users to τ pilot groups such that each group has one user.
2: Assign the unscheduled user i to group π_i, where $\pi_i = \arg\min_{t\in\mathcal{T}} \sum_{\pi_j=t,\, j\in\mathcal{K}} \delta\left(\mathbf{R}_i, \mathbf{R}_j\right)$.
3: Repeat step 2 until each of the remaining $(K - \tau)$ users is assigned to a pilot group.

The SGPS algorithm [13] can be obtained by modifying the first step of Algorithm 1 into constructing a set of τ users with high spatial correlation and assigning them into τ different groups while keeping other steps unchanged. Considering that the users who tend to experience severe interference are assigned to distinct pilot groups in the initialization of the SGPS algorithm, it is intuitively expected to be more effective than the Max-τ-Cut algorithm in terms of pilot contamination suppression.

From the preceding discussion, both the Max-τ-Cut algorithm and the SGPS algorithm inherently share the same optimization objective as problem (P1), while the SGPS algorithm can be regarded as a refined version of the former, which is subsequently validated by the simulation results in Section IV.

III. Proposed Algorithm

In this section, we present an algorithm to solve (P1) for an arbitrary pilot assignment pattern $\{\pi_k\}_{k=1}^K$. We first analytically derive an upper bound along with a lower bound for (16a) so as to transfer problem (P1) into a more tractable one. Then we utilize an alternating optimization approach to obtain at least a locally optimal solution for the new problem. Specifically, we alternately optimize the position of one receive FA, with all the other FAs being fixed. For each subproblem in the AO algorithm, we use the SCA technique to convert the nonconvex subproblem into a series of convex problems. Finally, we discuss the convergence and computational complexity of the proposed algorithm.

A. The Upper and Lower Bounds of the Spatial Correlation

By rewriting the denominator of any summation term in (16a) as

$$\left\|\mathbf{R}_i\right\|_F \left\|\mathbf{R}_j\right\|_F = \sqrt{\mathrm{Tr}\left(\mathbf{R}_i\mathbf{R}_i\right)\mathrm{Tr}\left(\mathbf{R}_j\mathbf{R}_j\right)}, \tag{20}$$

we have

$$\begin{aligned}\mathrm{Tr}\left(\mathbf{R}_i\right)\mathrm{Tr}\left(\mathbf{R}_j\right) &\overset{(a)}{\geq} \left\|\mathbf{R}_i\right\|_F \left\|\mathbf{R}_j\right\|_F \\ &\overset{(b)}{\geq} \frac{\mathrm{Tr}\left(\mathbf{R}_i\right)\mathrm{Tr}\left(\mathbf{R}_j\right)}{\sqrt{\mathrm{rank}\left(\mathbf{R}_i\right)\mathrm{rank}\left(\mathbf{R}_j\right)}},\end{aligned} \tag{21}$$

where (a) holds since $\left[\mathrm{Tr}\left(\mathbf{X}\right)\right]^2 \geq \mathrm{Tr}\left(\mathbf{X}^2\right), \forall \mathbf{X} \succeq 0$, and (b) holds due to $\mathrm{Tr}\left(\mathbf{X}^2\right) \geq \left[\mathrm{Tr}\left(\mathbf{X}\right)\right]^2 / \mathrm{rank}\left(\mathbf{X}\right), \forall \mathbf{X} \succeq 0$ [41].

Recalling (4), for the k-th user, $\mathbf{R}_k$ is expressed as $\mathbf{R}_k = \sigma_{h,k}^2 \mathbf{F}_k^H \mathbf{F}_k$, hence we have $\mathrm{Tr}\left(\mathbf{R}_k\right) = \sigma_{h,k}^2 \left\|\mathbf{F}_k\right\|_F^2 = \sigma_{h,k}^2 L_k^{\mathrm{r}} M$ and $\mathrm{rank}\left(\mathbf{R}_k\right) = \mathrm{rank}\left(\mathbf{F}_k\right) \leq \min\left(L_k^{\mathrm{r}}, M\right)$. Based on the results above, we arrive at

$$\begin{aligned}\sigma_{h,i}^2\sigma_{h,j}^2 L_i^{\mathrm{r}} L_j^{\mathrm{r}} M^2 &\geq \left\|\mathbf{R}_i\right\|_F \left\|\mathbf{R}_j\right\|_F \\ &\geq \sigma_{h,i}^2\sigma_{h,j}^2 \frac{L_i^{\mathrm{r}} L_j^{\mathrm{r}} M^2}{\sqrt{\min\left(L_i^{\mathrm{r}}, M\right)\min\left(L_j^{\mathrm{r}}, M\right)}} \\ &\geq \sigma_{h,i}^2\sigma_{h,j}^2 L_i^{\mathrm{r}} L_j^{\mathrm{r}} M.\end{aligned} \tag{22}$$

By applying the inequality in (22) to each summation term in (16a) and further conducting inequality relaxation, the upper bound and the lower bound of the objective function are given by

$$\begin{aligned}&\frac{1}{\left(L_{\max}^{\mathrm{r}}\right)^2 M^2} \sum_{i\in\mathcal{K}} \sum_{j\in\mathcal{I}_i'} \mathrm{Tr}\left\{\mathbf{F}_i^H \mathbf{F}_i \mathbf{F}_j^H \mathbf{F}_j\right\} \\ &\quad\leq \sum_{i\in\mathcal{K}} \sum_{j\in\mathcal{I}_i'} \frac{\mathrm{Tr}\left\{\mathbf{R}_i^H \mathbf{R}_j\right\}}{\left\|\mathbf{R}_i\right\|_F \left\|\mathbf{R}_j\right\|_F} \\ &\quad\leq \frac{1}{\left(L_{\min}^{\mathrm{r}}\right)^2 M} \sum_{i\in\mathcal{K}} \sum_{j\in\mathcal{I}_i'} \mathrm{Tr}\left\{\mathbf{F}_i^H \mathbf{F}_i \mathbf{F}_j^H \mathbf{F}_j\right\},\end{aligned} \tag{23}$$

where $L_{\max}^{\mathrm{r}}$ and $L_{\max}^{\mathrm{r}}$ denote the maximum and the minimum of $\{L_k^{\mathrm{r}}\}_{k=1}^K$, respectively. It should be emphasized that the upper bound is merely a constant scaling of the lower bound, with the scaling factor no larger than M if $L_{\max}^{\mathrm{r}} = L_{\min}^{\mathrm{r}}$.

The upper and lower bounds in (23) constitute the theoretical basis for reformulating the fractional programming problem (P1) as

$$\begin{aligned}(\mathrm{P2}) \quad \min_{\mathbf{r}} \quad & \sum_{i\in\mathcal{K}} \sum_{j\in\mathcal{I}_i'} \mathrm{Tr}\left\{\mathbf{F}_i^H\left(\mathbf{r}\right)\mathbf{F}_i\left(\mathbf{r}\right)\mathbf{F}_j^H\left(\mathbf{r}\right)\mathbf{F}_j\left(\mathbf{r}\right)\right\} \\ s.t. \quad & \text{(16b), (16c)}.\end{aligned} \tag{24}$$

Although the original objective function is replaced by its numerator, the optimization problem is still very challenging since it has a non-convex objective function (24) with non-convex constraints (16c).

B. Alternating Optimization for Problem (P2)

The AO method is utilized to solve (P2) in an alternate manner. Specifically, in each iteration, we optimize the objective function with respect to the position of a certain FA while keeping the remaining variables fixed. All variables are iterated until the convergence condition is satisfied. The AO algorithm can at least guarantee a locally optimal solution for problem (P2).

Given $\{\mathbf{r}_n, n \neq m\}_{n=1}^M$, $\mathbf{r}_m$ becomes the only optimization variable. To stress the variation of the objective function in (24) with $\mathbf{r}_m$, we devide $\mathbf{F}_k\left(\mathbf{r}\right)$ into $\bar{\mathbf{F}}_{k,m} = \left[\mathbf{f}_k\left(\mathbf{r}_1\right), \mathbf{f}_k\left(\mathbf{r}_2\right), \ldots, \mathbf{f}_k\left(\mathbf{r}_{m-1}\right), \mathbf{f}_k\left(\mathbf{r}_{m+1}\right), \ldots, \mathbf{f}_k\left(\mathbf{r}_M\right)\right]$ and $\mathbf{f}_k\left(\mathbf{r}_m\right)$. Subsequently, a certain summation term in (24) can be rewritten as

$$\begin{aligned}&\mathrm{Tr}\left\{\mathbf{F}_i^H\left(\mathbf{r}\right)\mathbf{F}_i\left(\mathbf{r}\right)\mathbf{F}_j^H\left(\mathbf{r}\right)\mathbf{F}_j\left(\mathbf{r}\right)\right\} \\ &\quad= \mathrm{Tr}\Big\{\left[\mathbf{f}_i\left(\mathbf{r}_m\right)\mathbf{f}_j^H\left(\mathbf{r}_m\right) + \bar{\mathbf{F}}_{i,m}\bar{\mathbf{F}}_{j,m}^H\right] \\ &\qquad\qquad \left[\mathbf{f}_j\left(\mathbf{r}_m\right)\mathbf{f}_i^H\left(\mathbf{r}_m\right) + \bar{\mathbf{F}}_{j,m}\bar{\mathbf{F}}_{i,m}^H\right]\Big\} \\ &\quad= \underbrace{2\,\mathrm{Re}\left\{\mathbf{f}_i^H\left(\mathbf{r}_m\right)\mathbf{C}_{ij}\mathbf{f}_j\left(\mathbf{r}_m\right)\right\}}_{h_{ij}(\mathbf{r}_m)} \\ &\qquad+ \underbrace{\mathrm{Tr}\left\{\bar{\mathbf{F}}_{i,m}\bar{\mathbf{F}}_{j,m}^H\bar{\mathbf{F}}_{j,m}\bar{\mathbf{F}}_{i,m}^H\right\} + L_i^{\mathrm{r}}L_j^{\mathrm{r}}}_{\text{constant}},\end{aligned} \tag{25}$$

where $\mathbf{C}_{ij} \in \mathbb{C}^{L_i^{\mathrm{r}} \times L_j^{\mathrm{r}}}$ is a constant matrix expressed as

$$\mathbf{C}_{ij} = \bar{\mathbf{F}}_{i,m}\bar{\mathbf{F}}_{j,m}^H = \sum_{n=1, n\neq m}^{M} \mathbf{f}_i\left(\mathbf{r}_m\right)\mathbf{f}_j^H\left(\mathbf{r}_m\right). \tag{26}$$

According to (25), when $\{\mathbf{r}_n, n \neq m\}_{n=1}^M$ are fixed, the objective function in (24) turns out to be determined by

$$h_{ij}\left(\mathbf{r}_m\right) \triangleq 2\,\mathrm{Re}\left\{\mathbf{f}_i^H\left(\mathbf{r}_m\right)\mathbf{C}_{ij}\mathbf{f}_j\left(\mathbf{r}_m\right)\right\}. \tag{27}$$

An important observation is that $h_{ij}\left(\mathbf{r}_m\right)$ closely resembles a quadratic form which could provide convenience for analysis. Inspired by this, we rewrite $h_{ij}\left(\mathbf{r}_m\right)$ as

$$\begin{aligned}h_{ij}\left(\mathbf{r}_m\right) =& 2\,\mathrm{Re}\left\{\mathbf{f}_i^H\left(\mathbf{r}_m\right)\mathbf{C}_{ij}\mathbf{f}_j\left(\mathbf{r}_m\right)\right\} \\ =& \mathbf{f}_i^H\left(\mathbf{r}_m\right)\mathbf{C}_{ij}\mathbf{f}_j\left(\mathbf{r}_m\right) + \mathbf{f}_j^H\left(\mathbf{r}_m\right)\left(\mathbf{C}_{ij}\right)^H\mathbf{f}_i\left(\mathbf{r}_m\right) \\ =& \mathbf{f}_{ij}^H\left(\mathbf{r}_m\right)\mathbf{A}_{ij}\mathbf{f}_{ij}\left(\mathbf{r}_m\right),\end{aligned} \tag{28}$$

where $\mathbf{f}_{ij}\left(\mathbf{r}_m\right) = \left[\mathbf{f}_i^T\left(\mathbf{r}_m\right) \quad \mathbf{f}_j^T\left(\mathbf{r}_m\right)\right]^T \in \mathbb{C}^{\left(L_i^{\mathrm{r}}+L_j^{\mathrm{r}}\right)\times 1}$ and

$$\mathbf{A}_{ij} = \begin{bmatrix}\mathbf{0} & \mathbf{C}_{ij} \\ \left(\mathbf{C}_{ij}\right)^H & \mathbf{0}\end{bmatrix} \in \mathbb{C}^{\left(L_i^{\mathrm{r}}+L_j^{\mathrm{r}}\right)\times\left(L_i^{\mathrm{r}}+L_j^{\mathrm{r}}\right)}. \tag{29}$$

To conclude, given $\{\mathbf{r}_n, n \neq m\}_{n=1}^M$, the subproblem with respect to $\mathbf{r}_m$ can be equivalently reformulated as

$$(\mathrm{P3}) \quad \min_{\mathbf{r}_m} \quad \sum_{i\in\mathcal{K}} \sum_{j\in\mathcal{I}_i'} \mathbf{f}_{ij}^H\left(\mathbf{r}_m\right)\mathbf{A}_{ij}\mathbf{f}_{ij}\left(\mathbf{r}_m\right) \tag{30a}$$

$$s.t. \quad \mathbf{r}_m \in \mathcal{C}_{\mathrm{r}}, \tag{30b}$$

$$\left\|\mathbf{r}_m - \mathbf{r}_n\right\|_2 \geq D, \quad n = 1, 2, \ldots, M, \quad n \neq m. \tag{30c}$$

Although expressed as a quadratic form, $h_{ij}\left(\mathbf{r}_m\right) = \mathbf{f}_{ij}^H\left(\mathbf{r}_m\right)\mathbf{A}_{ij}\mathbf{f}_{ij}\left(\mathbf{r}_m\right)$ is still neither convex with respect to $\mathbf{r}_m$ nor convex with respect to $\mathbf{f}_{ij}\left(\mathbf{r}_m\right)$. A tremendous obstacle is that $\mathbf{A}_{ij}$ is bound to be an indefinite matrix. One can confirm

this fact by proving that the non-zero eigenvalues of $\mathbf{A}_{ij}$ must emerge in positive-negative pairs. To be specific, if λ is a non-zero eigenvalue of $\mathbf{A}_{ij}$ and $\mathbf{v}_{ij} = \begin{bmatrix}\mathbf{v}_i^T & \mathbf{v}_j\end{bmatrix}^T$ is the corresponding eigenvector, where $\mathbf{v}_i \in \mathbb{C}^{L_i^{\mathrm{r}}\times 1}$, $\mathbf{v}_j \in \mathbb{C}^{L_j^{\mathrm{r}}\times 1}$, then $-\lambda$ is also a eigenvalue of $\mathbf{A}_{ij}$, considering that

$$\begin{cases}\mathbf{A}_{ij}\begin{bmatrix}\mathbf{v}_i\\ \mathbf{v}_j\end{bmatrix} = \begin{bmatrix}\mathbf{C}_{ij}\mathbf{v}_j\\ \left(\mathbf{C}_{ij}\right)^H\mathbf{v}_i\end{bmatrix} = \lambda\begin{bmatrix}\mathbf{v}_i\\ \mathbf{v}_j\end{bmatrix},\\ \mathbf{A}_{ij}\begin{bmatrix}-\mathbf{v}_i\\ \mathbf{v}_j\end{bmatrix} = \begin{bmatrix}\mathbf{C}_{ij}\mathbf{v}_j\\ -\left(\mathbf{C}_{ij}\right)^H\mathbf{v}_i\end{bmatrix} = -\lambda\begin{bmatrix}-\mathbf{v}_i\\ \mathbf{v}_j\end{bmatrix}.\end{cases} \tag{31}$$

C. Successive Convex Approximation for Problem (P3)

The indefiniteness of $\mathbf{A}_{ij}$ considerably increases the intractability of problem (P3). As a countermeasure, we adopt an SCA approach to solve (P3). The key idea is to optimize the objective function in an iterative manner. In each iteration, a surrogate function is constructed at the current feasible point, satisfying the upper-bound property and convex property, and then the next feasible point is obtained through minimizing the surrogate function. The value of objective function is non-increasing during the iterations, and convergence is guaranteed.

In order to derive an upper bound for (30a) and construct the surrogate function, we start by defining a negative semi-definite matrix $\mathbf{B}_{ij}$ as

$$\mathbf{B}_{ij} \triangleq \mathbf{A}_{ij} - \lambda_{ij}^m\mathbf{I} \preceq 0, \tag{32}$$

where λ_{ij}^m is the largest eigenvalue of $\mathbf{A}_{ij}$. Since $\mathbf{B}_{ij} \preceq 0$, the function $p_{ij}\left(\mathbf{r}_m\right) \triangleq \mathbf{f}_{ij}^H\left(\mathbf{r}_m\right)\mathbf{B}_{ij}\mathbf{f}_{ij}\left(\mathbf{r}_m\right)$ is concave with respect to $\mathbf{f}_{ij}\left(\mathbf{r}_m\right)$, thus $p_{ij}\left(\mathbf{r}_m\right)$ can be globally upper-bounded by its first-order Taylor expansion at the point $\mathbf{r}_m^t$ as

$$\begin{aligned}p_{ij}\left(\mathbf{r}_m\right) &= \mathbf{f}_{ij}^H\left(\mathbf{r}_m\right)\mathbf{B}_{ij}\mathbf{f}_{ij}\left(\mathbf{r}_m\right)\\ &\le \mathbf{f}_{ij}^H\left(\mathbf{r}_m^t\right)\mathbf{B}_{ij}\mathbf{f}_{ij}\left(\mathbf{r}_m^t\right)\\ &\quad + 2\,\mathrm{Re}\left\{\mathbf{f}_{ij}^H\left(\mathbf{r}_m^t\right)\mathbf{B}_{ij}\left[\mathbf{f}_{ij}\left(\mathbf{r}_m\right) - \mathbf{f}_{ij}\left(\mathbf{r}_m^t\right)\right]\right\},\end{aligned} \tag{33}$$

where $\mathbf{r}_m^t \in \mathbb{R}^2$ is a constant vector denoting the local point in the t-th iteration of SCA. It should be noted that the inequality in (33) holds for any $\mathbf{r}_m$. By rearranging the terms in (33), we obtain an upper bound on $h_{ij}\left(\mathbf{r}_m\right)$ as

$$\begin{aligned}h_{ij}\left(\mathbf{r}_m\right) =& \mathbf{f}_{ij}^H\left(\mathbf{r}_m\right)\mathbf{A}_{ij}\mathbf{f}_{ij}\left(\mathbf{r}_m\right)\\ \le& \underbrace{2\,\mathrm{Re}\left\{\mathbf{f}_{ij}^H\left(\mathbf{r}_m^t\right)\mathbf{B}_{ij}\mathbf{f}_{ij}\left(\mathbf{r}_m\right)\right\}}_{g_{ij}\left(\mathbf{r}_m\right)}\\ &+ \underbrace{\lambda_{ij}^m\left(L_i^{\mathrm{r}} + L_j^{\mathrm{r}}\right) - \mathbf{f}_{ij}^H\left(\mathbf{r}_m^t\right)\mathbf{B}_{ij}\mathbf{f}_{ij}\left(\mathbf{r}_m^t\right)}_{\text{constant}},\end{aligned} \tag{34}$$

where $\mathbf{r}_m^t$ represents the value of $\mathbf{r}_m$ in the t-th iteration.

Exploiting the upper bound on $h_{ij}\left(\mathbf{r}_m\right)$ given in (34), minimizing problem (P3) can be converted to minimizing its upper bound. To this end, in the t-th iteration of SCA, the optimization problem (P3) is relaxed as

$$\begin{aligned}(\mathrm{P4})\quad \min_{\mathbf{r}_m}\ & \sum_{i\in\mathcal{K}}\sum_{j\in\mathcal{I}_i'}\mathrm{Re}\left\{\mathbf{f}_{ij}^H\left(\mathbf{r}_m^t\right)\mathbf{B}_{ij}\mathbf{f}_{ij}\left(\mathbf{r}_m\right)\right\}\\ s.t.\ & \text{(30b), (30c)}.\end{aligned} \tag{35}$$

Although $g_{ij}\left(\mathbf{r}_m\right) \triangleq \mathrm{Re}\left\{\mathbf{f}_{ij}^H\left(\mathbf{r}_m^t\right)\mathbf{B}_{ij}\mathbf{f}_{ij}\left(\mathbf{r}_m\right)\right\}$ is a linear function over $\mathbf{f}_{ij}\left(\mathbf{r}_m\right)$, it is still not convex over $\mathbf{r}_m$. Fortunately, the problem-solving methodology in [42] can be leveraged to construct a convex surrogate function that locally approximates $g_{ij}\left(\mathbf{r}_m\right)$.

Any summation term in (35) can be rewritten as $g_{ij}\left(\mathbf{r}_m\right) = \mathrm{Re}\left\{\mathbf{b}_{ij}^H\mathbf{f}_{ij}\left(\mathbf{r}_m\right)\right\}$ by defining a constant variable $\mathbf{b}_{ij} \in \mathbb{C}^{\left(L_i^{\mathrm{r}}+L_j^{\mathrm{r}}\right)\times 1}$ as $\mathbf{b}_{ij} \triangleq \mathbf{B}_{ij}\mathbf{f}_{ij}\left(\mathbf{r}_m^t\right)$. Then the gradient vector and Hessian matrix of $g_{ij}\left(\mathbf{r}_m\right)$ over $\mathbf{r}_m$, i.e., $\nabla g_{ij}\left(\mathbf{r}_m\right) \in \mathbb{R}^2$ and $\nabla^2 g_{ij}\left(\mathbf{r}_m\right) \in \mathbb{R}^{2\times 2}$, can be expressed in a closed form [42]. Specifically, recalling that $\mathbf{r}_m = \left[x_m, y_m\right]^T$, the gradient vector $\nabla g_{ij}\left(\mathbf{r}_m\right) = \left[\frac{\partial g_{ij}(\mathbf{r}_m)}{\partial x_m}, \frac{\partial g_{ij}(\mathbf{r}_m)}{\partial y_m}\right]^T$ is provided in (36) at the bottom of this page. In (36), we drop the subscript of $\mathbf{b}_{ij}$ and denote its l-th entry as $b_l = |b_l|\,e^{j\angle b_l}$, with $|b_l|$ and $\angle b_l$ being the amplitude and the phase, respectively.

By constructing a positive real number $\delta_{ij}^t = 8\pi^2\left\|\mathbf{b}_{ij}\right\|_1/\lambda^2$, it is guaranteed that $\nabla^2 g_{ij}\left(\mathbf{r}_m\right) \preceq \delta_{ij}^t\mathbf{I}$ for any $\mathbf{r}_m$ [42]. Thus $g_{ij}\left(\mathbf{r}_m\right)$ is a function with Lipschitz continuous gradient [43], which can be globally upper-bounded by a quadratic function as

$$\begin{aligned}g_{ij}\left(\mathbf{r}_m\right) \le& g_{ij}\left(\mathbf{r}_m^t\right) + \nabla g_{ij}\left(\mathbf{r}_m^t\right)^T\left(\mathbf{r}_m - \mathbf{r}_m^t\right)\\ &+ \frac{\delta_{ij}^t}{2}\left(\mathbf{r}_m - \mathbf{r}_m^t\right)^T\left(\mathbf{r}_m - \mathbf{r}_m^t\right)\\ =& \underbrace{\frac{\delta_{ij}^t}{2}\mathbf{r}_m^T\mathbf{r}_m + \left(\nabla g_{ij}\left(\mathbf{r}_m^t\right) - \delta_{ij}^t\mathbf{r}_m^t\right)^T\mathbf{r}_m}_{\tilde{g}_{ij}(\mathbf{r}_m)}\\ &+ \underbrace{g_{ij}\left(\mathbf{r}_m^t\right) + \left(\frac{\delta_{ij}^t}{2}\mathbf{r}_m^t - \nabla g_{ij}\left(\mathbf{r}_m^t\right)\right)^T\mathbf{r}_m^t}_{\text{constant}}.\end{aligned} \tag{37}$$

Therefore, minimizing (35) can be converted to minimizing $\sum_{i\in\mathcal{K}}\sum_{j\in\mathcal{I}_i'}\tilde{g}_{ij}\left(\mathbf{r}_m\right)$, where $\tilde{g}_{ij}\left(\mathbf{r}_m\right) \triangleq \frac{\delta_{ij}^t}{2}\mathbf{r}_m^T\mathbf{r}_m + \left(\nabla g_{ij}\left(\mathbf{r}_m^t\right) - \delta_{ij}^t\mathbf{r}_m^t\right)^T\mathbf{r}_m$. As a result, in the t-th iteration of SCA, problem (P4) is converted to

$$\begin{aligned}(\mathrm{P5})\quad \min_{\mathbf{r}_m}\ & \frac{\delta^t}{2}\mathbf{r}_m^T\mathbf{r}_m + \left(\nabla g\left(\mathbf{r}_m^t\right) - \delta^t\mathbf{r}_m^t\right)^T\mathbf{r}_m\\ s.t.\ & \text{(30b), (30c)},\end{aligned} \tag{38}$$

$$\begin{aligned}\frac{\partial g_{ij}\left(\mathbf{r}_m\right)}{\partial x_m} =& -\frac{2\pi}{\lambda}\sum_{l=1}^{L_i^{\mathrm{r}}}|b_l|\sin\theta_{i,l}\cos\phi_{i,l}\sin\left(\frac{2\pi}{\lambda}\rho_{i,l}\left(\mathbf{r}_m\right) - \angle b_l\right) - \frac{2\pi}{\lambda}\sum_{l=1}^{L_j^{\mathrm{r}}}\left|b_{L_i^{\mathrm{r}}+l}\right|\sin\theta_{j,l}\cos\phi_{j,l}\sin\left(\frac{2\pi}{\lambda}\rho_{j,l}\left(\mathbf{r}_m\right) - \angle b_{L_i^{\mathrm{r}}+l}\right),\\ \frac{\partial g_{ij}\left(\mathbf{r}_m\right)}{\partial y_m} =& -\frac{2\pi}{\lambda}\sum_{l=1}^{L_i^{\mathrm{r}}}|b_l|\cos\theta_{i,l}\sin\left(\frac{2\pi}{\lambda}\rho_{i,l}\left(\mathbf{r}_m\right) - \angle b_l\right) - \frac{2\pi}{\lambda}\sum_{l=1}^{L_j^{\mathrm{r}}}\left|b_{L_i^{\mathrm{r}}+l}\right|\cos\theta_{j,l}\sin\left(\frac{2\pi}{\lambda}\rho_{j,l}\left(\mathbf{r}_m\right) - \angle b_{L_i^{\mathrm{r}}+l}\right).\end{aligned} \tag{36}$$

where δ^t and $\nabla g\left(\mathbf{r}_m^t\right)$ are respectively obtained by

$$\delta^t = \sum_{i\in\mathcal{K}}\sum_{j\in\mathcal{I}_i'}\delta_{ij}^t = \frac{8\pi^2}{\lambda^2}\sum_{i\in\mathcal{K}}\sum_{j\in\mathcal{I}_i'}\left\|\mathbf{b}_{ij}\right\|_1, \tag{39}$$

and

$$\nabla g\left(\mathbf{r}_m^t\right) = \sum_{i\in\mathcal{K}}\sum_{j\in\mathcal{I}_i'}\nabla g_{ij}\left(\mathbf{r}_m^t\right). \tag{40}$$

Problem (P5) is a constrained quadratic programming problem with convex objective function. If the constraints are ignored, the closed-form global optimum can be obtained as

$$\mathbf{r}_{m,t+1}^\star = \mathbf{r}_m^t - \frac{1}{\delta^t}\nabla g\left(\mathbf{r}_m^t\right). \tag{41}$$

If $\mathbf{r}_{m,t+1}^\star$ satisfies constraints (30b) and (30c), it is the global optimum for (P5). Otherwise, to obtain a feasible solution to (P5), we convert the non-convex constraint (30c) to a convex constraint in the following.

Since $\left\|\mathbf{r}_m-\mathbf{r}_n\right\|_2$ is a convex function over $\mathbf{r}_m$, its first-order Taylor expansion at $\mathbf{r}_m^t$ provides a lower bound on $\left\|\mathbf{r}_m-\mathbf{r}_n\right\|_2$ as

$$\begin{aligned}\left\|\mathbf{r}_m-\mathbf{r}_n\right\|_2 \geq& \left\|\mathbf{r}_m^t-\boldsymbol{r}_n\right\|_2 + \frac{\left(\mathbf{r}_m^t-\boldsymbol{r}_n\right)^T}{\left\|\mathbf{r}_m^t-\boldsymbol{r}_n\right\|_2}\left(\mathbf{r}_m-\mathbf{r}_m^t\right)\\ =&\frac{1}{\left\|\mathbf{r}_m^t-\boldsymbol{r}_n\right\|_2}\left(\mathbf{r}_m^t-\boldsymbol{r}_n\right)^T\left(\mathbf{r}_m-\boldsymbol{r}_n\right).\end{aligned} \tag{42}$$

With (42), if $\mathbf{r}_{m,t+1}^\star$ in (41) is not a feasible solution to (P5), we can alternatively transform problem (P5) into

$$\begin{aligned}(\text{P6})\quad \min_{\mathbf{r}_m}\quad & \frac{\delta^t}{2}\mathbf{r}_m^T\mathbf{r}_m + \left(\nabla g\left(\mathbf{r}_m^t\right)-\delta^t\mathbf{r}_m^t\right)^T\mathbf{r}_m && \text{(43a)}\\ s.t.\quad & \frac{1}{\left\|\mathbf{r}_m^t-\boldsymbol{r}_n\right\|_2}\left(\mathbf{r}_m^t-\boldsymbol{r}_n\right)^T\left(\mathbf{r}_m-\boldsymbol{r}_n\right)\geq D,\\ & \qquad n=1,2,\ldots,M,\quad n\neq m, && \text{(43b)}\\ & \text{(30b)}.\end{aligned}$$

Problem (P6) involves convex quadratic programming (QP) and can be efficiently solved by using either CVX [44] or quadprog [45].

The proposed algorithm for solving problem (P5) is summarized in Algorithm 2. Therein, the relative decrease of the objective function in the i-th iteration is defined as $\left(f^{i-1}-f^i\right)/f^{i-1}$, where f^{i-1} and f^i denote the function value in the $(i-1)$-th iteration and the i-th iteration, respectively. The initial values of $\mathbf{r}$ can be obtained by generating random points that satisfy the constraints (30b) and (30c). To achieve a more refined solution, it is beneficial to set multiple random initial points and then select the most favorable solution from the corresponding pool of solutions. In our simulation experiments, setting ten initial points can yield satisfactory results.

Algorithm 2 Alternating Optimization for Problem (P2)

Input: M, K, λ, $\mathbf{r}$, $\{\pi_k\}$, $\{L_k^{\mathrm{r}}\}$, $\{\theta_{k,l}\}$, $\{\phi_{k,l}\}$, $\mathcal{C}_{\mathrm{r}}$, D, ϵ_1, ϵ_2.
Output: $\mathbf{r}$.

while Relative decrease of the objective function value in (P2) is above ϵ_1 **do**
  for $m=1\rightarrow M$ **do**
    Keep $\{\mathbf{r}_n, n\neq m\}_{n=1}^M$ fixed, initialize $t=0$ and $\mathbf{r}_m^0=\mathbf{r}_m$.
    Compute $\mathbf{B}_{ij}$ for each pair of pilot-sharing users via (32).
    while Relative decrease of the objective function value in (P3) is above ϵ_2 **do**
      Update δ^t via (39).
      Update $\nabla g_{ij}\left(\mathbf{r}_m\right)$ according to (36), compute $\nabla g\left(\mathbf{r}_m^t\right)$ via (40).
      Obtain $\mathbf{r}_{m,t+1}^\star$ via (41).
      if $\mathbf{r}_{m,t+1}^\star$ statisfies (30b) and (30c) **then**
        $\mathbf{r}_m^{t+1}=\mathbf{r}_{m,t+1}^\star$.
      else
        Obtain $\mathbf{r}_m^{t+1}$ by solving (P6).
      end if
    end while
    Update $\mathbf{r}_m\leftarrow\mathbf{r}_m^{t+1}$.
  end for
end while

D. Convergence and Complexity Analysis

Overall, the convergence of Algorithm 2 depends on the convergence of the SCA-based algorithm in the inner loop and the AO-based algorithm in the outer loop.

For the inner loop introduced in Section III-C, we denote the two constant terms in (34) and (37) by $\Gamma_{ij}^1\left(\mathbf{r}_m^t\right) = \lambda_{ij}^m\left(L_i^{\mathrm{r}}+L_j^{\mathrm{r}}\right) - \mathbf{f}_{ij}^H\left(\mathbf{r}_m^t\right)\mathbf{B}_{ij}\mathbf{f}_{ij}\left(\mathbf{r}_m^t\right)$ and $\Gamma_{ij}^2\left(\mathbf{r}_m^t\right) = g_{ij}\left(\mathbf{r}_m^t\right)+\left(\frac{\delta_{ij}^t}{2}\mathbf{r}_m^t-\nabla g_{ij}\left(\mathbf{r}_m^t\right)\right)^T\mathbf{r}_m^t$, respectively. Since the Taylor expansions in (34) and (37) are tight at $\mathbf{r}_m^t$, we obtain

$$h_{ij}\left(\mathbf{r}_m^t\right) = 2\left(\tilde{g}_{ij}\left(\mathbf{r}_m^t\right)+\Gamma_{ij}^2\left(\mathbf{r}_m^t\right)\right)+\Gamma_{ij}^1\left(\mathbf{r}_m^t\right). \tag{44}$$

Then, for the t-th iteration of SCA, the objective function in (30a), denoted as $h\left(\mathbf{r}_m\right)$, is non-decreasing due to

$$\begin{aligned}h\left(\mathbf{r}_m^t\right) =& \sum_{i\in\mathcal{K}}\sum_{j\in\mathcal{I}_i'}h_{ij}\left(\mathbf{r}_m^t\right)\\ \overset{(a)}{\geq}& \sum_{i\in\mathcal{K}}\sum_{j\in\mathcal{I}_i'}\left(2\left(\tilde{g}_{ij}\left(\mathbf{r}_m^{t+1}\right)+\Gamma_{ij}^2\left(\mathbf{r}_m^t\right)\right)+\Gamma_{ij}^1\left(\mathbf{r}_m^t\right)\right)\\ \overset{(b)}{\geq}& \sum_{i\in\mathcal{K}}\sum_{j\in\mathcal{I}_i'}\left(2g_{ij}\left(\mathbf{r}_m^{t+1}\right)+\Gamma_{ij}^1\left(\mathbf{r}_m^t\right)\right)\overset{(c)}{\geq} h\left(\mathbf{r}_m^{t+1}\right),\end{aligned} \tag{45}$$

where (a) holds because we have minimized the value of $\sum_{i\in\mathcal{K}}\sum_{j\in\mathcal{I}_i'}\tilde{g}_{ij}\left(\mathbf{r}_m\right)$ in (P5). (b) and (c) hold because $\tilde{g}_{ij}\left(\mathbf{r}_m\right)+\Gamma_{ij}^2\left(\mathbf{r}_m^t\right)$ and $2g_{ij}\left(\mathbf{r}_m\right)+\Gamma_{ij}^1\left(\mathbf{r}_m^t\right)$ are upper bounds on $g_{ij}\left(\mathbf{r}_m\right)$ and $h_{ij}\left(\mathbf{r}_m\right)$ at $\mathbf{r}_m^{t+1}$, respectively.

For the outer loop, the objective function value in (24) is also non-decreasing during the process of alternating optimization. Besides, since $\mathrm{Tr}\left\{\mathbf{R}_i^H\mathbf{R}_j\right\}\geq 0$ for any two positive semi-definite matrices $\mathbf{R}_i$ and $\mathbf{R}_j$, the objective function value is lower-bounded by zero. Thus, the proposed algorithm is guaranteed to converge.

The computational complexity of Algorithm 2 is analyzed as follows. We assume $L_k^{\mathrm{r}}=L_k^{\mathrm{t}},\forall k$ for simplicity. The number of user pairs that share identical pilot is estimated as

$P = \eta(\eta-1)/2 = K(K/\tau-1)/2$, assuming that users are evenly distributed among the pilot groups. In Step 4, we update $\{\mathbf{C}_{ij}\}$, $\{\mathbf{A}_{ij}\}$, and $\{\mathbf{B}_{ij}\}$ sequentially, the resulting computational complexity is $\mathcal{O}\left((M+\gamma_{\mathrm{p}})L_{\mathrm{r}}^2 P\right)$, where γ_{p} is the number of iterations to obtain λ_{ij}^m through power iteration [46]. From Steps 5 to Step 14, the corresponding complexity to update r_m through SCA is $\mathcal{O}\left(L_{\mathrm{r}}^2 P\gamma_{\mathrm{i}} + M^{1.5}\ln(1/\beta)\gamma_{\mathrm{qp}}\right)$, where γ_{i} represents the maximum numbers of inner iterations concerning Steps 5 to 14, γ_{qp} is the maximum number of iterations required to solve the quadratic programming problem, and β is the accuracy of the inner-point method [42]. Finally, supposing that the maximum number of outer iterations covering Steps 1-17 is γ_{o}, the total computational complexity is $\mathcal{O}\left(\left((M+\gamma_{\mathrm{p}}+\gamma_{\mathrm{i}})L_{\mathrm{r}}^2 P + M^{1.5}\ln(1/\beta)\gamma_{\mathrm{qp}}\right)M\gamma_{\mathrm{o}}\right)$.

IV. Simulation Results

In the simulation, we consider a single-cell scenario where $K = 10$ single-FPA users are served by the BS. The carrier frequency is set as $f_c = 7.5$ GHz so that the wavelength is $\lambda = 0.04$ m. The numbers of transmit and receive paths are set as $L_k^{\mathrm{r}} = L_k^{\mathrm{t}} = 10$. We assume that all the users are of equal distance from the BS so that $\sigma_{h,k}^2 = 1/L_k^{\mathrm{r}}, \forall k$. Users are uniformly distributed in a 120° sector, i.e., $\bar{\theta}_k$ and $\bar{\phi}_k$ are both uniformly distributed in the angle interval $[\pi/6, 5\pi/6]$. Considering that larger ASDs lead to a higher probability of AoA intervals' overlapping and indicate higher spatial correlation among users, we set the elevation and azimuth ASDs as $\sigma_\theta = \sigma_\phi = 20^\circ$ to depict a scenario with severe pilot contamination.

Unless otherwise specified, the remaining simulation parameters are set as follows. The length of the coherence block is $T_c = 20$, and a total of $\tau = 5$ orthogonal pilots are assigned to $K = 10$ users based on the SGPS algorithm proposed in [13]. For the FA scheme, the BS is equipped with M FAs, and the minimum distance between FAs is set as $D = \lambda/2$. The convergence thresholds for the relative decrease in Algorithm 2 are set as $\epsilon_1 = \epsilon_2 = 10^{-3}$.

Fig. 3 presents the convergence behavior of Algorithm 2. The curve shown in the figure is obtained by averaging results from 100 randomly initiated starting points. It is observed that the objective function value of (P2) monotonically decreases, and the relative decrease goes below 10^{-3} after 20 iterations, which demonstrates fast convergence. Moreover, we plot the intra-group correlation expressed by (15), which exhibits a similar decreasing trend to the objective function value, affirming the effectiveness of the bounds derived in (23). Specifically, the intra-group correlation is reduced by approximately 42.91% compared with the initial value.

In the subsequent part, the channel estimation and data transmission performances of the proposed FA schemes are compared with those of the conventional FPA schemes. All the following results are obtained from 50 independent cell realizations and averaged.

Fig. 4 demonstrates the influence of pilot assignment algorithms on the NMSE-CE performance. The compared schemes are summarized as follows:

- **Proposed M-FA**: the BS is equipped with M FAs, and Algorithm 2 is executed to minimize the intra-group correlation. The side length of the moving region is set as $A = 3\lambda$ for $M = 16$, and $A = 6\lambda$ for $M = 32$, respectively.
- **N-FPA**: the BS is equipped with a ULA composed of N FPAs spaced by $\lambda/2$.
- **Random**: the pilot assignment strategy that randomly assigns pilots to the users.
- **Max-τ-Cut**: the pilot assignment algorithm given in Algorithm 1.
- **SGPS**: the pilot assignment algorithm proposed in [13].

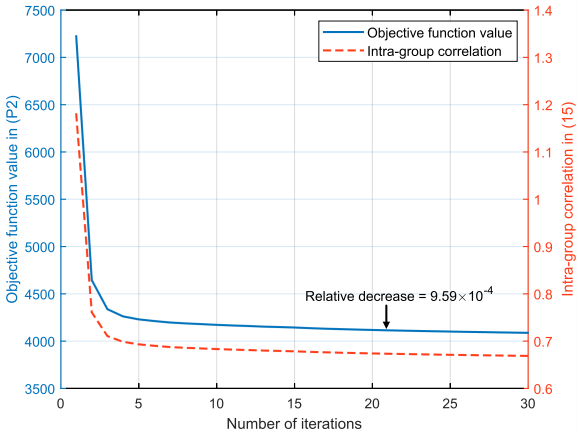


Fig. 3. Convergence behavior of Algorithm 2.

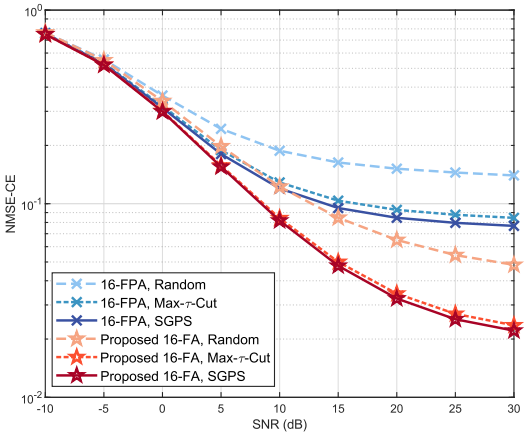
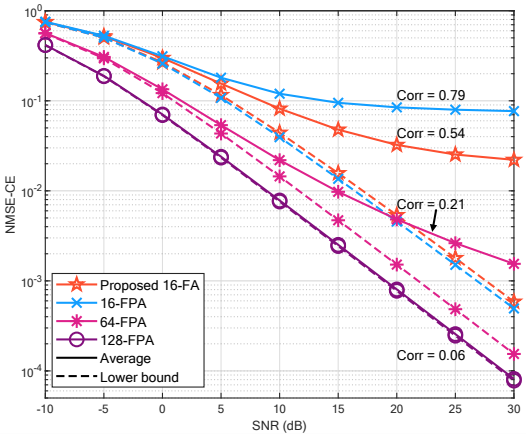


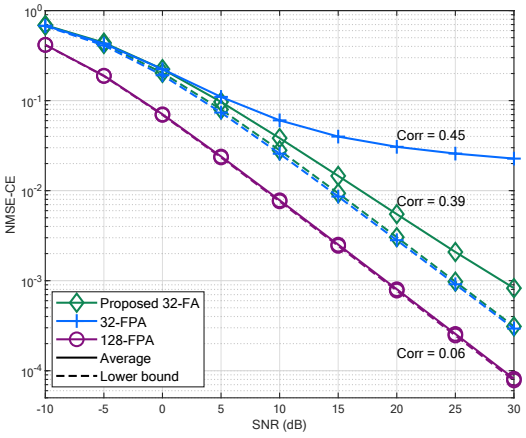
Fig. 4. NMSE-CE of different pilot assignment algorithms.

According to Fig. 4, we can conclude that for both FPA and FA schemes, an intelligent pilot allocation algorithm can yield considerable interference mitigation gain. It is worth noting that the FA scheme, even with random pilot assignment, can still provide better NMSE-CE performance than all FPA schemes at SNR levels above 10 dB, highlighting the impressive capability of fluid antennas in suppressing pilot interference. Furthermore, the SGPS algorithm always exhibits slightly superior performance to Algorithm 1, supporting the conclusion in Section II-D that the SGPS algorithm is an upgraded version of Algorithm 1.

Fig. 5 illustrates the NMSE-CE performances of different schemes for different SNR values, with the pilot assignment algorithm being SGPS. The intra-group correlation of each



(a) $M = 16$



(b) $M = 32$

Fig. 5. NMSE-CE of different schemes versus SNR.

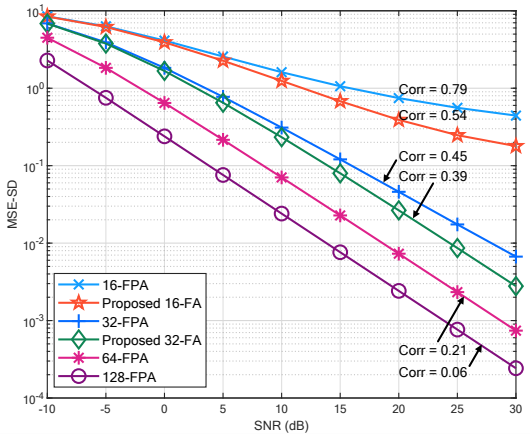


Fig. 6. MSE-SD of different schemes versus SNR.

scheme is also evaluated via (15) and noted next to the corresponding curve. The newly added legends are explained as follows:

- **Average**: the user-averaged NMSE-CE calculated through (13), and the results are further averaged over 50 independent cell realizations.
- **Lower bound**: the minimum NMSE-CE given in (14), which can be achieved when pilot contamination is completely nonexistent. The closer the NMSE curve is to the minimum NMSE curve, the weaker the effect of pilot contamination is.

From Fig. 5(a), we have the following observations: 1) in the high-SNR regime, the NMSE-CE of both 16-FPA and 16-FA schemes no longer decreases as SNR increases. This is because the predominant limiting factor for channel estimation performance at high SNR values is pilot contamination rather than the noise; 2) in contrast, the performance of the 128-FPA scheme nearly matches its lower bound over a wide SNR region. This confirms the correctness of the theory in [5] that the covariances of pilot-sharing users are asymptotically linearly independent as $M \to \infty$, despite the pilot reuse among the users; 3) the 16-FA scheme outperforms the 16-FPA scheme for all SNR values. Specifically, when SNR = 20 dB, the FA scheme obtains 61.7% NMSE-CE performance gain over the 16-FPA scheme with the same pilot assignment strategy; 4) reducing pilot contamination brings negligible improvement to NMSE-CE at low SNR levels. The reason lies in that the noise serves as the main factor affecting NMSE-CE in this regime; 5) the intra-group correlation metric is highly correlated with NMSE-CE performance, i.e., a lower intra-group correlation often suggests a smaller NMSE-CE, thereby making minimizing the intra-group correlation a reasonable strategy for mitigating pilot contamination.

Additionally, Fig. 5(b) presents the comparison of NMSE-CE performances between the 32-FA and 32-FPA schemes. The 32-FA scheme performs remarkably well over a wide SNR range, approaching the pilot contamination-free case. Compared to the 16-FA scheme, the 32-FA scheme's notable superiority over its FPA counterpart emphasizes the benefits of increasing FA numbers for interference suppression.

In Fig. 6, by exploiting the MMSE receiver [13] during the UL data transmission stage, we evaluate the performances of different schemes in terms of MSE for symbol detection (MSE-SD). The intra-group correlation of each scheme is also noted next to the corresponding curve. Herein, the MSE-SD is defined as $\epsilon^{\mathrm{SD}} = \mathbb{E}\left\{\|\hat{\mathbf{s}} - \mathbf{s}\|_2^2\right\}$, where $\mathbf{s}$ is the data symbols simultaneously sent from K users and $\hat{\mathbf{s}}$ denotes the detected symbols at the BS. At SNR = 20 dB, the proposed 16-FA and 32-FA schemes outperform their FPA counterparts by 2.81 dB and 2.39 dB, respectively. In general, the FA schemes always perform better than the FPA ones with the same antenna numbers, implying that mitigating pilot contamination can improve not only the accuracy of channel estimation but also the performance of data transmission.

Fig. 7 shows how the NMSE-CE of different schemes changes with τ, i.e., the pilot length (which is also the number of orthogonal pilot sequences). The SNR is fixed at 30 dB to exclude the influence of noise on NMSE-CE performance. The conclusions are drawn as follows: 1) a trade-off exists between NMSE-CE performance and pilot overhead. Under specified antenna configuration and SNR value, a lower level of NMSE-CE is achieved at the cost of increased pilot overhead, i.e., a greater pilot length; 2) when the pilot reuse factor $\eta = K/\tau$ is large, the performance gap between different schemes becomes

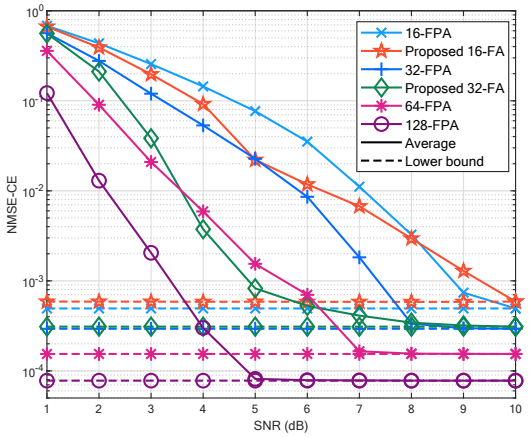


Fig. 7. NMSE-CE of different schemes versus pilot length τ, for SNR = 30 dB.

small. The reason is that the system's capability to suppress interference is exhausted when the same pilot is reused by quite a few users, such that even the FA scheme or the large-scale FPA scheme fails to prevent severe pilot contamination; 3) on the other hand, for small pilot reuse factors, the proposed 16-FA scheme loses to the 16-FPA scheme. Nonetheless, the pilot contamination is negligible in such a case, because the minimum MSE-CE condition in (11) is readily satisfied when only a few pilots are reused among users; 4) the proposed FA scheme provides remarkable interference mitigation gain for moderate pilot reuse factors around two. Specifically, with the reuse factor set to $\eta = 2$, the 16-FA and 32-FA schemes both obtain the highest NMSE-CE performance gains over their FPA baselines, with respective gains of 71.23% and 96.58%. It is noteworthy that at this point, the 32-FA scheme's NMSE-CE performance even surpasses that of the 64-FPA one. This suggests that the proposed FA scheme is particularly effective for addressing pilot contamination when there is a moderate level of pilot reuse with $K/\tau = 2$.

V. Conclusion

In this paper, we mitigated pilot decontamination for single-cell pilot reuse by establishing an FA position optimization problem that minimizes the novel objective function called intra-group correlation. We first derived analytical upper and lower bounds for the objective function to simplify the problem. Then a double-loop based algorithm was proposed to find a locally optimal solution. Specifically, the positions of M receive FAs at the BS were alternately optimized in the outer loop and each subproblem was solved in the inner loop utilizing the successive convex approximation technique. Numerical results showed the remarkable interference mitigation gains of the proposed FA schemes over the conventional FPA-based schemes.

References

[1] P. Yang, Y. Xiao, M. Xiao, and S. Li, "6G wireless communications: Vision and potential techniques," *IEEE Netw.*, vol. 33, no. 4, pp. 70–75, Jul./Aug. 2019.

[2] Y. Xiao, Z. Ye, M. Wu, H. Li, M. Xiao, M.-S. Alouini, A. Al-Hourani, and S. Cioni, "Space-air-ground integrated wireless networks for 6G: Basics, key technologies, and future trends," *IEEE J. Sel. Areas Commun.*, vol. 42, no. 12, pp. 3327–3354, Dec. 2024.

[3] W. Jiang, B. Han, M. A. Habibi, and H. D. Schotten, "The road towards 6G: A comprehensive survey," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 334–366, Feb. 2021.

[4] K. Shafique, B. A. Khawaja, F. Sabir, S. Qazi, and M. Mustaqim, "Internet of things (IoT) for next-generation smart systems: A review of current challenges, future trends and prospects for emerging 5G-IoT scenarios," *IEEE Access*, vol. 8, pp. 23 022–23 040, Jan. 2020.

[5] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO has unlimited capacity," *IEEE Trans. Wireless Commun.*, vol. 17, no. 1, pp. 574–590, Jan. 2018.

[6] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.

[7] Q. Spencer, C. Peel, A. Swindlehurst, and M. Haardt, "An introduction to the multi-user MIMO downlink," *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 60–67, Oct. 2004.

[8] E. Björnson, E. G. Larsson, and T. L. Marzetta, "Massive MIMO: Ten myths and one critical question," *IEEE Commun. Mag.*, vol. 54, no. 2, pp. 114–123, Feb. 2016.

[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] T. L. Marzetta, "How much training is required for multiuser MIMO?" in *Proc. 40th Asilomar Conf. Signals, Syst. Comput. (ACSSC)*, Pacific Grove, CA, USA, Oct./Nov. 2006, pp. 359–363.

[11] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "Pilot contamination problem in multi-cell TDD systems," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Seoul, South Korea, Jun./Jul. 2009, pp. 2184–2188.

[12] 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.

[13] L. You, X. Gao, X.-G. Xia, N. Ma, and Y. Peng, "Pilot reuse for massive MIMO transmission over spatially correlated rayleigh fading channels," *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3352–3366, Jun. 2015.

[14] L. Ribeiro, M. Leinonen, H. Djelouat, and M. Juntti, "Channel charting for pilot reuse in mMTC with spatially correlated MIMO channels," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Taiwan, Dec. 2020, pp. 1–6.

[15] L. Ribeiro, M. Leinonen, H. Al-Tous, O. Tirkkonen, and M. Juntti, "Exploiting spatial correlation for pilot reuse in single-cell mMTC," in *Proc. IEEE 32nd Annu. Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, Helsinki, Finland, Sep. 2021, pp. 654–659.

[16] ——, "Channel charting aided pilot reuse for massive MIMO systems with spatially correlated channels," *IEEE Open J. Commun. Soc.*, vol. 3, pp. 2390–2406, Dec. 2022.

[17] L. Sanguinetti, E. Björnson, and J. Hoydis, "Toward massive MIMO 2.0: Understanding spatial correlation, interference suppression, and pilot contamination," *IEEE Trans. Commun.*, vol. 68, no. 1, pp. 232–257, Jan. 2020.

[18] L. S. Muppirisetty, H. Wymeersch, J. Karout, and G. Fodor, "Location-aided pilot contamination elimination for massive MIMO systems," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, San Diego, CA, USA, Dec. 2015, pp. 1–5.

[19] N. Akbar, S. Yan, N. Yang, and J. Yuan, "Mitigating pilot contamination through location-aware pilot assignment in massive MIMO networks," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Washington, DC, USA, Dec. 2016, pp. 1–6.

[20] L. S. Muppirisetty, T. Charalambous, J. Karout, G. Fodor, and H. Wymeersch, "Location-aided pilot contamination avoidance for massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 4, pp. 2662–2674, Apr. 2018.

[21] H. Echigo, T. Ohtsuki, W. Jiang, and Y. Takatori, "Fair pilot assignment based on AOA and pathloss with location information in massive MIMO," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Singapore, Dec. 2017, pp. 1–6.

[22] P. Li, Y. Gao, Z. Li, and D. Yang, "User grouping and pilot allocation for spatially correlated massive MIMO systems," *IEEE Access*, vol. 6, pp. 47 959–47 968, Aug. 2018.

[23] L. You, M. Xiao, X. Song, Y. Liu, W. Wang, X. Gao, and G. Fettweis, "Pilot reuse for vehicle-to-vehicle underlay massive MIMO transmission," *IEEE Trans. Veh. Technol.*, vol. 69, no. 5, pp. 5693–5697, May 2020.

[24] 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.

[25] H. T. Dao and S. Kim, "Vertex graph-coloring-based pilot assignment with location-based channel estimation for massive MIMO systems," *IEEE Access*, vol. 6, pp. 4599–4607, Jan. 2018.

[26] S. E. Hajri, M. Assaad, and G. Caire, "Scheduling in massive MIMO: User clustering and pilot assignment," in *Proc. 54th Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Sep. 2016, pp. 107–114.

[27] W. Zeng, Y. He, B. Li, and S. Wang, "Pilot assignment for cell free massive MIMO systems using a weighted graphic framework," *IEEE Trans. Veh. Technol.*, vol. 70, no. 6, pp. 6190–6194, Jun. 2021.

[28] X. Gao, O. Edfors, F. Rusek, and F. Tufvesson, "Massive MIMO performance evaluation based on measured propagation data," *IEEE Trans. Wireless Commun.*, vol. 14, no. 7, pp. 3899–3911, Jul. 2015.

[29] K.-K. Wong, W. K. New, X. Hao, K.-F. Tong, and C.-B. Chae, "Fluid antenna system—Part I: Preliminaries," *IEEE Commun. Lett.*, vol. 27, no. 8, pp. 1919–1923, Aug. 2023.

[30] K.-K. Wong, A. Shojaeifard, K.-F. Tong, and Y. Zhang, "Fluid antenna systems," *IEEE Trans. Wireless Commun.*, vol. 20, no. 3, pp. 1950–1962, Mar. 2021.

[31] W. K. New, K.-K. Wong, H. Xu, K.-F. Tong, and C.-B. Chae, "An information-theoretic characterization of MIMO-FAS: Optimization, diversity-multiplexing tradeoff and q-outage capacity," *IEEE Trans. Wireless Commun.*, vol. 23, no. 6, pp. 5541–5556, Jun. 2024.

[32] K.-K. Wong and K.-F. Tong, "Fluid antenna multiple access," *IEEE Trans. Wireless Commun.*, vol. 21, no. 7, pp. 4801–4815, Jul. 2022.

[33] G. Hu, Q. Wu, K. Xu, J. Ouyang, J. Si, Y. Cai, and N. Al-Dhahir, "Fluid antennas-enabled multiuser uplink: A low-complexity gradient descent for total transmit power minimization," *IEEE Commun. Lett.*, vol. 28, no. 3, pp. 602–606, Mar. 2024.

[34] X. Pi, L. Zhu, Z. Xiao, and R. Zhang, "Multiuser communications with movable-antenna base station via antenna position optimization," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Kuala Lumpur, Malaysia, Dec. 2023, pp. 1386–1391.

[35] H. Qin, W. Chen, Z. Li, Q. Wu, N. Cheng, and F. Chen, "Antenna positioning and beamforming design for movable-antenna enabled multi-user downlink communications," 2023, arXiv:2311.03046.

[36] Y. Wu, D. Xu, D. W. K. Ng, W. Gerstacker, and R. Schober, "Movable antenna-enhanced multiuser communication: Jointly optimal discrete antenna positioning and beamforming," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Kuala Lumpur, Malaysia, Dec. 2023, pp. 7508–7513.

[37] Z. Cheng, N. Li, J. Zhu, X. She, C. Ouyang, and P. Chen, "Sum-rate maximization for fluid antenna enabled multiuser communications," *IEEE Commun. Lett.*, vol. 28, no. 5, pp. 1206–1210, May 2024.

[38] S. Sahni and T. Gonzalez, "P-complete approximation problems," *J. Assoc. Comput. Mach.*, vol. 23, no. 3, pp. 555–565, Jul. 1976.

[39] L. Zhu, W. Ma, and R. Zhang, "Modeling and performance analysis for movable antenna enabled wireless communications," *IEEE Trans. Wireless Commun.*, vol. 23, no. 6, pp. 6234–6250, Jun. 2024.

[40] M. Herdin, N. Czink, H. Ozcelik, and E. Bonek, "Correlation matrix distance, a meaningful measure for evaluation of non-stationary MIMO channels," in *Proc. IEEE 61st Veh. Technol. Conf.*, Stockholm, Sweden, May/Jun. 2005, pp. 136–140.

[41] H. Wolkowicz and G. P. H. Styan, "Bounds for eigenvalues using traces," *Linear Algebra Appl.*, vol. 29, pp. 471–506, Feb. 1980.

[42] W. Ma, L. Zhu, and R. Zhang, "MIMO capacity characterization for movable antenna systems," *IEEE Trans. Wireless Commun.*, vol. 23, no. 4, pp. 3392–3407, Apr. 2024.

[43] Y. Nesterov, *Lectures on convex optimization*. Berlin, Germany: Springer, 2018, vol. 137.

[44] M. Grant and S. Boyd. (2014) CVX: MATLAB Software for Disciplined Convex Programming, Version 2.1. [Online]. Available: https://cvxr.com/cvx

[45] B. A. Turlach and A. Weingessel. quadprog: Functions to Solve Quadratic Programming Problems. 2019, R package version 1.5-7. Accessed: Jul. 8, 2020. [Online]. Available: https://CRAN.R-project.org/package=quadprog

[46] B. N. Parlett and W. G. Poole, "A geometric theory for the QR, LU and power iterations," *SIAM J. Num. Anal.*, vol. 10, no. 2, pp. 389–412, Apr. 1973.